Learning and Emotional Intelligence in Agents

Magy Seif El-Nasr
Assistant Lecturer & Graduate Student
Texas A&M University
Computer Science Department
College Station, TX 77844-3112
magys@cs.tamu.edu

Thomas R. Ioerger
Assistant Professor
Texas A&M University
Computer Science Department
College Station, TX 77844-3112
ioerger@cs.tamu.edu

John Yen
Professor
Texas A&M University
Computer Science Department
College Station, TX 77844-3112
yen@cs.tamu.edu

Abstract

Neurological evidence was uncovered by A. Demasio revealing the existence of 'emotional intelligence' and its importance. Following this breakthrough many computational models of emotions were developed. Although, psychological research on emotions recognized memory and experience as the main factors that define and shape the complexity of the emotional process, existing computational models of emotion did not incorporate experience or learning. We are proposing a model of an agent named PETEEI - a PET with Evolving Emotional Intelligence. PETEEI was modeled to produce emotions according to its own experience. Furthermore, PETEEI filters and expresses emotions according to its own moods and previous experience.

Introduction

The emergence of the 'emotional intelligence' concept answered the question that many researchers posed on the importance of simulating emotions. A. Demasio (Demasio 1994) has uncovered some neurological evidence, which emphasized the role that emotions play in the human decision-making process. Not long after, AI researchers, who were searching for a more accurate model of human intelligence, started modeling emotions (Shiida 1989, Velasquez 1997, Shibata, Ohkawa, and Tanie 1996, Reilly 1996, Masuyama 1994). Researchers modeled emotions for many reasons. Some were searching for a complete model of human intelligence, some were searching for an ultimate simulation of human intelligence; some others, such as Shibita et al., were trying to use emotions to improve their systems. Shibata used frustration as a fitness function to help robots cooperate in performing certain tasks (Shibita, Ohkawa, and Tanie 1996).

Research on emotion defined it as a complex process that dynamically changes through personality, experience and time (Goleman 1995). This dynamic nature of emotions was not incorporated in any of the computational models of emotion. To introduce our model we will first review some of the previous computational models of emotion and point out where these models fail to capture the complexity of the emotional process. We will then outline an alternative technique that we have developed.

Research on emotion has taken many alternative paths. Each path concentrated on simulating or understanding only one element in the big world of emotions. Some models were developed to understand and develop the link between events and emotions; these models were called event appraisal models (Ortony, Clore, and Collins 1988, Rosman, Jose, and Spindel 1990). Other models developed a pure physiological simulation of emotions, where they described each emotion in terms of its physiological reactions (Picard 1997). A few others modeled the interactions between emotions, including fear and pain (Bolles and Fanselow 1980).

Event appraisal models, such as Ortony et al.'s model (Ortony, Clore, and Collins 1988), relied on variables that affect the synthesis of emotions, including expectations, effort and predictions of the occurrence of events. For example, in Ortony's model joy was defined as the occurrence of a desirable event. In contrast, Rosman et al. formalized emotions according to event categories (Rosman, Jose, and Spindel 1990). These categories depend on many factors, including the likelihood of events to occur and the consistency measure of an event in terms of goals. In addition, they further refined the model to include self-perception. The idea that a subject may perceive him/herself as weak or strong in terms of an event may trigger different emotions.

These two models show how two or three variables could effectively trigger different emotions. However, they only consider a subset of the emotional process. For example, it is clear that at a given moment different emotions may be triggered with different intensities. However, it is unclear how other emotions or other motivational states, including thirst, hunger and pain, may affect emotions. These models are also limited to conscious emotions; subconscious emotions were not discussed. Additionally, since these models were not really developed for simulating emotions, they did not provide a way to compute the internal states or variables used in their models. For example, emotions, such as hope or fear, depend on the events expected and the subject's certainty of their occurrence. If we are going to simulate these emotions, we need to find a way to calculate variables, such as expectations.

Looking at the interactions between emotions and how several emotions may inhibit or empower others, Bolles and Fanselow (Bolles and Fanselow 1980) developed a model of two motivational states, namely, 'fear' and 'pain.' According to their findings, fear inhibits pain in some situations and pain inhibits fear in some others. Even though the model was small and only modeled two motivational states, it represents a very important and missing element in modeling emotions; emotions and motivational states do affect each other in various situations. Thus, for example pain may inhibit other emotional states such as pride, shame, and fear.

Due to the emergence of believable and social agents, a number of computational models of emotions have been proposed within the agent's community. We will review one of these models, namely, the OZ project (Bates 1992, Reilly 1996) at CMU. The OZ project simulated believable emotional and social agents. Each agent initially has preset attitudes towards certain objects in the environment. Furthermore, each agent has some initial goals and a set of strategies that it can follow to achieve each goal. The agent perceives an event in the environment. This event is then evaluated according to the agent's goals, standards and attitudes. After an event is evaluated, special rules are used to produce an emotion with a specific intensity. These rules are based on Ortony et al.'s model (Ortony, Clore, and Collins 1988). The emotions triggered are then mapped, according to their intensity, to a specific behavior (Reilly 1996).

Even though the model has many strong points, it still has some weaknesses. The emotion triggered is only triggered when the event perceived affects the standards, attitudes or is associated with a goal's success or failure. In other words, emotions such as hope, fear, disappointment and relief were not simulated in their model since these emotions rely on expectations. Realistically, an event does not normally cause a goal to fully succeed or fully fail. The idea of partial success and failure is an important factor that was missing from the model.

Many other models were developed to tackle different elements in the emotional process; for example, Simon at MIT (Velasquez 1997) and Picard's work on affective wearables (Picard 1997). In our work, we tried to take a more general perspective and link these different approaches and elements of the emotional process together. For example, we linked the work on emotion interactions with the work on event appraisals. In doing so, we get a more complete and realistic picture of emotions. Furthermore, we have looked at the link between emotions and experience. In the following section, we will present the contributions of our model in light of previous models.

Proposed Model

Model Architecture

The proposed model consists of three major components: the learning component, the emotional component and the decision-making component. Figure 1 shows the architecture of the model. It should be noted that there might be some missing links, such as, the link from the emotional process to the learning process and the link from the environment to the emotional process (drugs, etc.). However, since we chose to focus on the emotional component and the use of experience in the emotional process, we chose to ignore those factors and any others that may influence the emotional process.

The agent first perceives certain events from the environment, as depicted from the right hand side of the figure. These events are then passed to both the emotional component and the learning component (on the left-hand side of the figure). The learning component keeps track of the history of events that occurred and the consequent change in the environment. Accordingly, the learning component computes internal variables, such as expectations and the impact of events perceived. The emotional component uses information about the perceived event and some of the learning component outcomes, including events expected and event-goal associations, to produce an emotional behavior. The emotional behavior is then returned back to the decision-making component to make a decision on what action to take. The decision is made according to the situation and the emotional behavior.

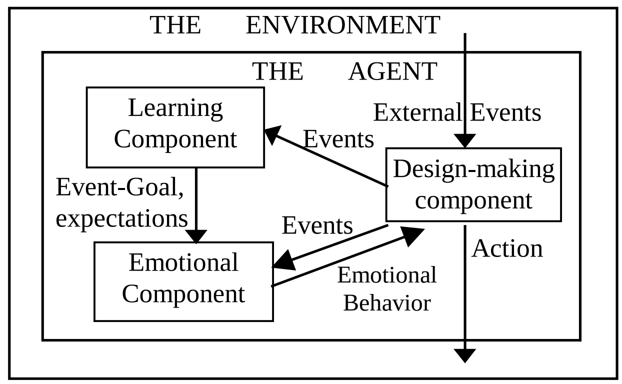

Figure 1. System's Architecture

Emotional Component

The emotional component is shown in more detail in Figure 2. Boxes represent different processes within the model. The perceived event taken from the decision-making component is first evaluated. The evaluation process consists of two sequential steps. Firstly, the experience model determines which goals are affected by the event and the degree of its impact. Secondly, desirability level of the event is computed according to the measure calculated by the first step and the importance of the goals involved. In other words, the event desirability measure is calculated from two major factors: the impact of the event on the agent's goals and the importance of these goals at the given time. Fuzzy rules are used to infer the desirability measure of an event according to these two criteria.

The desirability measure, once calculated, will be passed to an event appraisal model to further determine the emotional state of the agent. The event appraisal model is based on Ortony et al.'s model (Ortony, Clore, and Collins 1988). An emotion or a mixture of emotions will be triggered using the event desirability measure and the expectation measure, which is produced from the learning component. For example, hope was defined as the occurrence of an unconfirmed desirable event, and joy was defined as the occurrence of a desirable event. Thus, to trigger an emotion, such as hope, we need the event desirability measure and the expectation measure. The intensity of hope will then be calculated as a function of these two measures.

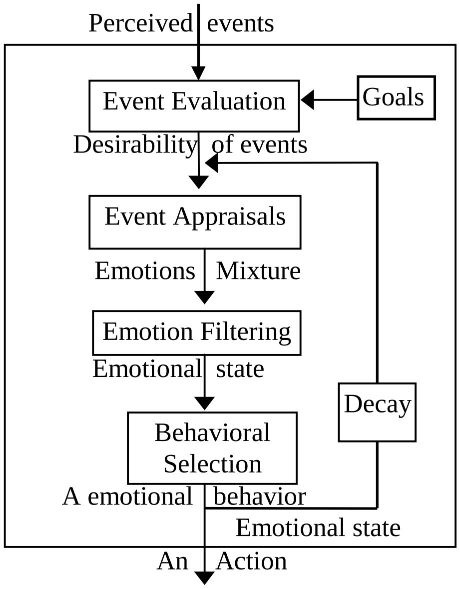

Figure 2. The Emotional Process

The mixture of emotions will be filtered to produce an emotional state. The filtering process uses a variation of Bolles and Fanslow's (Bolles and Fanselow 1980) model to produce a coherent mixture of emotions. In essence, the emotional state will be a list of emotions that apply at a specific time. The emotional state is passed to a behavioral selection phase. Through this phase, a behavior is chosen according to the situation and the emotional state. The behavior selection process is simulated using fuzzy rules. The emotional state will be decayed and fed back to the system for the next iteration. Additionally, there are other paths by which a behavior is produced. Some events or objects may trigger a conditioned behavior (LeDoux 1996), as we shall discuss in the learning section.

In the next couple of paragraphs, we will explore the learning process and its importance in simulating the emotional process. For more details on the filtering technique used and the use of fuzzy logic in the event evaluation and behavioral selection processes, the reader is referred to (El-Nasr 1998, El-Nasr and Yen 1998).

Learning Process

Learning has a major impact on the emotional process. The importance of learning will become very clear when we detail the learning process to explore the various ways that learning can change the emotional and behavioral state. We modeled four types of learning: (1) learning about events and likelihood of events to occur at any given situation, (2) learning about the user, (3) learning about actions and categorizing them as good or bad, and (4) conditioning, where a specific object is associated with a certain motivational state or emotion.

Learning Expectations

At any given time, an agent will need to know what event to expect, how certain it is that this event will happen and how desirable the event is. This information is crucial for the emotional process. It was noted in (Ortony, Clore, and Collins 1988, Rosman, Jose, and Spindel 1990) that the identification of emotions and emotional intensities depend heavily on expectations.

Furthermore, desirability of events are normally measured by their impact on a set of goals. It is often the case that a given event does not have any impact on any specific goal directly, but a sequence of events may eventually have an impact on some goals. For example, consider a goal of being rich, an event of being fired might not affect the goal directly. However, if you are fired, you will no longer receive your salary, and thus you will not get rich. Therefore, the problem is no longer a simple one to one link between events and the goals they affect, but rather we need to identify the link between an event or a sequence of events and the corresponding goals affected. Identifying that link was recognized as a very complex task to accomplish in the past (Reilly 1996). We knew that learning could potentially solve this problem. Our main idea is that the agent can potentially learn to link events or sequence of events with the goals affected by using reinforcement learning (Kaelbling, Littman, and Moore 1996, Mitchell 1996); we will briefly outline one of the reinforcement algorithms, namely the Q-learning algorithm. The reader is referred back to (Kaelbling, Littman, and Moore 1996, Mitchell 1996) for more details.

The agent represents the problem space using a table of Q-values in which each entry corresponds to a state-action pair. The table can be initially filled with default initial values. The agent will begin from a state *s*. It will take an action, *a*, which takes it to a new state s'. The agent may obtain a reward *r* for its action (Mitchell 1996). If it receives a reward it updates its table. The Q-value of the previous state-action pair depends on the Q-value of the new state-action pair. Thus, the Q-value is updated according to (1) the expected probability of the reward given the action taken and (2) a discounted product of the maximum value of the Q-values of the immediate previous states-action pair and its probability.

This algorithm is guaranteed to converge only for deterministic Markov Decision Processes (MDP). However, in our application there is a one-to-one alternation between the agent and the user. It is almost impossible to predict the reward given a state and an action pair, because it may change according to the user and environment. For example, at a state, s_0, the agent may take an action, *x*, and as a consequence the user gives him a positive reward. At a later time step, the agent is in the same state, s_0, so he takes action, *x*, thinking it is the best action to take, but this time the user may reward him negatively. Therefore, the user introduces a non-deterministic response that will affect the agent's reward. Thus, we treated the reward as a probability distribution over outcomes based on the state-action pair. An example is shown in Figure 3. The agent starts off in state, s_0, and takes an action, a_0, then the user can take either action, a_1, that puts the agent in state, s_1, or action, a_2, that puts the agent in state, s_2. The dotted lines show the non-determinism induced by user's actions, while the straight black lines shows the state-action transition as represented in the Q-table. The probability of the user's actions is calculated by using the user action patterns detailed in the next section.

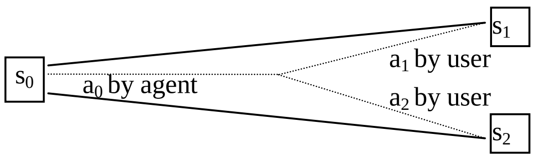

Figure 3. Non-deterministic Q-Learning

At any given time, the agent will be faced with different actions to take with the possibility of different outcomes and different rewards. The Q-learning algorithm is used to obtain the maximum expected reward given that the agent is at a particular state.

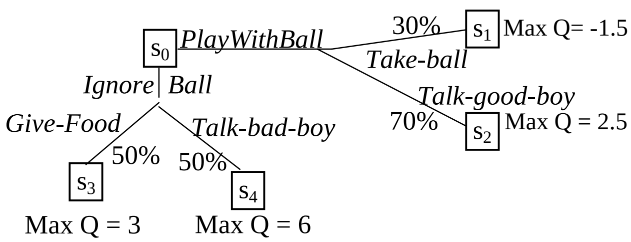

Figure 4. An example of reinforcement Learning

To illustrate the effect of the formula and the algorithm, we will adopt an example. Suppose the agent is trying to decide between two actions. These actions are illustrated in Figure 4. If it plays with the ball then there is a 70% chance that it will end up in state, s_2, which has a max Q-value of 2.5, but there is also a 30% chance of ending up in state, s_1, which has a max Q-value of -1.5. While if it ignores the ball then there is an equal chance of getting to state, s_4, which has a max Q-value of -3, or state, s_5, which has a max Q-value of 6. Thus, if we use regular probability calculation, action P*layWithBall* will have an expected Q-value of $0.3 \times -1.5 + 0.7 \times 2.5 = 1.3$, and *IgnoreBall* will have a value of $0.5 \times 3 + 0.5 \times 6 = 4.5$. Thus, the action, *IgnoreBall*, is considered more desirable than *PlayWithBall.*

Since we are trying to simulate a believable agent, we looked at how humans make decisions. We found that emotions and moods have a great impact on their decisions (Bower and Cohen 1982, Damasio 1994, Gardner 1983 and Goleman 1995). As noted in (Bower and Cohen 1982), when the agent is in a positive mood it will tend to be more optimistic, so it will naturally expect desirable events to occur. To replicate this phenomenon, we refined the expectation mechanism to reflect the mood. In a particular situation, the agent will be looking at numerous alternative paths, some of them may lead to failure and some of them may lead to success. If the agent's mood is positive then it will expect desirable events with a degree of β more than the negative events and vice versa.

Using this model the agent will be able to formulate its expectations according to its moods and the situation it faces. Furthermore, the agent will be able associate an event to a particular goal down the path.

Forming a User Model

Since the agent is interacting with the user, the agent will have to learn about the user's patterns of actions. A heuristic approach is used to define a pattern and to further define a probability of an action, a_1, to occur given that an action, a_2, has occurred. We focused on patterns of length three, i.e. the user did action a_1, then action, a_2, and finally action, a_3. This concept can be further illustrated using the pet and owner example. The owner goes into the kitchen *(a_1)*, takes out the pet's food from the cupboard *(a_2)* and feeds the pet *(a_3)*. These three consecutive actions led to the pet being fed. Thus, if the owner goes to the kitchen again, the pet would probably expect to be fed to a certain degree. As it can be seen from the example given above, learning sequences of actions can be very useful to predict other people's actions.

Most of the actions are very small, quick and specific, such as *ThrowBall*, and thus, an action or two consecutive actions by themselves are not meaningful. Therefore, in order to capture a meaningful pattern, we set the length of the pattern to a lower bound of two. Considering that the human mind can handle seven plus or minus two chunks (Miller 1956); we thought that handling a pattern of three consecutive actions is suitable for our agent considering that we decided to make the agent believable rather than intelligent.

There are two different operations done on a pattern. Firstly, the pattern is created or reinforced by the number of times it is repeated. Every pattern, *p*, will have an initial value, v_0, set when created. As soon as an action is taken, the history of actions is checked against the patterns stored in the agent's pattern tables. If a pattern exist then it is incremented by a unit. If it does not exist then a pattern is created and initialized to the initial value, v_0. Secondly, a probability of an event is calculated according to the expected action given that a sequence of one or two previous actions has occurred.

Learning Pleasing and Displeasing Actions

External reinforcement is used to learn what actions are pleasing or displeasing to the user. Thus, we are letting the user give the agent some feedback on his actions by saying *bad* or *good*. The agent assesses actions in terms of a function:

$$Quality(x, s)=v,$$

where x is an action that happens in a particular situation s, and v is the value given to the action-situation pair. When externally reinforced the agent searches its knowledge base for the action-situation pair. If they were not found, a value, v_a, is given to the action-situation pair. However, if the action-situation pair was found and the Quality matches the Quality that was documented then the pair is reinforced by η, otherwise the agent becomes confused and decrements the pair by η.

This type of learning is two folded; (1) it creates a basis by which the agent can please or displease the other when the situation arises, and (2) it creates the baseline by which the agent can set its own standards. These standards are then used to determine emotions, such as pride or shame. In addition, these standards are used to evaluate the user's actions and trigger emotions, including admiration and reproach. The intensities of these emotions are calculated as a function of the value, *v*, of the learnt rule.

Classical Learning or Conditioning

Associating objects with an emotion or a motivational state, forms yet another type of learning (LeDoux 1996). For example, if the agent experiences pain when an object, *g*, touches it, then a motivational state of pain will be associated with the object *g*. This kind of learning does not depend on the situation per se, but it depends on the object-emotion/ motivational state association.

Each of these associations will have a count, which is incremented by the repetition of the object-emotion occurrence. This type of learning will provide the agent with another type of expectation triggered by the object rather than the event. Using the intensity of the emotion triggered, the agent can calculate the expectation level. We used the formula shown below:

$$\langle Intensity(e|o)\rangle = \frac{\sum_{event(o)} Intensity(e)}{\sum_{i=o} event(i)}$$

where *e* is the emotion triggered, and *o* is the object introduced. In essence, the formula is averaging the intensity of the emotion in the events where the object, *o*, was introduced.

To illustrate this process we will adopt the following example. Consider a needle, which was introduced to the agent. The first time the needle was introduced it caused the agent 30% pain. So the agent will associate the needle with 30% pain. The next time we introduced the needle, the agent will expect pain with a level of 30%. Let's say that we did not inflict any pain on the agent, and we introduced the needle 99 times without inflicting any pain. Thus, the next time the needle is introduced, the agent will expect pain with a level of 0.3%. As you can see, the level of expectation of the intensity of a particular emotion, *e*, is decreasing by the number of times the object, *o*, was introduced without inflicting the emotion, *e*. If, however, we shocked the agent with the needle 90 times (30% of pain each time) and only introduced it 10 times without inflicting any pain, then the expectation of pain will increase to 27%.

Simulation & Results

It can be seen from the discussion above that several types of learning may indeed affect the emotional process and so affect the believability of the agent simulated. To validate this hypothesis, we had set up three models. One model simulated the emotional process without learning, the other simulated emotions with learning, and a third one simulated random set of behaviors and emotions. The third model was used to establish a baseline to compare the other two models. We used the same interface for all models. The interface, shown in Figure 5, is based on a Role Playing Games' interface, where the agent is simulated as a pet. The user can interact with the pet through various actions, including (1) introducing objects, (2) taking objects, (3) hitting objects, including the pet, and (4) talking aloud.

Figure 5. Model's Interface

Twenty-one users were assembled to evaluate the model and the hypothesis involved. The users were told that their evaluation would be used to further enhance the model rather than evaluate the quality of the model. This deception mechanism was made to make sure that none of the users was inclined or biased to evaluate the model positively due to his/her appreciation of the work involved or for other reasons. The users were chosen as first year undergraduates to decrease the bias inflicted by specialized background knowledge.

These users met with the principle investigator for a period of two and a half-hours. During this time, they ran the different simulations and answered a questionnaire. The questionnaire included questions about the pet's intelligence, learning and adaptation. The questions were in the form of 1 to 10 rating questions. For example, (1) 'did the pet meet your expectations? (Rate your answer from 1 to 10, 10 being it met all your expectations),' (2) 'did you think that pet's behavior was believable? (Rate your answer from 1 to 10, 10 being very believable).' Table 1 shows the user's answers to the believability question:

Table 1. Behavior Ratings

Model Type	**Behavior ratings**	
	Mean	**Conf. Interval**
Random	1	0.67 - 1.33
No Learning	5.43	4.86 - 6
Learning	8.095	7.65 - 8.542

As you can see from the table above, the ratings increased from an average of 5.43 (no learning) to an average of 8.095 (with learning). The interval, shown in the table, is calculated as a 95% confidence interval using the standard deviation calculated from the sample data. As shown by the interval column, the increase caused by the learning model is indeed significant. Thus, we can conclude that introducing learning increased the believability perceived by the users.

Future Work & Conclusion

The model is mainly an expansion on the many ideas that were presented in previous psychological and computational models. Our experiments indicated that incorporating various learning techniques improved the believability of the agent and produced a more dynamic emotional state. Nevertheless, the model still falls short in many aspects.

Personality has been recognized by many theatre artists to be the one factor that helps in the development and creation of believable characters. Furthermore, psychologists recognized it to be the factor that influence and is being influenced by the learning and the emotional processes. By developing experience, we defined one way of modeling personality, however, we identify personality as another important component that has to be added to our model in the near future.

Moreover, many other important factors are missing from the model, including self-perception, self-awareness and self-esteem. These factors were recognized by psychologists to have a major impact on the subject's emotional state and responses (Roseman, Jose, and Spindel 1990, Goleman 1995).

The emotional process and the decision-making process communicate in a much more complex fashion than what is displayed in the model above. There is a lot to be gained by looking at the link between these components in more detail. However, we decided to keep that for future work.

In conclusion, our experiments showed that learning shaped and helped in capturing the complex nature of the emotional process. We think that other types of learning, if simulated, can also affect the emotional process, including social learning opening up another door for future research - building a social entity - a social agent.

References

Bates, J. 1992. The Role of Emotion in Believable Agents. *Communications of the ACM* 37:122-125.

Bolles, R. C., and Fanselow, M. S. 1980. A Perceptual Defensive Recuperative Model of Fear and Pain. *Behavioral and Brain Sciences* 3:291-301.

Bower, G. H., and Cohen, P. R. 1982. Emotional Influences in Memory and Thinking: Data and Theory. *Affect and Cognition*, 291-331. Clark, M., and Fiske, S. eds. London: Lawrence Erlbaum Associated Publishers.

Damasio, A. R. 1994. *Descartes' Error: Emotion, Reason, and the Human Brain.* NY: G.P. Putnam.

Gardner, H. 1983. *Frames of Mind.* NY: Basic Books.

Goleman, D. 1995 *Emotional Intelligence.* NY: Bantam Books.

Kaelbling, L. P., Littman, M. L., and Moore, A. W. 1996. Reinforcement Learning: A Survey. *Journal of Artificial Intelligence Research* 4:237-285.

LeDoux, J. 1996. *The Emotional Brain.* NY: Simon & Schuster.

Maes, P. 1997. Agents that Reduce Work and Information Overload. *Communications of ACM* 37:37-45.

Masuyama, A. 1994. A Number of Fundamental Emotions and Their Definitions. *In Proceedings of IEEE International Workshop on Robot and Human Communication*, 156-161. Tokyo, Japan.

Miller, A. 1956. The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information. *Psychological Review* 63: 81-97.

Minsky, A. 1986. *The Society of the Mind.* NY: Simon and Schuster.

Mitchell, T. 1996. *Machine Learning.* NY: McGraw-Hill.

Mizogouhi, H., Sato, T., and K. Takagi. 1997. Realization of Expressive Mobil Robot. In *Proc. IEEE Int. Conf. of Robotics and Automation*, 581-586. New Mexico, USA.

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

Picard, W. 1997. *Affective Computing.* MA.: MIT Press.

Price, D. D., Barrell, J. E., and Barrell, J. J. 1985. A Quantitative-Experiential Analysis of Human Emotions. *Motivation and Emotion* 9:19-38.

Reilly, S. 1996. Believable Social and Emotional Agents. Ph.D. diss., Dept. of Computer Science, Carnegie Mellon Univ.

Rosman, J., Jose, P. E., and Spindel, M. S. 1990. Appraisals of Emotion-Eliciting Events: Testing a Theory of Discrete Emotions. *Journal of Personality and Social Psychology* 59:899-915.

Seif El-Nasr, M. 1998. Modeling Emotion Dynamics in Intelligent Agents. Master's thesis, Dept. of Computer Science, Texas A&M Univ.

Seif El-Nasr, M., and Yen, J. 1998. Agents, Emotional Intelligence and Fuzzy Logic. In *Proceedings. of NAFIPS*. FL, USA.

Shibata, Ohkawa, K., and Tanie, K. 1996. Spontaneous Behavior of Robots for Cooperation - Emotionally Intelligent Robot System-. In *Proceedings of IEEE Int. Conf. on Robotics and Automation,* 2426-2431. Tokyo, Japan.

Shiida. 1989. An Attempt to Model Emotions on a Machine. *Emotion and Behavior* 2:275-287.

Velasquez, J. 1997. Modeling Emotions and Other Motivations in Synthetic Agents. In *Proceedings of the AAAI Conference*, 10-15. RI, USA.

Yen, J., and Langari, R. 1998. *Fuzzy Logic: Intelligence, Control and Information.* Forthcoming.