Point of View: Modeling the Emotions of Others*

Clark Elliott
The Institute for the Learning Sciences
Northwestern University
1800 Maple Avenue
Evanston, Illinois, 60201
and
The Institute for Applied Artificial Intelligence
DePaul University, 243 South Wabash Ave.
Chicago, Illinois, 60604
email: elliott@ils.nwu.edu

Andrew Ortony
The Institute for the Learning Sciences
Northwestern University
email: ortony@ils.nwu.edu

Abstract
When people reason about the behavior of others they often find that their predictions and explanations involve attributing emotions to those about whom they are reasoning. In this paper we discuss the internal models and representations we have used to make machine reasoning of this kind possible. In doing so, we briefly sketch a simulated-world program called the Affective Reasoner. Elsewhere, we have discussed the Affective Reasoner's mechanisms for generating emotions in response to situations that impinge on an agent's concerns, for generating actions in response to emotions, and for reasoning about emotion episodes from cases [Elliott, 1992]. Here we give details about how agents in the Affective Reasoner model each other's point of view for both the purpose of reasoning about one another's emotion-based actions, and for "having" emotions about the fortunes (good or bad) of others (i.e., feeling sorry for someone, feeling happy for them, resenting their good fortune, or gloating over their bad fortune). To do this, agents maintain Concerns of Others representations (COOs) to establish points of view for other agents, and use cases to reason about those agents' expressions of emotions.

Introduction
The Affective Reasoner is a program that reasons about the emotions of agents in a simulated multi-agent world. Agents in this world are given a simple emotional life, consisting of twenty-four emotion classes and approximately 1200 different expressions of these emotions. They are given idiosyncratic personalities which allow them to have different interpretations of identical situations in their own and different response tendencies for identical emotional states. In addition to "having" and "expressing" their own emotions, agents also reason about, and have emotions in response to the emotions of other agents.

To date the Affective Reasoner has been used primarily to simulate the interpersonal interactions of taxi drivers with one another and with their passengers in an instantiation called TaxiWorld [Elliott and Ortony, 1992a]. However, experiments can be and have been conducted in other simple areas of interpersonal relations including the one used in a story about a rookie quarterback discussed later in this paper. In fact, all the Affective Reasoner requires is a domain in which emotions can arise, provided the domain can be represented using a discrete-event simulator.

We consider reasoning about emotions to have both a strong-theory component, from emotion-inducing situations to the emotions they engender, and a weak-theory component, from emotions to actions. To represent the different characteristics of individual agents we accordingly break our rudimentary personality representations into two parts. The first of these, the interpretive personality, is used for determining whether some event, act, or object is of concern with respect to the goals, standards, or preferences (hereafter GSPs) of some agent. Rule-based, strong-theory reasoning is used to tie interpretations of situations to emotional states. The second part of each personality representation is the manifestative personality component which is used for determining how an agent will act, or "feel," in response to these emotional states. This component contains a set of temperament traits which are dynamically activated so as to tune an agent's action tendency profile. Case-based, weak-theory reasoning is used to reason back from observed actions to an agent's emotional states. After a brief introduction to these two constructs, this paper focuses on how agents form, and maintain, internal representations of the personalities of other agents.

To understand how the interpretive personality works, suppose a basketball player on a team misses a shot at the buzzer in a game so that the team loses by one point. One observing agent might be unhappy that his team had lost the game, whereas another observing agent, being a fan of the victorious team, might be happy that her team

---

*Preparation of this article was supported in part by National Science Foundation grant IRI-8812699, and in part by Andersen Consulting through Northwestern University's Institute for the Learning Sciences. Citations appear in bib.
had won. This is an example of how the same situation (the missed shot at the buzzer) can be interpreted either as one of having achieved, or of having failed to achieve, a goal. There are other alternative or additional interpretations that might be made. For instance, an observing agent might have admiration for the player for making a heroic attempt to win the game, an example of mapping the act into the perception of a principle being upheld, or an observing agent might simply be enthralled by the beauty of the move to the basket, an example of mapping the move into the perception of an appealing object, and so on.\(^1\)

One important aspect of the Affective Reasoner’s representation of agents’ interpretive personalities is that it treats them as modular data structures, representing them as inheritance hierarchies of frames.\(^2\) Any leaf node frame (i.e., one of the frames used to match and interpret events, acts, and objects in the world), and its inheritance path, may be combined with any other leaf node frame, and its inheritance path, when forming an interpretive personality representation. As a consequence, multiple and even conflicting interpretations of a situation may arise for some agent. This not a limitation of our system, but rather a requirement of emotion representation. For example, we may see a woman sad over the death of her favorite aunt, and yet relieved because she knows that her inheritance will be her financial salvation. So, the same situation can give rise to conflicting emotions which, in turn, could even give rise to similar actions (people can cry from sadness or joy). To sum up, the interpretive component of personality representations in the Affective Reasoner exploits the fact that individuals have different goals, standards, and preferences (GSPs) by allowing these different concerns to be the basis for interpreting emotion-inducing situations.

We turn now to a brief account of the manifestative personality component. Suppose that some agent is feeling proud about some admirable act she has performed. If she is a quiet type, she may simply manifest this emotion through a quiet somatic response (e.g., a feeling of general well-being). If she is verbally inclined, she may express her pride through verbal communication with another agent (e.g., telling someone about how proud she is). If she tends to be manipulative she may manifest her pride by attempting to modulate the emotions of others (e.g., seeking to have them admire her by calling attention to her praiseworthy act). An agent in the Affective Reasoner can have many of these different temperament traits active at the same time. Together these give the agent its idiosyncratic manifestative personality.

There is at best only a loose mapping from emotional states to particular actions; few actions are unambiguously indicative of the emotions that initiated them. For example, people can smile because they are happy, because they are gloating over the misfortune of an adversary, or because they are pretending not to be afraid. In addition people can express the same emotion in many different ways (e.g., frowning because they are angry and want it to be known, or smiling because they are angry and do not want it to be known). Complicating this is the fact that individuals have different emotion expression styles, and are affected by moods as well. Because of such ambiguities we have chosen a case-based approach to reasoning in this portion of the emotion domain.

People model one another’s points of view. This enables them to both explain and predict the responses of others to situations, and to have emotions regarding the fortunes of those others. In the Affective Reasoner these points of view are captured in the two-part rudimentary personality representation just discussed. To model such points of view, agents can maintain internal models of both the concern structures of other agents (i.e., their interpretive personalities), and their response action structures (i.e., their manifestative personalities). A model of the former allows the agent to make inferences about emotions other agents are likely to have in certain situations. A model of the latter allows the agent to make explanatory inferences about antecedent emotions when seeing other agents acting in a certain way. These two aspects of capturing and maintaining knowledge about other agents are discussed in the next two sections.

**Representing the Concerns of Others**

For an agent to understand how another agent is likely to construe a situation, he or she must see that situation from the other agent’s point of view. Because in the Affective Reasoner an agent’s interpretations of the world are derived from its GSP database it follows that an observing agent must also have some internal representation of the observed agent’s GSPs. This knowledge is captured in data structures known as Concerns-of-Other (COO) databases. They are, essentially, imperfect copies of other agents’ GSPs, and represent their concerns as modeled by the observing agent.

Thus, in addition to the GSP database representing an agent’s own concerns, a COO database can be maintained for each other agent the observing agent is modeling. Using the same machinery that causes emotions to be generated by the system for some agent when a situation is filtered through that agent’s interpretive personality (i.e., its own GSP database) that agent may instead filter the situation through the internally modeled interpretive personality of the observed agent (i.e., its COO representation for that other agent) to see the situation from the other agent’s (supposed) point of view.

A perfect COO representation, of course, would be an exact duplicate of the observed agent’s GSP database, and would always lead to the same interpretations that the observed agent has. However, because as discussed above, GSPs are built out of interpretation modules (i.e., frames), partial COOs can be created incrementally. Even though they are imperfect, these partial represen-
tions are useful because they allow the observing agent to interpret at least some situations correctly. For example, Harry might know that Sarah is a passionate Cubs fan, and that if the Cubs lose she will be upset, and yet not know more about her. Still, if the Cubs do lose, and Sarah is jumping up and down, then Harry probably knows why.

Because the Affective Reasoner was developed as a general research platform, several options are available with respect to the establishment of COO databases. They may be constructed at start-up time as part of the initial composition of agents, or they may be learned as the simulation proceeds and as agents come into contact with one another. In the former case a number of difficulties are avoided, such as having to work out the details of when agents are permitted to observe each other. In the latter case, many useful knowledge acquisition issues can be studied. For example, if the system is to be used to store knowledge about interesting agents and to study emotions that arise when they interact, then the domain-analysis investment required for setting up the COO learning process will have little return. On the other hand, if one is studying user-modeling from the domain-analysis viewpoint, then the concern structure for the agent. Remove it from the COO.

Collecting Construal Frames for COOs

When the Affective Reasoner is set up so that agents learn about one another’s concerns through interactions, COOs are built up incrementally by locating and incorporating construal frames that seem to explain another agent’s emotional states in response to observed situations. For example, when Harry sees that Sarah is always unhappy when the Cubs lose he might infer that Sarah construes some aspect of this situation as blocking one or more of her goals. Harry might then try to determine exactly which goals are involved: is Sarah a Cubs fan, or has she just been betting on them to win? In the following algorithm, which describes this process, we assume that the observing agent has already discovered the emotion(s) present in the other agent. The observing agent now attempts to explain that emotion in terms of the eliciting situation, and possible construals of that situation. To do this the observing agent first consults its COO for the observed agent, and then, if necessary, a set of databases containing alternate construal frames (see Defaults). Here is the algorithm for incrementally building COO representations:

1. Locate the Concerns-of-Other representation (COO) for the observed agent. If one does not exist or it does not contain an interpretation for this type of eliciting situation then go to 5,

2. Filter the situation through the COO, producing an emotion. If this emotion is the same as the emotion that actually was present in the observed agent then the COO has probably given a correct interpretation of the eliciting situation so go to 8,

3. Because the interpretation produced by the COO is incorrect (i.e., the emotion based on the COO’s interpretation of the eliciting situation does not match the emotion known to be present in the agent whose concerns it is supposed to represent) the construal frame used to make the interpretation should not be part of the concern structure for the agent. Remove it from the COO.

4. Mark the construal frame ineligible for this agent. Eligible frames are those frames that can produce interpretations for this type of situation. Ineligible frames are previously eligible frames which have been found to produce incorrect interpretations.

5. Search through the global (or default) database for the next eligible interpretation of this situation,

6. Evaluate the situation using the new interpretation as a filter. If the resultant emotion is not the same then go to 4,

7. Add the construal frame to the COO,

8. Generate an explanation based on the current construal frame.

Once a COO has been established for some other agent it can be used for two purposes. First, it is now possible for an observing agent to have emotions based on its perceptions of the fortunes of the second agent. In the Affective Reasoner this may come about if the agents are in one of the following three (possibly only unidirectional) relationships: friendship, animosity, and empathetic unit. For example, if the observing agent knows that a second agent is a Cubs fan, then if they are friends the observing agent can feel sorry for the second agent when the Cubs lose. On the other hand, if they are adversaries then the observing agent can gloat when the Cubs lose. Lastly, should the bond between the two agents be so strong in some situation that the observing agent temporarily takes some of the second agent’s concerns on as its own then an empathetic unit has been formed. The observing agent will temporarily suspend its own GSP database, using its COO for the second agent to generate direct emotions instead. Note that even in this case the observing agent might actually be wrong about the import of a particular situation for the observed agent, since it is the observing agent’s representation of the observed agent’s concerns that is being used to generate emotions, not the actual GSPs of the observed agent. The second

---

5 These three relationships have a very specific meaning here. Friendship means that an agent will tend to have similarly valenced emotions in response to the emotions of another agent. Animosity means that the emotions will tend to be oppositely valenced. Empathetic unit means that the particular situation is seen “through the eyes” of the other agent, so that the emotions are experienced as the observing agent’s own.
use of COOs is that once they are established it is possible to explain, and sometimes predict, the emotional responses of other agents based on the eliciting condition rules, as in the previously discussed case of Sarah the fan, which opened this section.

Satellite COOs

Agents in the Affective Reasoner may be set up to do more than model the simple, direct concerns of agents whom they observe. The need for more complex internal models is illustrated by the following story, on which one of our simulation runs was based.

A rookie quarterback is, as usual, sitting on the bench during a football game. His brother and a woman friend of his brother are in the stands. Suddenly the starting quarterback goes down with a knee injury. The woman smiles because she is happy for her friend who’s brother will now be placed in the game.

When reasoning about the emotions that arise in this situation we must consider the following sets of concerns and relationships:

1. The actual concerns of the rookie quarterback, i.e., his GSPs. Implied in the story is that he will be pleased about achieving a getting-to-play goal.

2. The supposed concerns of the rookie quarterback as represented by his brother (i.e., the COO representing the brother’s beliefs about the GSPs of the rookie quarterback).

3. The relationship between the rookie quarterback and his brother. Specifically the friendship relationship, or even an empathetic unit relationship.

4. The friendship relationship between the woman and the brother.

5. The supposed concerns of the brother as represented by the woman. This must include, recursively, her supposed supposed concerns of the brother for the quarterback as well, and the supposed empathetic relationship between the brother and the quarterback. In other words, the woman must have a belief that the brother will believe that the rookie quarterback will be happy about the starting quarterback’s injury. Furthermore, she must believe that the relationship between the brother and the rookie quarterback is such that a positive outcome for the rookie quarterback maps to a positive outcome for the brother.

Because the story gives no clues as to the emotional states of either the quarterback or his brother it should be obvious that neither the actual concerns of the quarterback nor those of his brother are necessary for understanding the episode. To make this clear, consider the following possible continuation to the story:

...But the smile quickly fades when the brother says, “Oh no, I told him he shouldn’t have drunk that case of beer at lunch.”

Clearly the woman’s beliefs leading to emotional states and action expressions of those states are not dependent upon all of the actual facts. Similarly, even if her understanding of the facts is correct, this still does not mean that her emotions have to be in line with them. Consider the following alternate continuation to the story,

Instead of playing the rookie quarterback, however, the coach puts in a third-string quarterback. The woman, who unbeknownst to the brother had consumed a case of beer with the rookie quarterback at lunch and was sworn to secrecy about it, is relieved. The brother, however, feels terrible for the rookie quarterback because he will not get to play. Consequently the brother is very unhappy. The woman is sorry to see him in this state.⁶

In this case the woman knows that the brother’s beliefs are incorrect, and she does not share them, but she still is capable of having emotions based on the brother’s fortunes, which in turn are based on those incorrect beliefs.

It can be seen then, that for observing agents to represent the fortunes of another agent, to have emotions regarding those fortunes, and to interpret their actions with regard to those fortunes, the observing agents must not only be able to represent the concerns of the observed agents, but sometimes must also be able to model the observed agents’ own representations of the concerns of those important to them. In the Affective Reasoner we capture such knowledge in a second-level set of COOs, called satellite COOs. These are used in conjunction with a set of supposed relationships between the observed agents and those whom the satellite COOs are intended to represent. For example, if Harry believes that Sarah is in relationships with both Joan and Eva, then Harry will maintain a COO for Sarah and satellite COOs for Joan and Eva as seen through Sarah’s eyes.

The three distinct interpretive personality representations used by the system are all structurally and functionally the same. It does not matter whether the representation is to be used as a system-level GSP or as an agent-level COO or satellite COO. The emotion machinery that is applied to GSPs for the generation of direct emotions may also be applied to COOs used for the generation of the fortunes-of-others emotions and to satellite COOs used for representing an agent’s beliefs about another agent’s beliefs.

The process for making use of each of these GSP and COO databases is also the same in all cases. The elicitation...
ing event, act, or object is filtered through each respective GSP or COO database to produce an interpretation with respect to the antecedents of emotions. In the direct case, the result is emotions that the system generates for the agent. In the once-removed case the result is an interpretation based on imagining what it is like for the other agent (possibly incorrectly), which, when combined with a relationship, may yield a fortunes-of-other emotion in the observing agent. In the twice-removed case, when combined with beliefs about relationships, the interpretation may lead to a belief about the emotional state of the other agent, which in turn may also lead to fortunes-of-others emotions.

**Defaults**

In some cases little may be known about another agent. Nonetheless, one may feel sorry for a stranger, and one certainly may wish to explain the actions of strangers. Thus we must give agents a mechanism by which they may still reason about the emotions of other agents, even if nothing specific is known about those agents.

Because, for the purpose of generating emotions in the Affective Reasoner, one GSP database is as good as another, and because even the component construal frames may be mixed at will, we may use a system of defaults for reasoning under uncertainty. Two of these are rather obvious. The first is a system-wide default GSP which corresponds to the knowledge source one might consult in addressing such questions as How might a typical agent interpret this situation? The interpretations produced by this default database are useful when producing explanations such as When someone is hit they get mad and Losing money increases distress. The next obvious default GSP is an agent's own GSP database, which corresponds to the knowledge source one consults when asking How would I interpret and react to this situation? The resulting emotional states can then be projected onto the other agent.

In addition to these two defaults, observing agents in the Affective Reasoner might also make use of an extant COO for some third agent, provided that one exists and that it is consistent with what has already been observed about the new modeled agent. As long as this COO is suitable it remains in use. When the COOs diverge (i.e., when one of the construal frames in the existing COO is found to be incorrect for the new modeled agent) then a copy of the existing COO is made and the offending construal frame is removed. This becomes the current representation of the COO for the new agent. The use of COOs in this manner corresponds roughly to reasoning that because Agent A seems just like Agent B, then assume they are alike in all ways until learning differently.

Nor are we restricted to using only one COO when searching for an explanation. For agents then, the order of precedence is as follows: (1) search through the COO for the other agent to look for an interpretation of some eliciting situation; if there is none, or it is found to be in error, then (2) search through a COO for some other agent that appears to be similar to this one, if one exists; next, (3) search through the system default GSP database to see how a typical agent would interpret the situation; failing in this, then (4) search through one's own GSP database to see how I might interpret this situation, and lastly (5) search through the global shared database of construal frames for all possible interpretations of the situation.

**Representing the Action Tendencies of Others**

In the Affective Reasoner observing agents attempt to make sense of the way observed agents respond to situations that arise by using emotion-specific knowledge to limit the search space. Something happens, this gives rise to an emotion, or set of emotions, and these in turn give rise to emotion-induced reactions. So far in this paper we have discussed representations that agents keep of others' concerns. This knowledge may be used in two ways. First, reasoning backwards it may be used abductively to explain how an agent sees the world. Second, once established for an agent, it may be used deductively to predict what emotions that agent may have in response to future situations. By contrast, in this section we discuss an alternate source of knowledge which allows agents to observe features of a situation and some agent's response to that situation (i.e., an emotion episode), and reason back to an emotion category using past cases. This roughly corresponds to lines of thought such as Is the agent smiling? He is probably happy. Is he shaking his head? He may be reproachful.

With respect to such reasoning the Affective Reasoner may be run in three major modes, two of which will be discussed here. In the first mode agents make use of a heuristic classification component based on the Protos program developed by Bareiss [Bareiss, 1989]. Using their own set of cases drawn from past experience, agents make determinations about which emotion is present based on the features in an eliciting situation and on the observed agent's responses to that situation. In this mode agents are free to ask questions of the "teacher" (in this case the user) to acquire knowledge about the relationship of features to emotion categories. This knowledge includes such relationships as highly correlated, mutually exclusive, and so forth. For example, in this mode an agent may decide that another agent who is shouting is expressing anger, because the present case reminds the agent of a previous case of anger that had that feature. Should there be additional features present that the agent does not understand, or if the classification is made incorrectly, then the agent, through Protos, asks for domain knowledge from the teacher. Suppose for example that in addition to shouting, the observed agent was also represented as shaking its fist in the air. In this case the observing agent might ask for an explanation of how this feature relates to anger, and would then either update the present exemplar, and possibly the set of reminders that lead to its selection, or create an entirely new exemplar if the new one is sufficiently different from the old one. In the second mode Protos is used only to make
classifications for agents using an existing case base. No knowledge is acquired and the case base remains static. This mode is useful for running simulations without input from the researcher, where the case bases have reached a certain level of maturity. As with COOs, case bases may be established at start-up time as part of the initial composition of agents, they may be acquired entirely as part of the current simulation, or they may be established at start-up time and then enhanced as part of the current simulation.\footnote{In the third mode, classification of episodes is bypassed entirely, and observing agents are simply informed directly by the system what emotion(s) the observed agent was experiencing.}

In using this scheme for capturing weak-theory knowledge about the features of emotion episodes one might wish to store cases for each other agent, just as one maintains distinct COO representations for each other agent. Such knowledge would be equivalent to knowing that Tom shakes his fist when he gets angry, and Harry always wrings his hands when he gloats. We have not taken this approach in the Affective Reasoner. Instead, each agent maintains a single case base. This is the equivalent of having seen a case of anger before where the agent was shaking his fist. However, since the name of the agent is counted among the features of an episode which may be recorded by Protos, it is nonetheless still possible to capture some agent-specific knowledge about the expression of emotions through actions.

**Conclusion**

We have described components of a system, the Affective Reasoner, in which a number of simulated agents, each with their own “personality” interact. These agents respond to situations in their world in emotion-like ways, but they also respond to what they take to be the emotions of others. We conclude with a brief discussion of some caveats pertaining to this work, particularly with respect to relation between the “emotions” of our simulated agents and emotions as we know and experience them as human beings.

We do not claim that our simulated agents “have” or “feel” emotions. Such a claim would be uninterpretable at best and nonsense at worst. Human emotions are comprised of interacting cognitive components, behavioral components, physiological components, and phenomenal components. The Affective Reasoner only seeks to model (aspects) of the cognitive and behavioral components, and is best considered as an attempt to generate the ingredients required to reason about emotions rather than as an attempt to produce emotions. When we humans speak of having or feeling emotions, we are implicitly focusing on the phenomenal and, perhaps therefore by necessity, the physiological components. Human emotions are not cold cognitions leading to detached behaviors, they are hot cognitions integrated with (sometimes dysfunctional) behaviors. It is their physiological and phenomenal qualities that give them their special “feel,” and we neither attempted to, nor even would know how to begin to model these aspects of emotions. Rather, we concentrated on the more manageable aspects, namely the cognitive and behavioral ones. Human emotions are not randomly related to how people perceive their world. There is some order, and the Affective Reasoner seeks to capture some of that order by embodying a strong theory of the relation between construed situations and emotions. Similarly, whereas the linkage between human emotions and actions may be somewhat weak, still, the relation is not arbitrary. The Affective Reasoner incorporates a weak theory of emotion-to-action relations in an effort that we claim to be little more than a first step. Thus, the affective Reasoner could be viewed as a system that attributes emotions to its agents by reasoning, rather than as a system in which emotions simply arise in agents. In this paper, we have focused on this reasoning process at one level of embedding in that we have described what emotions the system attributes to its agents when they are interpreting the situated behaviors of others as being emotion-induced.

Trying to build systems that understand anything at all about emotions is not easy. In order to prevent the system from becoming unmanageable, consideration of many important aspects of emotions and emotion-related behavior had to be postponed. The most obvious of these is the omission of considerations of emotion intensity which, in future efforts, is likely to be handled using qualitative reasoning techniques. In fact, as experimenters and observers of human emotions we frequently use intensity-relevant inferences to predict and explain behaviors. For example, given knowledge about particular individuals and their emotional “styles” we can infer whether or not they will react with intense or mild emotions in particular types of situations. We can make similar inferences about emotion-induced behaviors because we know that generally speaking, mild emotions do not give rise to extreme behaviors, and so on. We hope to address these and other limitations in future work on this project.

**References**


