Modeling Users' Emotions During Interactive Entertainment Sessions

Piotr J. Gmytrasiewicz* **and Christine L. Lisetti†**
*Department of Computer Science and Engineering
University of Texas at Arlington,
Arlington, TX 76019-0015
†Department of Information Systems and Decision Sciences
University of South Florida
Tampa, FL 33620-7800
piotr@cse.uta.edu, lisetti@coba.ufs.edu

Abstract

We use the formalism of decision theory to develop principled definitions of emotional states of a user, and we postulate that these notions should be useful in designing interactive entertainment systems. First, it is essential that the system be aware of the user's emotional state to be able to react to it appropriately. Second, well defined notions of emotional states provide a well defined implementation-independent vocabulary the system and the user can use to communicate their internal states to each other. Using decision theory to formalize the notions of emotions provides a formal bridge between the rich bodies of work in cognitive science, on the one hand, and the high-end AI architectures for designing rational artificial agents, on the other hand.

Introduction

Our work is predicated on the thesis that the concept of emotions and feelings can be formalized and be made useful in designing interactive entertainment systems that are to flexibly and adaptively interact with human players. Our formalization starts from the formal description of a rational agent based on decision theory, according to which agents act so as to maximize the expectation of their performance measure. This decision-theoretic model of decision making can be used to formally define the emotional states and feelings of a rational agent. Our definitions show how emotional states transform the agent's decision-making situation, say, by making the agent more short-sighted, by altering the agent's subjective performance measure (preferences), or by modifying the probabilities that the agent assigns to states of the world.

[0]This research is supported, in part, by the Office of Naval Research grant N00014-95-1-0775, and by the National Science Foundation CAREER award IRI-9702132.

Having the formal definitions of emotional states allows us to show how, and why, they are useful for interactive entertainment systems. First, well defined emotional states of synthetic chracters and others are crucial in these agents' interactions with humans. Frequently, such interactions are impeded by the system or the synthetic character being hopelessly out-of-step with the emotional state of the human player. However, the player' emotional state, such as anger, fear, boredom, panic, surprise, joy, or excitation, can be assessed using measurable and inferred factors (facial expression recognition, vocal intonation, prosody, galvanic skin response, heart rate and breathing patterns, haptic and tactile feedback, body posture (Cottrell & Metcalfe 1991; Hayes-Roth *et al.* 1998; Lisetti 1998; Lisetti & Schiano 1999; Picard 1997)). Thus, it is important for the machine to model the effect the player's emotional state has on the his/her decision-making and his/her tendency for action during play. Second, emotions are valuable when the system or a synthetic agent finds it useful to inform the human player or other agents about its own internal state. Instead of having to describe the details of its internal state, and running the risk of being misunderstood if the other agents are engineered differently, the agent can use more abstract and universal terms. For example, notions of stress or panic may be convenient to express the fact that the urgency of the situation forced the agent to look at only short-term effects of its actions. Thus, in the context of communication with human players and other agents, the formal definitions of emotions serve as semantics of emotional terms, with which the agents can express their own internal states and understand the states the other agents are in.

We envision that results of our work will be useful in Human-Computer Interaction (HCI) and applications will range from intelligent tutoring systems and distance learning support systems (with recognition of expressions signaling interest, boredom, confusion), to stress and lie detectors, to monitors of pilots and drivers' state of alertness, to software product support systems (with recognition of users being dis/pleased with software products), to entertainment and computer games (enjoyment, confusion), to ubiquitous computing and smart rooms with awareness (of all of the above expressions).

Decision-Theoretic Preliminaries

The objective of our research is to develop a fundamental understanding of the role and usefulness of the concept of emotional states in designing human-computer interaction systems. The main hypothesis of our work is that concept of emotions and feelings can turn out to be very useful in designing competent synthetic agents that should operate within complex uncertain environments populated by other artificial and human agents. Our approach draws on and combines an emerging technology of rational agent design in Artificial Intelligence on the one hand (Dennett 1986; Doyle 1992; Russell & Norvig 1995; Simon 1958; Wooldridge & Anand Rao 1999), with research on human emotions in cognitive science and psychology on the other hand (Dimasio 1994; James 1884; 1894; Johnson-Laird & Oatley 1992; Leventhal & Scherer 1987; Ortony, Clore, & Collins 1988; Simon 1967; Sloman & Croucher 1981; Sloman 1990).

Our work accepts the decision-theoretic paradigm of rationality, according to which rational agent should behave so as to maximize the expected utility of its actions (see (Doyle 1992; Russell & Norvig 1995) and other references therein). The expected utilities of the alternative courses of action are computed based on their possible consequences, the desirability of these consequences to the agent[1], and the probabilities with which these consequences are thought by the agent to obtain.[2] The main thrust of our work is to examine ways in which components of the decision-theoretic model, i.e., the utility functions, the set of behavioral alternatives, and the probabilities of consequences, can be transformed in ways that that has been recognized in cognitive science as interactions between emotional states and decision-making, and to utilize these concepts to make the synthetic interactive entertainment systems more believable.

A rational agent formulates its decision making situation in terms of a finite set, A, of the alternative courses of action, or behaviors, it can execute. An alternative behavior, say a_i, can be thought of as a plan consisting of consecutive actions extending into the future time t_{a_i}, which we will call the time horizon of this particular plan. Thus , each plan has its own time horizon and their set is $A = \{a_1, a_2, ...\}$. Alternative courses of action in set A can further stand for abstract actions as well as for detailed elaborations; increasing the level of abstraction facilitates keeping the size of A down to manageable proportions. We demand that the actions be distinct. We also demand the the set A be exhaustive, i.e., that all of the possible behaviors be accounted for. Sometimes an "all-else" behavioral alternative is used for compactness, and represents all other possible behaviors except the ones explicitly enumerated. We will denote this alternative as $\mathbf{a}$, and set A can be written as $A = \{a_1, a_2, ..., a_m, \mathbf{a}\}$, and will be called the agent's *action space*.

Apart from what it can possibly do, an agent currently finds itself in some state of the world, but due to the fact that the environment may not be fully observable the agent may be uncertain about the state. The fact that the actual state may be unknown to the agent can be formalized by specifying the set of all possible states of the world, S, together with a family of probability distributions, $\mathbf{P}(S)$, over these states. One of these distributions, say $P_c(S)(\in \mathbf{P})$, specifies which of these states are currently possible and how likely they are. Thus $P_c(S)$ fully describes the information the agent has about the present state of the world. Given the distribution $P \in \mathbf{P}$ it will be convenient to introduce the entropy function $Ent(P)$. Entropy is an inverted measure of uncertainty; if the agent is certain about the state of the world, the probability distribution is binary, and the entropy is zero. If the agent has no information at all, the distribution is uniform and the entropy is maximized.

Now, the agent can ponder the consequences of its alternative actions. Due to possible nondeterminism each action, $a_i \in A$, may lead to many resulting possible states. The likelihoods of the resulting states can be specified by another probability distribution, $P_i(S)(\in \mathbf{P})$, also over S. The process of determining the probabilities of different results, i.e., the distribution P_i has been called a probabilistic temporal projection. Thus, projection is a function $Proj : \mathbf{P}(S) \times A \rightarrow \mathbf{P}(S)$, so that the result of projecting the results of action a_i given the current information about the state $P_c(S)$ results in the projected information about the resulting state, $P_i(S)$: $Proj(P_c(S), a_i) = P_i(S)$.

Of importance are also the desirabilities of the resulting states. The desirabilities of the states of the world to the agent are encoded using a utility function $U : S \rightarrow R$, which maps states of the world to real numbers. Intuitively, the higher the utility

[1]Such agents are sometimes called self-interested.

[2]The probabilities are, therefore, subjective.

value of a state the more desirable this state is to the agent.

The agent decision problem involves choosing which of the alternative actions in the set A it should execute. One of the central theorems of decision theory states that if the agent's utility function is properly formed, and the agent expresses its uncertain beliefs using probabilities, then the agent should execute an action, a^*, that maximizes the expected utility of the result.

$$a^* = ArgMax_{a_i \in A} \sum_{s \in S} p_i^j U(s^j), \tag{1}$$

where the p_i^j in the probability the projected distribution $P_i(S)$ assigns to a state $s^j \in S$. Frequently, it is convenient to represent the utility function, U, as depending on a small number of attributes of the states of the world, as opposed to depending on the states themselves. This is intuitive; humans may prefer, say, all of the states in which they have more money, are more famous, and are healthier. The attributes, say wealth, fame, and health are then convenient factors in terms of which the utility function can be expressed. Multi-attribute utility theory postulates that, in some simple cases, the utility of a state is a weighted sum of the utilities, $U(X_l(s))$ of individual attributes:

$$U(s) = \sum_{X_l \in Attributes} W_{X_l} U(X_l(s)), \tag{2}$$

where the W_{X_l} is the weight, or intuitively, the importance, of the attribute X_l. Having the weights of the attributes explicitly represented is convenient since it enables the tradeoffs among the attributes the agent may have to make. For example, the agent may have to give up some of its wealth to improve its health, and so on.

The elements defined above are sufficient to formally define a decision-making situation of an agent:

Definition 1: A decision-making situation of an agent is a quadruple: $D =< P_c(S), A, Proj, U >$, where S, $P_c(S)$, A, $Proj$ and U are as defined above.

The above quadruple fully specifies the agent's knowledge about the environment, the agent's assessment as to its possible courses of action, the possible results of the actions, and desirability of these results. Our definition here is closely related to that of stochastic processes, and in particular to Markov decision process (see (Boutilier, Dean, & Hanks 1999; Russell & Norvig 1995) and references therein), but it makes explicit the decision problem the agent is facing by enumerating the alternative action sequences the agent is choosing among.

Given its decision-making situation, and agent can compute its best action, a^*, as specified in Equation 1. It is clear that this computation can be fairly complex. In a multi-agent environment, for example, all of the information the agent has about the physical environment and about the other agents could be relevant and impact the expected utilities of alternative courses of action. Sometimes the agent may have information about the other agents' state of knowledge, which is also potentially relevant. Given these complexities it is clear that a mechanism for managing the agent's computational resources is needed. Here, we suggest that emotional states, as defined below, may provide for such ability. Other approaches, such as compilations of decision-theoretic reasoning, approximations, and simplifications were investigated in (Noh & Gmytrasiewicz 1999).

Emotional States: Classification and Dynamics

As we mentioned, we will view emotions as transformations of the decision-making situation defined above. First, we briefly describe a taxonomy of emotional states and a finite state machine model of dynamics of emotions which can assist agents in measuring and predicting the emotional state of other agents.

The taxonomy (Ortony, Clore, & Collins 1988) of emotional states in Figure 1 is aimed at differentiating among emotional states by using values of well-defined attributes. It is clearly desirable to be able to measure the values of the attributes to be able to determine the current emotional state of another agent. Given the emotional state, its impact on decision-making can be modeled. Presently, only some of the attributes differentiating various emotional states are directly measurable; in human agents for example, the positive and negative values of valence attribute may be measurable from facial expression recognition tools. Further, the taxonomy is incomplete in that we (and our colleagues in cognitive science) do not yet know what attributes and their values should be used to differentiate among some emotional states.

Our framework also uses a dynamic model of user's emotional states. Its purpose is twofold: First it is frequently useful not only to assess the current emotional state the other agent is in but also to predict what emotional state will obtain, given the current emotion on one hand, and a system's response or environmental event on the other hand. Second, since some of the variables that determine the current emotional state are not measurable, it may be possible to infer the current state from the previous state, if known, and the environmental input. In Figure 2 we present a very simple example model of

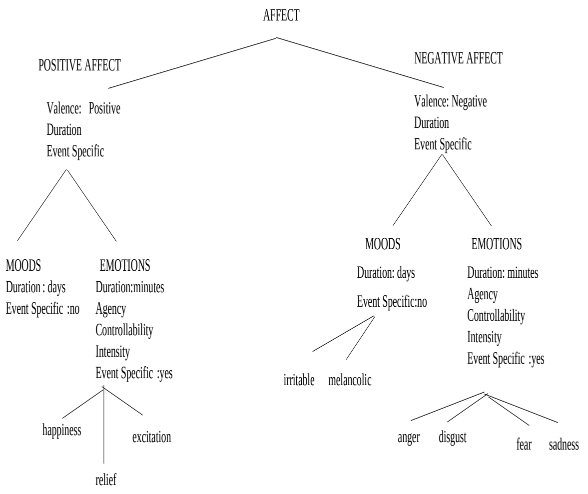

Figure 1: Partial Taxonomy of Emotional States

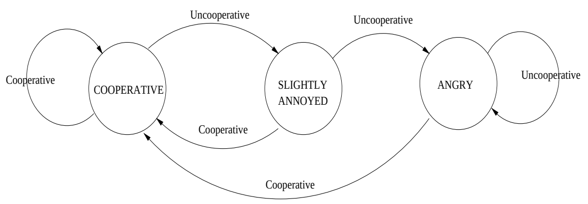

Figure 2: Simple Dynamic Model of an Agent's Emotional States

emotions' dynamics. It contains only three emotional states; COOPERATIVE, SLIGHTLY ANNOYED, and ANGRY. The transitions among the states are caused by environmental inputs or responses of the system, and they are divided into categories of Cooperative and Uncooperative. Using this dynamic model one can predict that an agent that is in COOPERATIVE emotional state will become SLIGHTLY ANNOYED given Uncooperative input. Further, the emotional state of SLIGHTLY ANNOYED will evolve into ANGRY if another Uncooperative response follows.

Dynamic models that use the formalism of finite state automata like the one in Figure 2 are common and used in the field of game theory. The one we present here coincides with the Tit-for-two-Tat strategy used for the Prisoner's Dilemma game (Axelrod 1984; Carmel & Markovitch 1996; Rubinstein 1986). The advantage of modeling emotional states of the user with the finite automata model is that models of this kind can be learned, for example using an unsupervised US-L learning algorithm (Carmel & Markovitch 1996; Littman 1994; Crites 1995).

Decision-Theoretic Definitions of Emotions

We now briefly outline some transformations of the decision-making situation of an agent. We call them *emotional transformations*, and their results are *emotional states*. In other words, an emotional transformation changes one decision-making situation, say a NEUTRAL emotional state, into another one, say an ANGRY emotional state. We assume that the emotional transformations themselves are triggered by some environmental input, IN, the agent experiences. We should caution that our identifying emotions with such transformations does not account for all of the richness of emotions in humans; it merely captures some of their intuitive functionality.

Let us denote as $\mathbf{D}$ the set of all decision situations, D, as defined by Definition 1. Given our decision-theoretic approach, we will assume that each $D \in \mathbf{D}$ defines an emotional state. This assumption may seem overly strong and suggests a one-to-one and onto mapping between emotional states and decision-making situations. In general such a mapping may not exist since, for example, many emotional states may map into the same decision situation. We justify it here by observing that emotional states that do not change decision-making situation of the agent are irrelevant from the point of view of the agent behavior, which is what we ultimately want to model here. Further, let $\mathbf{IN}$ be the set of all environmental inputs.

Definition 2: Emotional transformation is a function $EmotTrans : \mathbf{D} \times \mathbf{IN}^* \rightarrow \mathbf{D}$.

Thus, given an initial emotional state, D, and a, possibly empty, history of environmental inputs IN, the value of the $EmotTrans$ function is $EmotTrans(D, IN) = D'$, where D' is the agent's new emotional state. Examples of such emotional transformations are depicted in Figure 2. D' may differ from D in a number of ways. Below we look at some possibilities that correspond to some of the more intuitive emotional states.

Transformations of the action space A

Transformation of the action space A, for example by narrowing the set of alternative actions considered to encompass only a small subset of all of the actions, predisposes the agent to take action from this smaller set. This constitutes the *action tendency* that the emotion is invoking in the agent, as postulated, for example, by Fridja in (Fridja 1986). In the extreme, narrowing the set A to a single action implements a behavioral condition-response rule.

Formally, these are transformations $EmotTrans(D, IN) = D'$ such that $D =< P_c(S), A, Proj, U >$, and $D' =< P_c(S), A', Proj, U >$. An emotional transformation that implements an action tendency in one for which $A' \subset A$. For example, an agent becoming angry may result in it considering only a subset of its behavioral alternatives, say, ones of aggressive nature. A special case of this emotional transformation obtains when A' is a singleton set, containing only one behavior. This is an implementation of a emotional condition-action rule; all three emotional states that correspond to the Tit-for-two-Tats strategy in repeated Prisoner's Dilemma game in Figure 2 are of this kind since they result in the agent's being capable of performing only a single behavior.

Another intuitive special case of such transformation is one that results in the agent's deliberating in a more short-term fashion, such as it being rushed or panicked under time pressure. Formally we have: $\forall a_i^{'} \in A' : t_{a_i^{'}} \leq t_{a_i}$, which states that the time horizon of alternative plans considered has diminished. This is characteristic of human decision-makers; people frequently become more short-sighted when they are rushed or panicked, since they have no time to consider long-term effects of their alternative behaviors.

Transformations of the utility functions U

Intuition behind this transformation is that emotions and feelings both implement U, as well as modify it. Humans evaluate desirability of states by having positive or negative feelings about them. Positive or negative emotions or moods may alter these evaluations by, say, decreasing them, as in melancholic or depressed moods (when everything looks bleak), or increasing them, as in elated or happy moods. Other emotional states can change the weights of the factors contributing to the utility ratings (Equation 2).

Formally, these are transformations $EmotTrans(D, IN) = D'$ such that $D =< P_c(S), A, Proj, U >$, and $D' =< P_c(S), A, Proj, U' >$.

The special case of sadness or melancholy result in evaluation of the desirability of every state to diminish: $\forall s \in S : U'(s) \leq U(s)$.

Transformations of the probabilities of states

The intuition behind this transformation is that changing these probabilities, for instance by simplifying them, can be helpful and save time under time pressure. The most radical simplification is one that makes the most likely state to be the only possible state or result. This corresponds to considering only the most likely result of action and neglecting all less likely states and is often observed in human decision-makers.

Formally, these are transformations $EmotTrans(D, IN) = D'$ such that $D =< P_c(S), A, Proj, U >$, $D' =< P_c^{'}(S), A, Proj^{'}, U >$. The special case described above obtains is when the probability distribution $P_c^{'}$, as well as every projected distribution $P_i^{'}$ returned by the projection function $Proj^{'}$, assume only values of 0 or 1 and have zero entropy; $Ent(P_c^{'}) = Ent(P_i^{'}) = 0$.

Conclusions and Future Work

This paper outlined an approach to formally defining the notions of emotions of rational agents with the aim of using it in human computer interaction, and in particular in interactive entertainment systems. Following one of the recent approaches to designing rational agents based on decision theory (Russell & Norvig 1995), we attempted to define emotional transformations, and the resulting emotional states, as resulting from input the agents receive from the environment. The various emotional states are identified as various possible modifications of the decision-making situations the agents find themselves in.

We argue that our approach is one way to formally bridge a gap between the multitude of works in cognitive science on one hand, and between formal approaches to rational agent design in AI. More importantly, we argue that this direction is one that should be of interest for practitioners designing and building systems interacting with human users, when the system may need to adjust to the users emotional state, and may need a vocabulary that precisely defines what behavioral changes should be expected of users under various emotional conditions.

The approach we outlined serves as a point of departure for much of the needed future work. Clearly, the definitions of emotional transformations can be elaborated upon, and more intuitive special cases can be arrived at. These cases should ultimately find their way into the taxonomy depicted in Figure 1, and be well defined in terms of the attributes, as described in (Ortony, Clore, & Collins 1988). Further, the dynamic models of emotional states, like the one in Figure 2, can become far more elaborate, thus allowing the agents to predict the emotional states of users in much more detail. Finally, our future work will involve implementation and experimental validation of the approach proposed here.

References

Axelrod, R. 1984. *The Evolution of Cooperation.* Basic Books.

Boutilier, C.; Dean, T.; and Hanks, S. 1999. Decision-theoretic planning: Structural assumptions and computational leverage. *Journal of Artificial intelligence Research* 11:1–94.

Carmel, D., and Markovitch, S. 1996. Learning models of intelligent agents. In *Proceedings of the National Conference on Artificial Intelligence*, 62–67.

Cottrell, G., and Metcalfe. 1991. Empath: Face, emotion and gender recognition using holons. In *Advances in Neural Information Processing.* Morgan Kaufman Publishers.

Crites, T. S. R. H. 1995. Multiagent reinforcement learning and iterated prisoner's dilemma. *Biosystems Journal* 37:147–166.

Dennett, D. 1986. Intentional systems. In Dennett, D., ed., *Brainstorms.* MIT Press.

Dimasio, A. R. 1994. *Descartes' Error.* Grosset/Putnam.

Doyle, J. 1992. Rationality and its role in reasoning. *Computational Intelligence* 8:376–409.

Fridja, N. H. 1986. *The Emotions.* Cambridge University Press.

Hayes-Roth, B.; Ball, B.; Lisetti, C.; and Picard, R. 1998. Panel on affect and emotion in the user interface. In *Proceedings of the 1998 International Conference on Intelligent User Interfaces*, 91–94.

James, W. 1884. What is an Emotion? *Mind* 9:188–205.

James, W. 1894. The Physical Basis of Emotion. *Psychological Review* 1:516–529.

Johnson-Laird, P. N., and Oatley, K. 1992. Basic Emotions, Rationality, and Folk Theory. *Cognition and Emotion* 6(3/4):201–223.

Leventhal, H., and Scherer, K. 1987. The relationship of emotion to cognition:a functional approach to semantic controversy. *Cognition and Emotion* 1(1):3 – 28.

Lisetti, C., and Schiano, D. J. 1999. Automaic facial expression interpretation: Where human interaction, artificial intelligence and cognitive science intersect. *Pragmatics and Cognition, Special Issue on Facial Information Precessing and Multidisciplinary Perpective.*

Lisetti, C. 1998. Facial expression recognition using a neural network. In *Proceedings of the Florida Artificial Intelligence Research Symposium.*

Littman, M. L. 1994. Markov games as a frameowrk for multi-agent reinforcement learning. In *Proceedings of the International Conference on Machine Learning.*

Noh, S., and Gmytrasiewicz, P. J. 1999. Towards flexible multi-agent decision-making under time pressure. In *Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence*, 492–498.

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

Picard, R. 1997. *Affective Computing.* MIT Press.

Rubinstein, A. 1986. Finite automata play the repeated prisoner's dilemma. *Journal of Economic Theory* 39:83–96.

Russell, S., and Norvig, P. 1995. *Artificial Intelligence: A Modern Approach.* Prentice Hall.

Simon, H. A. 1958. *Rational Choice and the Structure of the Environment.* MIT Press.

Simon, H. 1967. Motivational and Emotional Controls of Cognition. *Psychological Review* 74:29–39.

Sloman, A., and Croucher, M. 1981. Why robots will have emotions. In *Proceedings of the Seventh IJCAI Vancouver, B.C.*, 197–202. San Mateo, CA: Morgan-Kaufmann.

Sloman, A. 1990. Motives, Mechanisms, and Emotions. In Boden, M., ed., *The Philosophy of Artificial Intelligence.* Oxford: Oxford University Press.

Wooldridge, M., and Anand Rao, E. 1999. *Foundations of Rational Agency.* Kluwer Academin Publishers.