Towards Affect-based Approximations to Rational Planning: A Decision-Theoretic Perspective to Emotions

Prashant Doshi and **Piotr Gmytrasiewicz**
Dept. of Computer Science, Univ. of Illinois, Chicago
{pdoshi, piotr}@cs.uic.edu

Introduction

We are interested in projecting emotions as a tool for approximating computationally intensive sequential decision-making models of single-agent systems. Specifically, we believe that affect can be used to tradeoff computational complexity with quality of the solution, thereby producing approximations that compare favorably with other approximation techniques. Our research is predicated on the thesis that cognitive appraisal of the environment induces internal emotions that in turn affect the agent's decision-making process. In order to validate our thesis, we require a computational model for capturing the dynamics of emotions and generating behaviors characteristic of several emotional states.

We are especially interested in modeling affective agent planning in stochastic domains where the agent is unaware of its current state (e.g. its physical location), but receives a series of percepts that may guide it. In this context we propose utilizing a decision-theoretic framework called **P**artially **O**bservable **M**arkov **D**ecision **P**rocess (POMDP). Solution of a problem modeled as a POMDP, produces a *policy* that controls agent planning. Exact solutions of POMDPs are intractable spurring interest into approximation techniques that compute near-optimal solutions using less running time. Towards this end, we propose employing affect as a justification for reducing the size of the input POMDP model, and therefore the computational complexity of its solution. In doing so, we not only demonstrate the feasibility of employing affect as a potential approximation technique, but also present a computational model that can generate behaviors characteristic of several emotions.

The primary contributions of this paper are twofold: We explore the feasibility of employing affect to produce near-optimal approximations to optimal plans in less time. Secondly, we propose a hybrid agent architecture that employs a reactive finite state machine based metareasoner and a deliberative POMDP based planner, to model different affective behaviors. We validate our claims through experimentation on an example toy problem domain.

The remainder of this paper is structured in the following manner. In the next section we discuss some related work and put our model in the proper context. Thereafter we introduce the planning model, followed by our affect-based approximation technique. Finally, we test our technique on an example problem domain and report our findings.

Related Work

In (Gmytrasiewicz & Lisetti 2001) a decision-theoretic approach to formalizing emotional behavior was introduced. This paper builds on the earlier work by utilizing a specific decision-theoretic model, and applying it to a problem domain. Fellous et al. (Fellous & Hudlicka 1996) have put forward a classification system for organizing different computational models of emotions. In this system, our decision-theoretic model will be classified as a high-level mechanism model capable of modeling planning behaviors characteristic of several affective states. Several bodies of research (Elliott 1992; Gratch 1999) have observed subjectively, the influence of emotions on immediate actions and planning. Velasquez (Vel'asquez 1998) specifically looks into emotion-based decision-making but does not utilize decision-theoretic models. In contrast, we present a *concrete* framework applicable to several domains, and show how it can be engineered to produce emotional behavior. Additionally, rather than reacting to an emotion eliciting condition, our model generates appropriate long-term plans.

Planning Model

We model the agent's planning process as a Partially Observable Markov Decision Process(POMDP) (Cassandra, Kaelbling, & Littman 1994). A POMDP describes a stochastic control process, and is formally defined as a sextuplet $P = (S, A, \Theta, T, O, R, H)$ where S is a set of world states; A is a set of agent actions; Θ is a set of observations; $T : S \times A \rightarrow \Pi(S)$ is a set of transition probabilities between states that describe the dynamic behavior of the modeled environment ($\Pi(\cdot)$ denotes a probability distribution); $O : S \times A \rightarrow \Pi(\Theta)$ gives a probability distribution over the observations given the action and the resulting state; $R : S \times A \times S \rightarrow \Re$ models payoffs associated with each transition; and $0 \leq H \leq \infty$ is the planning horizon or the future lookahead while planning. Since the agent is unaware of its state, it maintains a probability distribution (belief) over all possible states. The decision (or planning) problem in the context of POMDP requires one to find an action or a sequence of actions for one or more belief states that optimizes the objective reward function. The belief state compactly represents all information available to the agent at the time of selection of the optimal action.

The problem of designing optimal plans for POMDPs is PSPACE-complete (Papadimitriou & Tsitsiklis 1987). Several approximation techniques (Lovejoy 1991) attempt to tradeoff computational complexity with the quality of the solution. However, the resulting bounds on drop in quality are loose, and frequently produces highly sub-optimal behavior.

Approximation Method

Both classical and decision-theoretic planning mechanisms have typically followed a paradigm of rationality which shuns the effect of human-like emotions. However, Damasio's investigations (Damasio 1994) uncovered evidence that too little emotion can also impair decision-making. Motivated by these investigations, we are exploring the idea of using affect to approximate the "rational" exact model.

Affect-based Model Approximation

Our approximation method exploits the idea that certain emotions such as $Panic$, and $Anger$ cause the agent to deliberate less, consider only certain action alternatives, and/or change their goals (reward functions). Specifically, a $Panicked$ agent may consider only escape, and an $Angry$ agent may shorten its deliberation process. Subsequently, we use affect as a means of reducing the size of the POMDP model, and thereby the time and space its takes to produce a solution. Let *P' = (S, A', Θ', T', O', R', H')* be the planning model of an emotional agent. We term *P'* to be an approximation of *P*, *P'* $\approx$ *P*, and observe that *A'* $\subseteq$ *A*, Θ' $\subseteq$ Θ, and/or *H'* $\leq$ *H*. This potential reduction in the size of the original model directly translates into a reduction in the computation time of the solution.

The POMDP definition requires specification of several agent parameters. Tuning these parameters permits us to model behaviors characteristic of several emotions. Therefore, we put forward POMDPs as *sufficiently expressive* frameworks that are capable of modeling the cognitive aspects of many emotions. Rather than simply reacting to an emotion, these frameworks produce distinct plans that are characteristic of the agent's current emotional state.

Emotion Transition Model

Akin to a human being living his day-to-day life, an affective agent experiences a gamut of emotions, whose transitions are triggered by environmental percepts. Hence, computational models that capture emotional transformations are required to model affective behavior of an agent.

Picard (Picard 1995) proposes the use of HMMs for capturing the dynamics of emotional states. In this work, we utilize FSMs for modeling emotional transformations. One reason for selecting FSMs is that efficient algorithms exist (Trakhtenbrot & Brzdi 1973) for learning FSMs from a small set of labeled data.

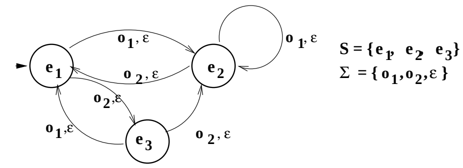

Figure 1: A deterministic FSM with three emotional states, and two environmental percepts.

A FSM is formally defined as a quadruplet $\mathcal{F} = \{\mathcal{S}, s_0, \Sigma, \mathcal{T}\}$ where $\mathcal{S}$ is the set of states; $s_0 \in \mathcal{S}$ is the start state; Σ is the input alphabet responsible for causing the transitions; and $\mathcal{T}$ is the transition function, $\mathcal{T} : \mathcal{S} \times \Sigma \rightarrow \mathcal{S}$. An agent whose emotional state transformation is modeled as a FSM, $\mathcal{F}$, exhibits behavior befitting of the current (emotional) state of $\mathcal{F}$. When an environmental percept, $o \in \Sigma$, arrives, the agent may alter its emotional state as governed by $\mathcal{T}$. The agent then displays behavior characteristic of its new emotional state. Here we assume that the environmental percepts satisfy the *emotion eliciting conditions* as defined by Elliott (Elliott 1992) and therefore activate the appropriate emotion in the agent.

Experimental Evaluation

To simulate a dynamic environment and emotion eliciting conditions we use a modified version of the **Wumpus World** toy problem given in Russell and Norvig (Russell & Norvig 1995).

Example Problem Domain

In the Wumpus World problem, an agent must reach the gold location, grab the gold, and arrive at a specific location with the gold while avoiding the Wumpi. We ran our experiments on a world that has 4×4 locations, one moving Wumpus, and a single stationary gold. The agent can perform 5 actions (move north, south, west, east, and grab gold) and receive 4 observations (stench, glitter, both, and none). We modify the original problem to a dynamic one by including a moving Wumpus. The Wumpus moves one location at a time in the horizontal direction, and on reaching the end of the row, it reappears at the start of the row and resumes its movement. Additionally, the Wumpus produces a stench that is evident to the agent when it is within two locations of the Wumpus in the horizontal or vertical direction (but not diagonal). The speed of movement of the Wumpus serves as a knob to control the dynamism of the Wumpus World. The gold produces a glitter which is observable by the agent when the agent is within two locations of the gold in the horizontal or vertical direction(but not diagonal). Fig. 2 shows our toy problem.

Both, the actions and the observations are deterministic. The agent receives a reward of 1000 points for arriving at the goal state, a penalty of 10,000 points for colliding with the Wumpus, and a penalty of 1 point for each other action taken. On reaching the goal state, or colliding with the Wumpus, the current trial ends and a new one begins.

The movement of the Wumpus is synchronized with a clock timer that ticks with a constant frequency. The agent moves only when its deliberation is complete and has yielded an action. Consequently, excessive deliberation by

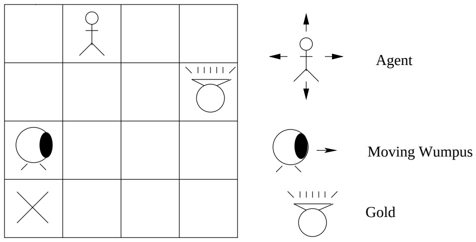

Figure 2: An example Wumpus World. The agent must arrive with the gold at the location marked with a "X".

the agent will result in a "no-op", though the environment may continue to change.

Affective Agent Design

The FSM shown in Fig. 3 captures the changes in the agent's emotional state and is formally defined as: $\mathcal{F} = \{\mathcal{S}, s_0, \Sigma, \mathcal{T}\}$ where $\mathcal{S}$ = {Contented,Panic,Fear,Elation}; s_0 = Contented; Σ = {Smell,Glitter,Both,None}; and $\mathcal{T}$ is as shown.

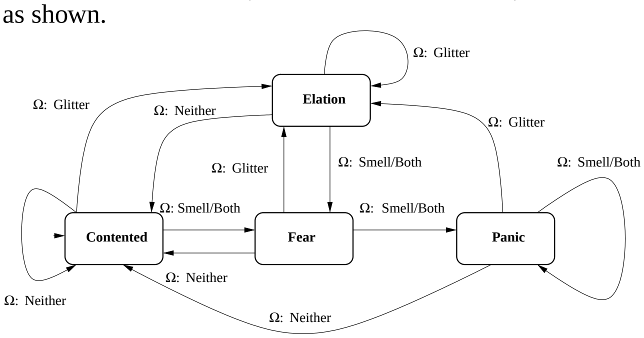

Figure 3: Emotional transformations in the Wumpus World

We have identified four basic emotions that an affective agent will realistically experience in our Wumpus World. Initially, the agent is **Contented**, but a percept of *Smell*, or *Both*, will induce an emotion of **Fear** in the agent, due to the presence of the Wumpus nearby. A further percept of *Smell*, or *Both* will induce **Panic** in the agent. A percept of *Glitter*, will cause the agent to feel **Elated** due to the absence of any nearby threats. The emotions of **Fear**, **Panic**, and **Elation** cause the agent to alter its original planning model.

The exact (rational) planning model, P, employed by an agent in the Wumpus World is defined as follows. The state of the agent, S, is completely captured using the variables: *Location of Agent, Agent Possesses Gold, Location of Wumpus*; the set of actions: A = {*North, South, West, East, Grab*}; the set of observations: Θ = {*Smell, Glitter, Both, Neither*}. Furthermore, both, actions and observations are deterministic. The reward function, R, is defined below:

$$R = \begin{cases} 1000 & \text{if agent arrives at the pre-defined} \\ & \text{location with the gold;} \\ -10,000 & \text{if the agent collides with a Wumpus;} \\ -1 & \text{otherwise.} \end{cases}$$

For each of the four emotional states introduced previously, a specific planning model is engineered. Each planning model is tailored to produce a plan that is distinctive of the corresponding emotion.

Contented The corresponding planning model will consume the maximum possible time in computing an optimal plan. The agent's decision-making process is described by the planning model, $P^c = P$, where P was defined above. The plans generated by a contented agent are similar to those generated by an affectless agent.

Fear The corresponding planning is myopic, solving the associated model over a reduced time horizon. The planning model is defined as $P^f = (S, A, \Theta, T, O, R, H^f)$; where $H^f < H$; and all other parameters retain their original meaning in P.

Panic The planning exhibits myopic behavior. Furthermore, the sole intention is to move away from the existing threat quickly. This disposition leads to a shortening of the planning horizon, and a reduction of the action-space to include only those actions that cause movement. The decision-making process is defined as $P^p = (S, A^p, \Theta, T^p, O^p, R, H^p)$ where $A^p = \{north, south, west, east\} \subset A$; $H^p < H$; T^p and O^p reflect the changed action set; and the remaining parameters retain their original meaning in P.

Elation The agent attempts to reach the gold location as quickly as possible. In the process, it ignores any existing threats, and blindly pursues the gold. This translates into a shortening of the planning horizon, and a transformation of the reward function to reflect its altered preferences. The planning model is defined as $P^e = (S, A, \Theta, T, O, R^e, H^e)$ where

$$R^e = \begin{cases} 1000 & \text{if agent grabs gold;} \\ -1 & \text{otherwise.} \end{cases}$$

$H^e < H$; and the remaining parameters retain their original meaning in P.

Performance Evaluation

In this subsection, we report on preliminary empirical results in support of our hypothesis that affect-based plans are close approximations of exact plans. We also compared the performance of an agent exhibiting several emotions, with that of an agent exhibiting a single emotion. Additionally, we varied the dynamism of the problem domain, and observed the performances of the different agents.

The performance of an agent is measured as the sum of rewards accumulated during a *run*. A run consists of an agent starting at some random location in a randomly generated Wumpus World, and attempting to reach the pre-defined location with the gold while avoiding the Wumpus. We experimented with three different agents: *Agent 1* which is the baseline program that lacks any emotional capability, and generates its plans using the exact planning model. Suppose that during its deliberation, *Agent 1* consumes time t_{a1} before coming up with the next best action; *Agent 2* which is the test program utilizes the FSM in Fig 3 that captures the emotional state transformations. Let its average deliberation time be t_{a2}; and *Agent 3* which remains in a constant **Panic** state. Let *Agent 3*'s deliberation time be t_{a3}. We observed the following ascending relationship between the deliberation times of the agents: $t_{a3} < t_{a2} < t_{a1}$

Our testing method consisted of isolating a sample of 16 randomly generated worlds from the population, fixing a speed of movement for the Wumpus, introducing the agents in the sample worlds, and noting the sum of rewards accumulated during each run of the agent.

In Fig. 4, we show the means with error bars (95% CI) of rewards accumulated by each of the 3 agents during their runs in the sample worlds. These results were compiled using the following observations: t_{a1} = 18 secs, t_{a2} = 6 secs, and t_{a3} = 2 secs. We used Wumpus speeds (t_w) of 18 secs, 9 secs, and 2 secs.

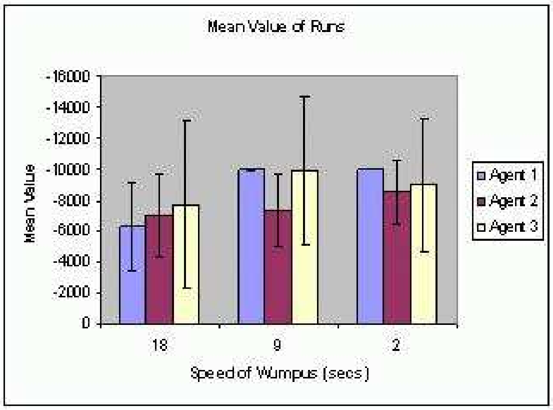

Figure 4: Histogram comparing performances of the agents for various Wumpus speeds.

When t_w=18 secs (no time pressure), the affective agent (*Agent 2*) does only marginally worse than the affectless agent (*Agent 1*). This result suggests that our affect-based approximation method produces plans that are near-optimal. Interestingly, when we increase the dynamism of the world (t_w=9 secs), *Agent 2* performs significantly better than *Agent 1*. Since $t_w < t_{a1}$, *Agent 1* gets killed by the Wumpus frequently. In contrast, *Agent 2*'s fast emotional responses prevent it from being killed frequently. Interestingly, *Agent 3* whose responses are also fast, performs poorly. One reason for its poor performance is the sub-optimal quality of the plan it constantly generates, causing it to take far more steps (and sometimes even get stranded in a corner) to reach the goal. On further increasing the speed of the Wumpus, no agent performed significantly better.

Our experimental results suggest that approximating plans using affective behavior does not significantly compromise the quality of these plans. Additionally, since the approximated models have smaller sizes, the corresponding plans consume less space and are generated in less time. However, the resulting computational savings may be somewhat offset by the increased number of steps that the sub-optimal plans may induce. In general though, we observed that an affective agent completes its run in less time, than an affectless agent.

Discussion and Future Work

In this paper, we adopted the normative method of decision-theoretic planning, for modeling emotional behavior. We showed the ability of POMDPs to generate plans characteristic of many discrete, pure emotions, simply by tuning the input parameters. Furthermore, we adopted finite-state machines as fast and memory-less computational devices for modeling emotional transitions. Our experiments show that the combination of a fast meta-level emotion transition model, and object-level planning model is capable of producing a variety of human-like affective behaviors in an agent. Rather than tediously crafting subjective interpretations of emotional behaviors, researchers can avail of principled frameworks such as POMDPs to produce emotional behavior. Additionally, our research shows that it is possible to produce affective behavior that closely approximates ”rational” affectless behavior.

In our current implementation we modeled only basic emotions. However, we believe the framework to be expressive enough to model emotions of different granularities, and mixtures. Additionally, several problem domains such as dialog management and robot navigation exhibit conditions that may elicit emotions and represent potential testbeds for emotional architectures such as ours. Ongoing work involves experimentation with one such real-world problem domain.

Acknowledgement

We thank Christine Lisetti for discussions related to the paper. This research was supported by ONR grant N00014-95-1-0775, and by the NSF CAREER award IRI-9702132.

References

Cassandra, A. R.; Kaelbling, L. P.; and Littman, M. L. 1994. Acting optimally in partially observable stochastic domains. In *Proceedings of the Twelfth National Conference on Artificial Intelligence*.

Damasio, A. R. 1994. *Descartes Error : Emotion, Reason, and Human Brain*. New York, NY: Gosset Putnam Press.

Elliott, C. D. 1992. *The Affective Reasoner: A Process Model of Emotions in a Multi-agent System*. Ph.D. Dissertation, North-Western University.

Fellous, J.-M., and Hudlicka, E. 1996. http://emotion.salk.edu/emotion/emores/compai/compai.html. Website.

Gmytrasiewicz, P., and Lisetti, C. 2001. Using decision theory to formalize emotions for multiagent systems. In *Workshop on Game Theoretic and Decision-theoretic Agents (ICMAS 2001)*.

Gratch, J. 1999. Why you should buy an emotional planner. In *Workshop on Emotion-based Agent Architectures (EBAA)*.

Lovejoy, W. S. 1991. Computationally feasible bounds for partially observed markov decision processes. *Operations Research* 39(1):162–175.

Papadimitriou, C. H., and Tsitsiklis, J. N. 1987. The complexity of markov decision processes. *Mathematics of Operations Research* 12(3):441–450.

Picard, R. W. 1995. Affective computing. Technical Report 321, MIT.

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

Trakhtenbrot, B. A., and Brzdi, J. 1973. *Finite Automata: Behavior and Synthesis*.

Vel'asquez, J. 1998. Modeling emotion-based decision-making. In *Fall Symposium on Emotion and Intelligence: The Tangled Knot of Cognition*, 164–169.