Probabilistic Assessment of User's Emotions During the Interaction with Educational Games

Cristina Conati
Department of Computer Science
University of British Columbia
2366 Main Mall
Vancouver, BC, V6T1Z4, CANADA,
conati@cs.ubc.ca

Abstract

We describe preliminary research on how to monitor a user's emotions and level of engagement during the interaction with educational games. We illustrate how the user's emotional state can be assessed through a probabilistic model that takes into account the context of the interaction, the user's personality and a variety of user's bodily expressions that are known to be directly influenced by emotional reactions. The probabilistic model relies on influence diagrams to flexibly leverage any evidence related to the user's emotional state, in order to estimate this state and any other related variable in the model. This is crucial in a modeling task in which the available evidence usually varies with the user and with each particular interaction. The probabilistic model we present is to be used by decision theoretic pedagogical agents to generate interventions aimed at achieving the best tradeoff between user's learning and engagement during the interaction with educational games .

Introduction

In recent years, there has been growing interest in studying how interactive computer systems can deal with emotional behavior, in addition to cognitive behavior. Much of this research has focused on investigating how software agents can display emotions. There has also been increasing attention to designing devices that allow an interactive system to monitor a user's emotional states by measuring direct evidence of emotional arousal coming from bodily expressions such has physiological responses, facial expression and intonation [17]. However, most of the work in this direction has so far focused on using a single bodily expression to recognize emotional states. This has been done either to recognize basic user's emotions in fairly constraining contexts, with the help of additional information on the task, the interaction and relevant user's traits [10] or to detect lower level measures of emotional reaction, such as the intensity and valence of emotional arousal (i.e., whether the emotion generates a positive or negative feeling) [1]. In this paper, we address the problem of how an interactive system can monitor the user's emotional state using multiple direct indicators of emotional arousal, in less constraining contexts in which a variety of emotional states can arise and cannot always be precisely diagnosed using additional information on the task, the interaction and the user.

We focus, in particular, on the interaction with pedagogical agents designed to improve the effectiveness of educational games. In the rest of the paper, we first describe why detecting emotions is important for educational games. We then illustrate how a Bayesian probabilistic user model can be used to integrate, in a principled way, different sources of ambiguous information on the user's emotional state. Finally, we briefly describe how the model's assessment can be used to devise decision-theoretic, socially intelligent pedagogical agents that can help users learn from educational games while maintaining a high level of engagement.

Socially Intelligent Agents for Educational Games

Several authors have suggested the potential of video and computer games as educational tools (e.g., [13, 20]). However, empirical studies have shown that, while educational games are usually highly engaging, they often do not trigger the constructive reasoning necessary for learning [4] [12].

For instance, studies performed by the EGEMS (Electronic Games for Education in Math and Science) project at the University of British Columbia have shown that the tested educational games were effective only when coupled with supporting classroom activities, such as related pencil and paper worksheets and group discussions. Without these supporting activities, despite enthusiastic game playing, the learning that these games generated was usually rather limited [12].

An explanation of these findings is that it is often possible to learn how to play an educational game effectively without necessarily reasoning about the target domain knowledge [4]. Insightful learning requires meta-cognitive skills that foster conscious reflection upon one's actions [3, 6], but reflective cognition is hard work. Possibly, for many students the high level of engagement triggered by the game activities acts as a distraction from reflective cognition, especially when the game is not integrated with external activities that help ground the game experience into the learning one. Also, educational games are usually highly exploratory in nature, and empirical studies on exploratory learning environments have shown that these environments tend to be effective only for those students that already possess the meta-cognitive skills necessary to learn from autonomous exploration [19] (such as self-monitoring, self-questioning and self-explanation [3]).

Given the above results, we believe that the effectiveness of educational games can be increased by providing them with the capability to:

- explicitly monitor how students interact with and learn from the games;
- generate tailored interventions to trigger constructive reasoning and reflection from those students that need them.

However, this must be done without interfering with the factors that make games fun and enjoyable, such as a feeling of control, curiosity, and challenge [13]. Thus, it is fundamental that the educational interventions be delivered within the spirit of the game, by agents that:

- are an integral part of the game plot;
- are capable of detecting user's emotions that can influence the user's engagement in the game;
- know how to take these emotions into account to maintain a good balance between engagement and learning.

In this paper, we focus on how these socially intelligent agents can assess relevant user's emotions during game playing and we briefly sketch how they can use this assessment to increase the effectiveness of educational games

Probabilistic assessment of relevant emotions during game playing

Emotional states can be detected because they often affect both visible bodily expressions, such as facial expressions, voice intonation, posture and gestures, as well as less observable ones, such as heart rate, blood pressure, skin conductance, color and temperature [17]. While often a single emotion affects multiple bodily expressions, several studies indicate that a single bodily expression is usually not sufficient to recognize a specific emotion. For instance, skin conductivity is a very good indicator of general level of arousal (i.e., the intensity of the emotion), but cannot identify the valence of the arousal [17]. Emotions with negative valence tend to increase heart rate more than emotions with positive valence [2], but heart rate provides little more information about specific emotions [7]. Also, which bodily expressions an emotion affects can depend on the intensity of the emotion, on the user's temperament and personality, as well as on the context in which the emotion is aroused.

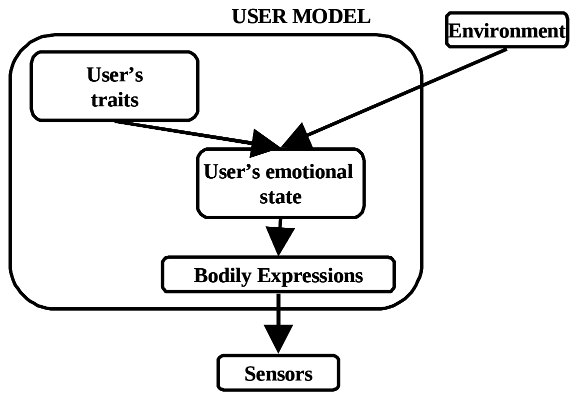

Figure 1:General Bayesian model to assess a user's emotional state

The above factors make emotion recognition a task frequently permeated with uncertainty, especially if the interaction context can induce a variety of different emotional states in different users, as it is the case in educational games. For instance, the exploratory nature of the game can be very exciting for people with a high level of curiosity and confidence, while it may cause frustration, shame, or fear in learners that are more passive and less confident. Also, a user's personality can strongly influence how the user reacts to a pedagogical intervention explicitly designed to help her learn more from the game.

To handle the high level of uncertainty involved in monitoring a user's emotional state during the interaction with an educational game, our approach is to explicitly model the probabilistic nature of the relations between emotional states, their causes and effects. We do so by using influence diagrams [8], an extension of the probabilistic reasoning framework of Bayesian networks [16] that can model variables representing an agent's choices in addition to random variables.

Figure 1 shows a high level description of our probabilistic model of emotions, in which only the general factors involved in assessing a user's emotional states are represented. The model indicates that the user's emotional states directly depend on the user's traits (e.g., personality traits as well as cognitive and meta-cognitive skills) and on the interaction with the environment (e.g., an educational game). The user's emotional states influence bodily expressions, which in turn influence the measurements taken with the available sensors. The advantage of having a model based on a Bayesian network is that it can leverage any evidence available on the variables related to emotional states to make predictions for any other variable in the model. So, for instance, the user's emotional state can be assessed using existing information on relevant user's personality traits and on the environment, even in absence of reliable sensors. Or, we can assess both emotional state and personality traits from reliable sensors and environment knowledge.

In the next section, we describe an example application on the above model in the context of Avalanche, the game we are developing as a test-bed for our research.

Figure 3: Prime Climb activity in Avalanche

Sample Emotion Model for Prime Climb

Avalanche is a multi-player game in which up to four players (grade 6 and 7 students) assume the roles of the leading citizens in a mountain ski town, and work through a series of mathematical puzzles to deal with the problems caused by a series of avalanches. Our goal is to devise socially intelligent agents that play different pedagogical roles in the game, designed to increase students' learning through reflection and effective collaborative behavior [5].

We are currently focusing on devising such agents for one of the Avalanche activities known as Prime Climb, which aims to help students learn issues related to number factorization. In Prime Climb, teams of two players must climb ice-faces by selecting numbers that do not share any factors with those occupied by other team members (see Figure 3), in order to identify possible sites for the next avalanche. Before climbing a mountain, the players need to get certified for that mountain climbing level, by doing practice climbs with an instructor known as "Cool Guy".

We conducted empirical evaluations of an early version of Prime Climb, in which Cool Guy provides a fixed sequence of instructions before a player starts climbing (see Figure 3, right side) and then directs the player's moves in the practice climbs to exemplify how legal moves can be computed through factor decomposition. These evaluations showed that students who do not know factor decomposition generally do not learn it from the practice climbs. This is evident from the fact that, when students start climbing in pairs, they often fall and cannot understand why. Although on-line help is available through a hypertextual game manual, students who tried to use the manual to understand how to avoid falling often failed to find the relevant information. Also, we noticed that many students did not even try to access the on-line help. They continued climbing by trial and error, exhaustively trying hexes until they made it to the top. To summarize, the empirical evaluations uncovered two main problems that limited the educational effectiveness of Prime Climb. First, the game fails to provide easy access to the information that may help students learn from it. Second, the game allows students to progress toward the game objectives without necessarily learning the target mathematical knowledge. In light of these findings, one of the agents we are designing to improve Prime Climb is a "smarter" Cool Guy that can provide more articulated, tailored help, both unsolicited and on demand, as a student is climbing, and that can do so without compromising the user's level of engagement in the game. We now show an illustrative example of how the general model in Figure 1 can be instantiated and used to allow Cool Guy to monitor a player's level of engagement and the emotional states that influence it.

Model structure

For the sake of simplicity, the illustrative model we present in this section includes only a subset of the variables that could be taken into account to assess the emotional reactions of a Prime Climb player. The goal is to give a sense of how the model is built and of its workings, but several more variables could be introduced to make it more accurate and comprehensive

Environment variables. The only environment variable that we consider in this example represents how Cool Guy can intervene when the student falls, with two values (1) suggest how to avoid falling; (2) do nothing (see *Environment* cluster in Figure 4). This is not a random variable, but rather a decision variable representing the rational decision that Cool Guy makes when selecting how to intervene in response to a student's error. Other relevant environment variables that could be included in this part of the model are, for instance, more information on the current state of the game and on the difficulty of the current activity.

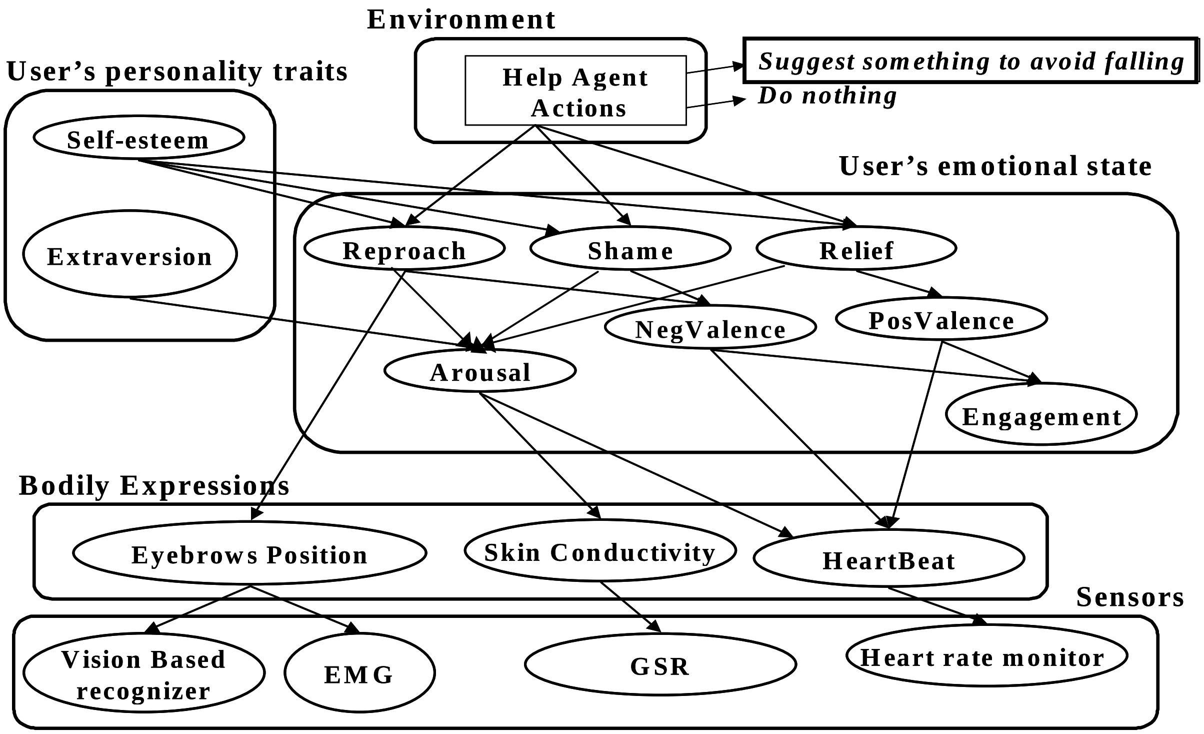

Figure 4: Sample model of emotional reaction

Variables describing the user's emotional state. As suggested by the Orthony, Clore and Collins (OCC) cognitive model of emotions [15], three of the possible emotional states that may arise as a consequence of the Cool Guy behavior, and that we include in our sample model are[2]: (i) reproach if the Cool Guy behavior interferes with a player's goal; (ii) *shame*, if the Cool Guy behavior makes the player self-conscious about her own errors; (iii) *relief* if the Cool Guy behavior helps the student overcome a problematic situation (see *user's emotional state* cluster in Figure 4)[3]. Because we are interested in assessing the student's level of *engagement* in the game, a corresponding variable is inserted into the model, along with links representing how this variable is influenced by the valence of a user's emotions (represented in Figure 4 by the deterministic nodes *PosValence* and *NegValence*). The corresponding conditional probabilities are defined to express the rather simplifying assumption that emotions with positive valence increase the level of engagement, while emotions with negative valence decrease it. In a more complete model, we may want to explicitly represent how specific emotions affect engagement. A node representing the level of *arousal* is also included in the model, because information on the level of arousal can be relevant to judge how much a given emotional state influences the user's behavior.

Variables describing the user's traits. For illustration, we consider only two user's traits in this example, describing the user's personality: *self-esteem* and *extraversion*. A complete model should include other relevant user's personality and cognitive traits, such as emotional stability and level of expertise in the task. As shown in Figure 4, self-esteem directly influences the emotional states that a Cool Guy action can induce. The conditional probability table (CPT) for the variable *reproach* (shown in Table 1[4]), for instance, models the fact that the player is likely to feel reproach if she has high self-esteem and Cool Guy provides help after a fall, because the agent's intervention may interfere with the student's beliefs that she can discover by herself how to

[2] Our model currently does not implement any of the appraisal mechanisms described in the OCC theory.

[3] This is of course only an illustrative subset of the emotional states that a user might experience when interacting with Cool Guy.

[4] All the conditional probabilities in the model are based on the author's estimates, derived from the qualitative relationships described in the literature.

play adequately. On the other hand, if the player has low self-esteem, the agent's provision of help can make the player self-conscious about his bad move, thus generating shame. There is also a chance that the same agent action could generate relief, because it allows a low self esteem player to recover from a situation that he might feel unable to handle alone.

Table 1: CPT for the variable *reproach*

AgentAction	GiveHelp		DoNothing	
SelfEsteem	High	Low	High	Low
T	0.9	0.1	0.2	0.05
F	0.1	0.9	0.8	0.95

The second user's personality trait variable, *extraversion*, directly influences level of arousal. The CPT for *arousal* encodes the finding that introverts are known to reach a higher level of arousal than extroverts , given the same stimuli [11].

Variables describing bodily expressions and sensors. Let's suppose that we have sensors to detect three types of bodily expressions: (i) *eyebrow position*, by using, for instance, vision software to detect facial expression and electromyogram (EMG, a way to detect muscle contraction); (ii) *skin conductivity*, through a sensor that detects galvanic skin response (GSR); (iii) *heart rate*, through a heart rate monitor. All these sensors can already be donned in a fairly non intrusive manner [17], and considerable research is being devoted to make these kinds of devices increasingly wearable. Each bodily expression B is linked to each sensor S that can detect it, as shown in Figure 4, and if multiple sensors are available, the Bayesian network propagation algorithms can automatically integrate evidence data coming from all of them. By encoding the probability of a sensor's value S given each value of bodily expression B, the conditional probability P(S|B) specifies the reliability of each sensor. Because this measure can be independently specified for each sensor and for the bodily expression that it detects, the model allows one to easily include new sensors as they become available.

Likewise, each conditional probability $P(B|E_1,..,E_n)$, indicates how a set of emotional states $E_1,..,E_n$ affects a given bodily expression B. As information on a bodily expression not yet considered in the model becomes available, a new variable for this expression can be added to the model and linked to the emotion variables that influence it, thus increasing the amount of evidence that can be used to detect the corresponding emotions. The conditional probabilities linking emotions and bodily expressions in our sample model represent the following findings [17]:

1. Frowning eyebrows are a very good indicator of negative emotions in the anger range, including reproach[5].
2. Skin conductivity is a very good indicator of the level of arousal.
3. Heartbeat increases more in the presence of emotions with negative valence (see CPT in Table 3).

Table 3: CPT for heart rate

NegativeValence	T					
PositiveValence	T			F		
arousal	normal	medium	high	normal	medium	high
Normal	0.8	0	0	0.8	0.05	0
High	0.05	0.2	0.9	0.1	0.15	0.8
Medium	0.15	0.8	0.1	0.1	0.8	0.2

NegativeValence	F					
PositiveValence	T			F		
arousal	normal	medium	high	normal	medium	high
Normal	1	0.4	0.1	1	0.2	0.2
High	0	0.05	0.3	0	0.4	0.4
Medium	0	0.55	0.6	0	0.4	0.4

Sample assessment

As we mentioned earlier, Bayesian networks and influence diagrams provide a flexible framework for reasoning under uncertainty. Given evidence on any subset of the random variables in a model, Bayesian propagation algorithms compute the conditional probability of any remaining random variables. We now give an example of how this process allows our model in Figure 4 to incrementally refine the assessment on the user's emotional state as more relevant user data become available, thus providing Cool Guy with increasingly accurate information to decide how to act to help the user interact best with the game.

Let's suppose that, at some point during the interaction with Prime Climb, the player falls and the agent decides to provide help. Let's also suppose that the only sensor signal available at this time comes from the heart rate monitor and indicates high heart rate. When this evidence is inserted in the model in Figure 4 and propagated, it increases the probability that the player's heart rate is high. High heart beat in turn increases the probability that the player is in an emotional state with negative rather than positive valence (because of the conditional probability representing finding 3 in the previous section). Although the available evidence cannot discriminate between the player feeling reproach or shame, high probability of negative valence is sufficient to raise the probability that the player's engagement is low. This can already be used by the agent to decide that

[5] Other kinds of facial expressions are generally good indicators of valence, if not of individual emotions. In our sample model, eyebrow position contributes indirect information on valence through the *reproach* variable.

its next action should focus on bringing engagement up again.

Let's now suppose that, in addition to high heart rate, we also detect high GSR. When propagated in the model, this evidence increases the probability of a high level of arousal (because of the conditional probability representing finding 2 in the previous section), and consequently the probability that our player is an introvert (because of the CPT in). This assessment provides the agent with additional information, that can be used to decide how to bring the player's level of engagement up, provided that the agent is given knowledge on how to deal with introverted vs. extroverted players.

Lastly, if our sensors also detect that the user is frowning, the probability of the player feeling reproach rather than shame increases (because of the conditional probability representing finding 1), providing further information for the agent to decide what to do to make up for its previous action. Indication that the player feels reproach also increases the probability that the player has high rather than low self esteem (because of the CPT in Table 1); this information can also be used to select the agent's next action. For instance, to revive the player's engagement, the agent could point out all the remaining challenges that the player needs to face. Had the evidence suggested a low-self-esteem player feeling shame, the next agent's action could point out to the player how other people make the same error he committed.

The model would have given a high probability to the user feeling reproach even if, instead of having evidence about the user frowning, it had evidence about the user having high self-esteem, collected, for instance, from the user's data available before the interaction with the game.

If contradictory evidence arises, such as knowledge that the player has low self-esteem but frowns upon provision of help from the agent, the model assessment will depend on the relative strength assigned to the different kinds of evidence by the model CPTs. However, in general the model probabilities will reflect a higher level of uncertainty on the user's emotional state, which also represents valuable information that the agent can use to decide how to act.

Model specification

One of the major difficulties in using Bayesian networks is defining the required prior and conditional probabilities. In the model in Figure 4, the only prior probabilities to be specified are those for variables representing user traits, which can be defined through existing statistics, specialized tests and stereotypes, or set to indicate lack of specific information. The conditional probabilities for the model have been defined by the author to encode the general qualitative information available in the literature, and can be refined for our particular application and user population (students in grade 6 and 7) through empirical evaluations.

We plan to continue combining the initial model specification based on existing findings and sensors with empirical model calibration, in order to include in our model additional personality traits, emotional states and bodily reactions that will provide a more complete model of user's emotional arousal for this particular type of interaction.

Using the model to decide how to act

As we mentioned in the previous section, the model we just described is based on an influence diagram, a Bayesian network augmented with nodes representing deterministic choices. In addition to computing the probability distribution over a model's random variables, influence diagrams compute the expected utility of each choice in a decision node. Thus, they allow implementing a decision-theoretic model of rational agents that always choose actions with maximum expected utility [18]. A decision theoretic approach has already been adopted to devise intelligent desktop assistants [9] and a computer tutor for coached problem solving [14]. We plan to implement our pedagogical agents for educational games as decision-theoretic agents that act to maximize the tradeoff between a user's learning and engagement. To do this, we need to extend the model in to include the player's relevant cognitive states (e.,g., game and math learning) and collaborative behaviors, as well as information about the temporal evolution of the interaction. By using influence diagrams, we can then compactly specify how the actions available to the different agents influence various elements in the student model. The utility function representing an agent's preferences can be defined in terms of these states, thus providing the agent with a normative theory of how to intervene in the students' game playing to achieve the best trade-off between engagement and learning. An agent's utility function will depend on the agent's role in the game. So, for instance, a help agent (e.,g., Cool Guy) will act to maximize the student's understanding of a specific activity (e.g., Prime Climb), while an agent in charge of eliciting effective collaboration will select actions that maximize this specific behavior.

Conclusions and Future Work

We have described a probabilistic model of a user's emotional reactions that integrates information on the possible causes of emotional arousal (e.g., stimuli from the environment and user's traits) as well as the behavioral effects of this arousal. The model is to be refined and extended to include additional cognitive and behavioral variables relevant to provide tailored support as a user interacts with educational computer games. This model's assessment will serve as input to intelligent pedagogical agents that use a decision-theoretic approach to decide how to intervene to help the user learn from the game while maintaining a high level of engagement.

We have collected videotapes of users playing an early version of Prime Climb, one of the educational activities we are developing to help students learn number factorization. We plan to use these videotapes, along with newly designed Wizard of OZ experiments, to identify common emotional reactions and the corresponding behavioral expressions that arise when users play with the original version of Prime Climb and with the version that includes the more intelligent help agent sketched in this paper. We will then decide which sensors to use to detect the relevant behaviors, and we will use this information to revise and complete the model of emotional reaction described in this paper.

References

1. Ball, G. and J. Breese. Modeling the emotional state of computer users. in *Workshop on 'Attitude, Personality and Emotions in User-Adapted Interaction', UM '99*. 1999.
2. Cacioppo, J.T., G.C. Berntson, K.M. Poehlmann, and T.A. Ito. The psychophysiology of Emotions, in *Handbook of Emotions (2nd edition)*, M. Lewis and J.M. Haviland-Jones, Editors. 2000, The Guilford Press.
3. Chi, M.T.H. Self-Explaining: The dual process of generating inferences and repairing mental models. *Advances in Instructional Psychology*, in press.
4. Conati, C. and J. Fain Lehman. EFH-Soar: Modeling education in highly interactive microworlds, in *Lecture Notes in Artificial Intelligence*. 1993, Springer-Verlag: New York.
5. Conati, C. and M. Klawe. Socially intelligent agents to improve the effectivess of educational games. in *AAAI Fall Symposium on Socially Intelligent Agents - The Human in the Loop*. 2000. Falmouth, MS, USA.
6. Conati, C. and K. VanLehn. Toward Computer-Based Support of Meta-Cognitive Skills: a Computational Framework to Coach Self-Explanation. *International Journal of Artificial Intelligence in Education*, 2000. 11.
7. Ekman, P., R.V. Levenson, and V.W. Friesen. Autonomic nervous system activity distinguishes among emotions. *Science*, 1983. 221.
8. Henrion, M., J. Breeze, and E. Horvitz. Decision Analysis and Expert Systems. AI Magazine, 1991.
9. Horvitz, E. Principles of mixed initiative interaction. in *CHI '99, ACM SIGCHI Conference on Human Factors in Computing Systems*. 1999. Pittsburgh, PA.
10. Hudlicka, E. Increasing SIA architecture realism by modeling and adapting to affect and personality. in *Fall AAAI Symposium on Socially Intelligent Agents: the Human in the Loop*. 2000. North Falmouth, MS, USA.
11. Kahneman, D. Arousal and attention, in *Attention and effort*. 1973, PrenticeHall: Englewood Cliffs, N.J. p. 2-49.
12. Klawe, M. When Does The Use Of Computer Games And Other Interactive Multimedia Software Help Students Learn Mathematics? in *NCTM Standards 2000 Technology Conference*. 1998. Arlington, VA.
13. Malone, T.W. and M.R. Lepper. Making learning fun: A taxonomy of intrinsic motivations for learning, in *Aptitude, learning and instruction: Volume III Conative and affective process analyses*, R.E. Snow and M.J. Farr, Editors. 1987, Erlbaum: Hillsdale, NJ.
14. Murray, C. and K. VanLehn. DT Tutor: A decision-theoretic dynamic apprach for optimal selection of tutorial actions. in *ITS 2000*. 2000. Montreal, Canada.
15. Ortony, A., G.L. Clore, and A.Collins. The cognitive structure of emotions. 1988: Cambridge University Press.
16. Pearl, J. Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. 1988, San Mateo, CA: Morgan-Kaufmann.
17. Picard, R. Affective Computing, MIT Press. 1997.
18. Russell, S. and P. Norvig. Artificial Intelligence: A Modern Approach. 1995, Los Altos, CA: Morgan-Kaufman.
19. Shute, V.J.. A comparison of learning environments: All that glitters..., in *Computers as Cognitive Tools*, S.P.L.a.S.J. Derry, Editor. 1993, Lawrence Erlbaum Associates: Hillsdale, NJ.
20. Silvern, S.B. Classroom Use of Videogames. *Educational Research Quarterly*, 1986.