A step toward irrationality: using emotion to change belief

Stacy Marsella
University of Southern California
Information Sciences Institute
4676 Admiralty Way, Marina del Rey, CA 90292
1-310-448-8407

marsella@isi.edu

Jonathan Gratch
University of Southern California
Institute for Creative Technology
13274 Fiji Way, Marina del Rey, CA 90292
1-310-448-0306

gratch@ict.usc.edu

ABSTRACT

Emotions have a powerful impact on behavior and beliefs. The goal of our research is to create general computational models of this interplay of emotion, cognition and behavior to inform the design of virtual humans. Here, we address an aspect of emotional behavior that has been studied extensively in the psychological literature but largely ignored by computational approaches, emotion-focused coping. Rather than motivating external action, emotion-focused coping strategies alter beliefs in response to strong emotions. For example an individual may alter beliefs about the importance of a goal that is being threatened, thereby reducing their distress. We present a preliminary model of emotion-focused coping and discuss how coping processes, in general, can be coupled to emotions and behavior. The approach is illustrated within a virtual reality training environment where the models are used to create virtual human characters in high-stress social situations.

Categories and Subject Descriptors

I.2 [**Computing Methodologies**]: Artificial Intelligence

General Terms

Algorithms, Design, Human Factors

Keywords

Emotion, Coping Strategies, Believability, Cognitive Modeling

1. INTRODUCTION

Imagine yourself as a young lieutenant in the U.S. Army on your first peacekeeping mission. You've been ordered to reinforce a sister unit, designated Eagle 1-6, that has encountered civilian unrest while supporting a U.N. weapons inspection team. Anxious to assist them, you arrive at a rendezvous point expecting the rest of your unit to be assembled and ready to proceed. Instead, you find your troops in disarray. Your platoon sergeant seems upset. Smoke is rising from one of your vehicles and a civilian car. A child lies on the ground, surrounded by a distraught woman

(his mother?) and a medic from your team. Dismayed, you ask your sergeant what happened. He pauses, casts an angry glance at the woman and responds, "they rammed into us sir. They just shot out from the side-street and our driver couldn't see them." How do you interpret this response? Take it at face value: was this intentional? Were they trying to prevent us from reaching the inspection site? Or is he (merely) being defensive, deflecting any potential blame for the delay. Such ambiguity is endemic in emotionally charged social interactions and learning how better to recognize and handle it would benefit not just young lieutenants but managers, spouses, parents; basically anyone that must interact with other people.

Our interest in emotion stems from a desire to create learning environments where one can experience high-stress social situations in the relative safety of virtual reality. Role-playing has long been used for psychotherapy, to teach social skills, or to practice decision-making under stress and researchers have begun to consider how to automate these methods (e.g., Marsella, Johnson & LaBore, 2000). Our focus is to support automation by creating virtual humans to play a variety of social roles and engage in wide ranging interactions with a human interactor via natural language. The interactive and emotionally charged nature of such social simulations presents serious challenges for agent design. In addition to the standard problem of artificial intelligence, these virtual humans must incorporate emotional models that are general and flexible enough to respond in reasonable ways to whatever circumstances the interactor is allowed to create. They must identify plausible emotions to express, model the typical coping strategies people use in emotional situations, and capture the dynamics of how emotions unfold over time.

The artificial intelligence community has made steady progress in creating intelligent agents that convey a sense of emotion. Approaches can be roughly characterized as being either communication-driven or appraisal-driven. Communication-driven methods choose emotional expression based on the agent's intended impact on some other agent or person. For example, Poggi and Pelachaud (1999) use facial expressions to convey the performative of a speech act, showing "potential anger" to communicate that the agent will be angry if a request is not fulfilled. In contrast, appraisal-driven theories focus on the agent's own emotional response and the apparent evaluative function it plays in reasoning. Appraisal theories view emotion as arising from some assessment of an agent's internal state vis-à-vis its external environment (e.g., is this event contrary to my desire?). Such appraisals can be used to guide decision-making and behavior selection or as a basis for communicating information about the agent's assessment, though not in the intentional way viewed by communicative models.

Appraisal methods are useful for giving coherence to an agent's emotional dynamics that can be lacking in purely communicative models. This coherence is essential for conveying a sense of believability and realism (Neal Reilly, 1996). These approaches are complementary, though few approaches have considered how to combine them into a coherent approach (cf. Marsella et al., 2000). In the discussion that follows, we will start with the appraisal perspective and yet show how an agent's coping mechanisms provide a basis for communication driven emotions.

The primary contribution of this paper is that it addresses one aspect of emotional behavior that has been studied extensively in the psychological literature but has been largely ignored by computational approaches. Emotions act as powerful motivators, but programs that attempt to model this function have largely focused on the problem of selecting actions to perform in the world. In contrast, the psychological findings indicate that, in addition to action selection (which has been termed problem-focused, or task-oriented, coping), people employ other strategies for dealing with strong emotions, termed emotion-focused coping (Lazarus, 1991) and suppression (Parkes, 1984; Wells and Matthews, 1994). Rather than acting on the world, emotion-focused coping works mainly by acting on an individual's beliefs to discharge negative or reinforce positive emotions. For example an individual may alter beliefs about the importance of a goal that is being threatened. The angry behavior of the sergeant at the accident site could be seen as a form of emotion-focused coping: dealing with guilt by placing blame on some other agent. Suppression is an avoidance strategy that attempts to avoid thinking about the problem. Our primary contribution is a preliminary model of emotion-focused coping, a functionality that is particularly important to model in the type of stressful social scenarios that are typically used in role-playing. We also discuss a range of extensions of our prior work (Gratch & Marsella, 2001) including (a) using the appraisal model to infer the emotions of others, (b) how coping mechanisms are tightly coupled to appraisals, and (c) how appraisals plus the coping mechanism provides a basis for a model of emotional expression that relates expression to both an agent's underlying emotion and intentions. We also briefly touch on how we currently relate coping mechanisms to the agent's dialog and the role of personality in biasing coping mechanisms.

2. THEORETICAL FRAMEWORK

Our work is motivated by psychological theories of emotion that emphasize the relationship between emotion and cognition (Scherer, K. 1984; Frijda, 1987; Ortony, Clore and Collins, 1988; Lazarus 1991), and is realized in terms of data structures and inference mechanisms that are particularly amenable to intelligent agent design. Figure 1 illustrates the basic outline of the model, which consists of two basic processes: appraisal and coping. Appraisal is the process that generates emotion by assessing the person-environment relationship (did an event facilitate or inhibit the agent's goals; who deserves blame or credit). Coping is the process of dealing with emotion, either externally (by forming intentions to act in the world) or internally (by changing the agent's interpretation of the situation). These processes interact and unfold over time, modeling the temporal character of emotion noted by several emotion researchers (Scherer, 1984; Lazarus 1991): an agent may "feel" distress for an event (appraisal), which motivates the shifting of blame (coping), which leads to anger (re-appraisal).
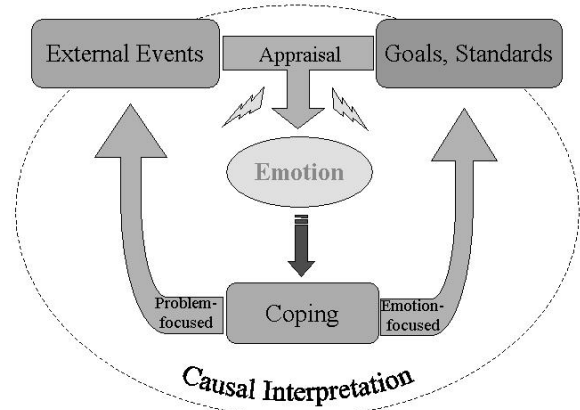

Figure 1: Theoretical framework

Underlying this model is a data structure, the *causal interpretation,* that facilitates the derivation of appraisals and coping strategies. The causal interpretation is essentially a causal model (analogous to the notion of Pearl, 2002) but cast from the agent's perspective (different agents may have different causal models based on their interpretation of the situation). The interpretation encodes a representation of events, world states, the causal relationship between states and events, as well as beliefs (truth values and probabilities over states), desires (utilities associated with states), intentions (over actions), and authority relationships. The interpretation also represents a restricted set of second-order beliefs (e.g., I believe that you desire this state). Reasoning mechanisms, such as planning, dialogue management, or perception, act to modify the causal interpretation (by adding plan steps, updating beliefs, etc.).

This representation has several advantages for modeling emotion. It makes a clean separation between domain-specific knowledge (e.g., specific action definitions, probabilities and utilities) from the domain-independent mechanisms that operate on these representations. It acts as a blackboard architecture, simplifying communication between appraisal and coping from other mechanisms (like planning) that operate on the interpretation. It facilitates reasoning about blame and indirect consequences of action (e.g., a threat to a sub-goal might be distressing, not because the sub-goal is intrinsically important, but because it facilitates a larger goal). It provides a uniform representation of past and future actions (this action caused an effect which I can use to achieve that goal). Finally, as discussed below, it facilitates reasoning about different agents' perspectives (I think this outcome is good but I believe you think it is bad).

Appraisal assesses the agent-environment relationship via features of the causal interpretation (as described in Gratch, 2000 and summarized below). The belief that another agent has caused an undesirable outcome should lead to distress and possibly anger. Coping similarly exploits the interpretation to uncover what features led to the appraised emotion, and what potential there may be for altering or reinforcing the interpretation.

3. ANNOTATED EXAMPLE

The subsequent discussion makes reference to the peacekeeping mission alluded to above and discussed in detail in (Swartout et al, 2001). The basic exercise involves a human platoon leader that must interact with several autonomous agents (his platoon ser-

Figure 2: A scene from the Mission Rehearsal Exercise

Figure 3: Portion of Causal Interpretation

geant, platoon members, and civilian characters). Although different characters in the virtual environment possess different domain knowledge, the sergeant serves as the main interface to the student and has the most comprehensive model of the domain. In addition to emotional reasoning, the sergeant incorporates speech understanding, plan reasoning (Rickel & Johnson, 1999), dialogue management (Traum & Rickel, 2001), natural language generation (Fleischman and Hovy, 2002), speech synthesis, and human figure animation. Figure 2 shows some of the characters in the scenario: the sergeant, the mother and the medic, from left to right.

At startup, the causal interpretation is initialized with knowledge that the platoon was moving to an assembly area to rendezvous with the lieutenant, the mother and squads were driving into the same intersection and that a collision resulted. No one is explicitly deemed responsible for the collision, but it has the effect that the driver has minor injuries and the mother's boy has critical injuries. The mother and the sergeant have indirect responsibility: they are individually responsible for the actions that established the pre-conditions for the collision. The sergeant has several desires provided by the exercise developer and represented as states with associated utilities. The goal of supporting Eagle 1-6 has high positive utility. Injuries to other agents in the environment have negative utility. The sergeant also models the desires of other agents, which can differ from his own (he believes that the mother attributes more dis-utility to the boy's injuries) and which do not necessarily reflect the true desires of the other agents. In the scenario, possible future actions involve different (and potentially conflicting) plans for helping the boy and/or helping Eagle 1-6. Figure 3 illustrates a portion of the causal interpretation associated with helping the boy.

Given the initial interpretation, the appraisal process identifies several instances of emotion that the sergeant should be feeling. The two dominant emotions are distress over the critical injuries of the boy and distress over the minor injuries of the driver. These arise from the fact that some event occurred (the collision) that had effects with negative utility. As no one (as of yet) has
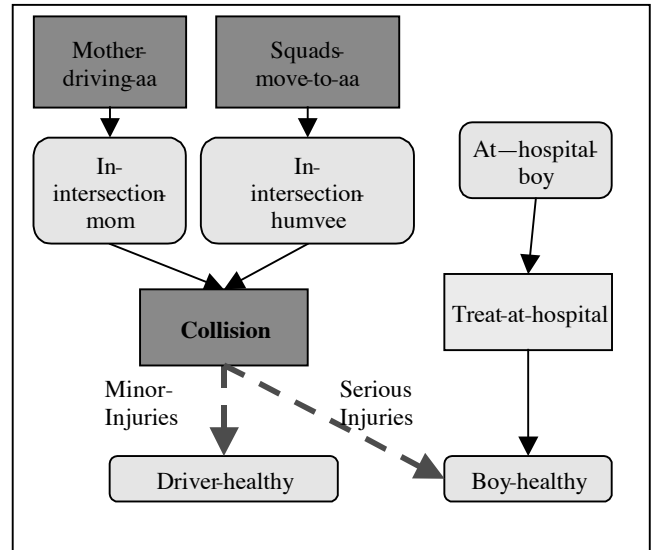
been deemed responsible for the collision, there are no associated feelings of anger or guilt. The sergeant is also distressed that the unit is not currently helping Eagle 1-6.

The sergeant has two (contradictory) plans in memory – to help the child and to help Eagle1-6, each of which is blocked pending authority of the lieutenant – and several emotions arise from these plans and their interactions. The sergeant has hope that the boy will be made healthy and that Eagle1-6 will be supported. The sergeant also has guilt arising from the fact that supporting Eagle1-6 is a viable plan and its execution would require abandoning the boy. In addition to his own emotions, the sergeant infers that the lieutenant is distressed that Eagle1-6 is not being supported and that the mother is extremely distressed about the child.

When the lieutenant arrives, the sergeant updates his belief about the lieutenant's location. Any emotions associated with this belief are brought into focus (in a sense to be discussed below). In this case there is only a small amount of joy associated with the lieutenant's arrival (as his presence is a sub-goal of achieving some of the sergeant's desires). This triggers a change in facial expression but the emotion is of insufficient intensity to require any coping.

When the lieutenant asks, "what happened here," this is recognized as an information request about events in the causal history that occurred at the present location. The request unifies with three events, the squads driving to the assembly area, the mother driving from the side street, and the collision. All emotions associated with these events are brought into focus: the sergeant's distress about the child and the driver, the lieutenant's distress over Eagle1-6, and the mother's distress over the child. Each of these emotions is detailed in an appraisal frame. Figure 3 shows the most intense appraisal frame, the distress over the child's health. In this case, the emotions are sufficiently strong to induce coping. A coping elicitation frame is also created to collect information about potential coping strategies. Domain independent rules note several social factors related to the events in focus: that the person asking the question is the sergeant's superior and that the agent driving the vehicle involved in the collision was under the sergeant's command. They also identify that the mother and

the sergeant are potentially responsible for the accident, in that no one has yet been assigned responsibility and they (or their subordinates) were responsible for events that led up to the collision (the mother and driver are the responsible agents for actions who's effects are preconditions of the collision task).

Three coping strategies are proposed to deal with the intense emotions brought into focus: make amends, accept responsibility, and shift responsibility. Make-amends is a form of problem-focused coping and is proposed if there is an action that could "undo" some negative emotion. As summarized in the underlying appraisal frame, the distress arises from the undesired consequence of a past action. Working backwards from the frame, the make amends strategy determines that the treat-at-hospital(child) action would undo this negative state. If selected, this strategy would form an intention to perform this action.

The accept-responsibility strategy is a form of emotion-focused coping and it is proposed if the agent has potential responsibility for the cause of the emotion. In this case, the sergeant has potential responsibility for the collision because his subordinate, the driver, was responsible for an action that was a precondition of the collision. If selected, this strategy would assert the belief that the sergeant is responsible for the accident, triggering a reappraisal of the collision. Appraisal rules will automatically fire, elaborating appraisal frames associated with the driver and child's injuries with the information that the sergeant is blameworthy. This, in turn, will cause new emotion instances to be created, indicating that the sergeant feels guilty and he believes that the lieutenant and mother will be angry with him.

Finally, the shift-responsibility strategy is a form of emotion-focused coping that is proposed if a superior (power relation) is asking about an event that the agent is potentially responsible for (the collision) and there is another agent that is also potentially responsible (the mother). If selected, this strategy would assert the belief that the mother is responsible for the accident. As in accept-responsibility, this triggers a reappraisal. In this case, the sergeant will feel anger towards the mother and will infer that the lieutenant will be angry as well, and that the mother should feel guilty.

After performing the coping strategy, the sergeant will answer the question. Speech, facial expressions, and gestures are modulated by the coping strategy and the current emotions in focus. The impact on natural language generation is via lexical choice, based on a method proposed by Hovy (1988) and elaborated by Fleischman and Hovy (2002). Gestures and head movements are determined by rules that consider the syntactic structure, the associated semantics and the emotional markup of concepts in the resulting utterance. The BEAT system (Cassell, Vilhjalmsson & Bickmore, 2001) is used to synchronize the gestures with the production of phonemes and visemes.

Currently, we use a simple personality model to assert preferences over these coping strategies. Depending on the traits that were pre-assigned to the sergeant, different strategies are preferred.

4. APPRAISAL

The causal interpretation is critical for uncovering the strong emotions that people experience when faced with this situation. Only by working through the consequences of the sergeant's potential actions will he uncover that helping the boy precludes helping Eagle1-6. Causal reasoning is also essential for propagating the consequences of new events and information through the plan

---

Appraisal Frame: distress332

  Perspective: sergeant

  Emotion-type: Distress

  Expected utility: -59.3 (out of range –100..100)

  Intensity: 59.3

  Type: facilitator

  Annotation: progress-towards-undesired-state

  Desire-self: undesirable

  Status: confirmed

  Object: health-status (child, critical-injuries)

  Cause: collision (mom,driver)

  Evaluation: potential-blame{sergeant, mom}

Figure 4: Example Appraisal Frame

network. The sergeant shouldn't experience great distress until he learns the boy is critically injured. Medical helicopters are hard to come by so he shouldn't experience much hope until he learns that one is available. This hope should increase as the various steps of his plan fall into place.

We use the Émile model of emotional appraisal to exploit the causal interpretation and the effects of plan reasoning to derive an agent's emotional state (Gratch, 2000). Inspired by Ortony, Clore and Collins (1988) and Lazarus (1991), Émile is based on the theory that emotions arise from the relationship between an agent's goals and external events. In general, the relationship between an event and a goal can be indirect and Émile relies on the causal interpretation for this assessment. Speaking loosely, events are judged by the extent to which they facilitate or inhibit the probability of goal or sub-goal achievement. Emotional appraisal of an event leads to one or more data structures, called appraisal frames, which characterize the agent's various emotional reactions to an event. Returning to the example from the previous section, Figure 4 illustrates the sergeant's most intense appraisal frame, his distress over the accident's impact on the child's health.

To support our current work, we have made a small number of alterations and improvements to Émile and the structure of the causal interpretation that underlies it. We now model conditional plans: plans may contain sensing actions with indeterminate effects, and conditional plans can be constructed to cover alternative possible outcomes of sensing actions. The model in Gratch (2000) was also unduly focused on goal achievement. It only tracked the utility of goals (which were assumed to have positive utility) and ignored possibility that an action could have an effect that was undesirable (even if it didn't inhibit any active goals). The current model allows the effects of actions to have positive or negative utility, which allows us to assess the impact of any non-goal related side-effects of actions. It also allows an agent to pursue a goal that it (locally) views as having negative utility, which frequently arises in social settings. For example, the sergeant might be given an order that he is loathe to carry out (sending two squads forward, thereby fracturing his outfit), work towards its achievement, but then be pleased/relieved to learn some event has rendered the goal unachievable or irrelevant (e.g., some other platoon can support the inspection). Finally, we extended Émile to appraise events not only from the perspective of the agent's own

emotions, but also appraise events from the perspective of other agents. Specifically, the causal interpretation can contain information about the utility other agents are presumed to attribute to states. In that way the sergeant can surmise that the lieutenant will be emotionally upset that the platoon is not helping Eagle1-6.

5. FOCUS OF ATTENTION

Émile provides a powerful mechanism for appraising the agent-environment relationship but, paradoxically, we have found it too powerful for our needs. Émile appraises and aggregates multiple features of the causal interpretation in parallel, independent of when those features were actually derived. This contrasts with the apparent serial nature of cognition and emotional dynamics, and tends to wash out subtle distinctions between emotions arising from different parts of the causal interpretation. Although this can be useful for computing an overall "mood" for the agent, it is not well suited to guiding the rapid dynamics people exhibit in their facial expressions and speech. In particular, it is not sufficient to support the complex dynamics in Marsella's IPD system (2000), where an animated agent switched from angrily blaming her child to feeling guilty towards her in a single utterance. Characters tied directly to Émile's aggregate emotions can appear wooden or frozen abnormally long in a certain emotion. What is needed is a mechanism that focuses on a small number of appraisals at a time.

Our solution, we feel, is quite elegant. Rather than inventing some new focus of attention mechanism, we found that we could tie the dynamics directly to the basic operations the agent must perform to interact with the student and the environment. To perform in this environment an agent must understand and generate speech, generate and repair plans and direct its sensors to perceive activities in the environment. All of these operations reference or modify the causal interpretation. For example, perception updates the certain beliefs. Each time one of these operations accesses an element of the causal interpretation it activates any emotional appraisals associated with the element. If, after arriving on the accident scene, the student asks the sergeant agent what happened, it activates the emotions of distress associated with the injuries to the driver and the boy. More specifically, the dialogue module, in interpreting the question, makes reference to every past event that occurred in the current location (the accident and the events leading up to it). All the emotions associated with these events are activated. These activated emotions are now available to influence the response to the question.

What focusing means in practice is that, whenever a reasoning mechanism makes reference to an object in the causal interpretation, any emotions associated with the object are made available as "concerns" for the coping process. Currently, we only keep track of the single most intense emotion associated with each known agent in the world. Note that this side-steps the issue of integrating multiple conflicting appraisals of the same event, something that Émile supports. These "in focus" emotions are replaced by any subsequent more intense emotions, or are "discharged" by coping.
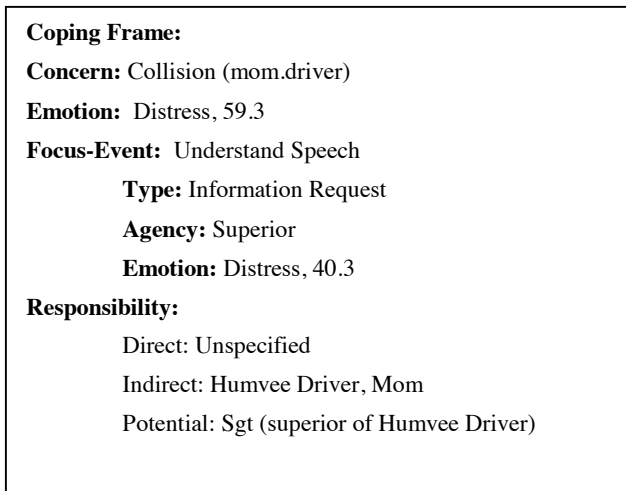
6. COPING

Emotions don't serve just to modulate facial expressions and lexical choice. They are also powerful motivators. Many theories of emotion focus on how people use coping mechanisms to mitigate negative emotions and reinforce positive ones. People typically cope with emotions by acting externally on the world (problem-focused coping), or acting internally to change their beliefs or attention (emotion-focused coping). In the Bosnia scenario, the sergeant is under extreme duress due to the negative emotionality arising from the collision. If and when the student lieutenant asks about the accident, this will bring these strong emotions into focus and creates the opportunity to perform a coping behavior to "discharge" the affect.

This view of coping tightly couples the process that leads to emotion, the appraisal, with the coping process that deals with them. In essence, coping is the inverse of appraisal. To discharge a strong emotion about some situation, one obvious strategy is to change one or more of the factors that contributed to the emotion. Coping operates on the same representations as the appraisals, the agent's beliefs, goals and plans, but in reverse, seeking to make a change, directly or indirectly, that would have the desired impact on appraisal. Coping could impact the agent's beliefs about the situation, such as the importance of a threatened goal, the likelihood of the threat, responsibility for the threat, etc. Further, the agent might form intentions to change external factors, for example, by performing some action that removes the threat. Indeed, our coping strategies, can involve a combination of such approaches. This mirrors how coping processes are understood to operate in human behavior whereby people may employ a mix of problem-focused coping and emotion-focused coping to deal with stress.

Table 1. Some typical coping strategies

Problem-Focused Strategies	Engage in problem solving
	Form intention to act
	Seek instrumental support
Emotion-Focused Strategies	Wishful thinking
	Resignation, Acceptance
	Denial
	Shift responsibility
	Assume responsibility

People employ a rich set of coping strategies and different individuals tend to adopt stable and characteristic "coping styles" that are correlated with personality type. Our preliminary model captures a small subset of these strategies and uses personality-inspired preference rules to model consistent differences in style across different agents. Table 1 illustrates some typical coping strategies. For instance, people may take preemptive action or seek instrumental support from others to circumvent a stressful factor. Alternatively, they may behaviorally disengage from attempts to achieve a goal that is being thwarted or threatened.

```
Coping Frame:

Concern: Collision (mom.driver)

Emotion:  Distress, 59.3

Focus-Event:  Understand Speech

          Type: Information Request

          Agency: Superior

          Emotion: Distress, 40.3

Responsibility:

          Direct: Unspecified

          Indirect: Humvee Driver, Mom

          Potential: Sgt (superior of Humvee Driver)
```

Figure 4: Part of Coping Elicitation frame

6.1 Coping Process: Elicitation & Selection

We model coping as a three-step process. First some focusing event occurs, such as being asked a question or perceiving a change in the simulated environment. This brings into focus some concern, a part of the plan, or causal history, relevant to the event. Assuming the agent has a strong appraisal with respect to a concern, a coping elicitation frame is constructed that ties together the concern, the appraisal, as well as the key situational/social context, such as what and who caused the agent to focus on this concern (e.g., some person asked a question). The elicitation frame also ties together the social relations between the various players in this concern, their emotions as well as their responsibilities with respect to the concern. Figure 4 shows the elicitation frame for the collision that results from the lieutenant's question to the sergeant, "What happened here?" The frame summarizes that the speaker is the agent's superior, as well what the agent believes the speaker feels about the event drawn into focus by the question (the collision). Also, the responsibility for the concern is annotated. The agent has not assigned responsibility for the event to anyone. However the drivers of the cars have indirect responsibility. Further, the agent has inferred that he has potential responsibility since in this case he is the superior of one of the drivers.

The next phase of the coping process involves the matching of potential coping strategies to the elicitation frames. As noted earlier, three strategies match the collision scenario; a form of problem solving (make-amends) and two emotion-focused strategies (assume-responsibility and shift-responsibility). We then use a simplistic model of personality traits to establish preferences over applicable coping strategies. Once a coping strategy is selected, the next phase applies the various internal and external behaviors that realize the strategy. Figure 5 illustrates the entire process, starting with the causal interpretation at the top, leading to the appraisal of the accident from the sergeant and lieutenants perspectives, motivating coping strategy selection, which leads to a change in the causal interpretation.

6.2 Coping Process: Behavior

The final step of the coping process determines the impact on agent's internal state and external behavior. Currently, our coping strategies can impact the agent by a combination of changing
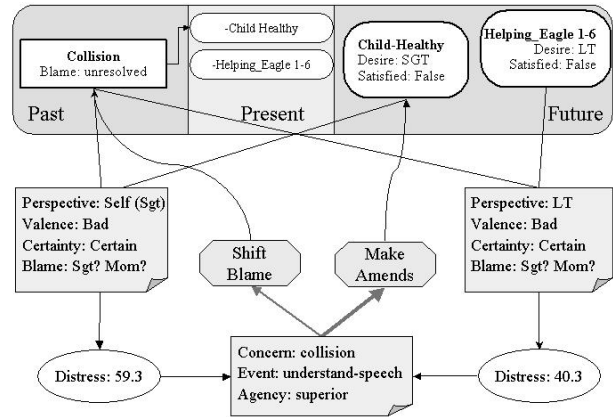
Figure 5: An illustration of appraisal and coping

beliefs, forming intentions to act, forming an intention to communicate and finally expressing the agent's emotions over the concern. Different strategies differentially exploit and emphasize these behaviors. For making-amends, the intention to act in order to address the concern is fundamental. Shift-responsibility relies more on modifying beliefs as well as the intention to communicate that belief.

The key common mechanisms shared across these alternative coping strategies is the formation of intentions and changes to beliefs. These mechanisms currently operate in a constrained fashion and thus the agent is not free to change any belief or form any intention. Here, we again rely on the causal interpretation to guide the process. The formation of new intentions is based on what is feasible given the current interpretation. For example, if the sergeant is following a make-amends strategy in response to the collision's injuring the boy, he will search for a possible action that addresses the boy's health, such as taking the boy to the hospital. He will then form an individual intention to perform that action. This intention is integrated within the rest of the agent's reasoning. So if the user (the human playing the lieutenant) asks the sergeant what they should do, the sergeant will propose taking the boy to the hospital.

Similarly, the agent is not free to change any belief. For example, the sergeant cannot simply start believing the boy's healthy in the face of contradictory information. Currently, we only allow beliefs to be changed for which there is no current contradictory belief formed via perception or dialog with teammates. Along similar lines, we should note that we distinguish between beliefs changed by perception and beliefs changed by coping, which are marked as derived and one might argue should be less permanent.

Note these constraints on intention formation and belief changes are conservative, which limits the range of coping strategies that can be represented. For instance, we could relax intention formation, to allow the formation of intentions for which there is not even any viable plan. If, for example, the intention addressed a concern that was stressing the agent, this "good intention" coping strategy by itself would in fact lead to appraisals that reduced distress regardless of whether the agent eventually acted on the intention. On the other hand, if the intention did not address the concern, the intention formation would nevertheless shift the agent's focus away from the concern via the Focus of Attention mechanism discussed earlier. This strategy mirrors a common human strategy of coping by distraction. In our agent, the effec-

tiveness of this strategy would hinge on whether the intention had a strong emotion associated with it or lead to subsequent intentions or actions.

Similarly, belief changes could and probably should be relaxed. For instance, extreme belief changes could be tied to emotional intensity so that an agent under intense duress might change or hold onto a belief in the face of contradictory information if that coping by denial helped them deal with the stress (e.g. Thagard 2002). Given the persuasive, communicative use of emotional expression, coping mechanisms can also employ feigned belief changes and intentions. For example, the sergeant's most intense concern may be that the lieutenant anger at him for the accident, and not his own concern for the accident. As such, he may cope by feigning the belief that the mother caused the accident. Although such subtlety, or perhaps duplicity, could readily be modeled in the current coping framework by adding a second shift-blame coping rule that simply did not changing beliefs, we have not yet added such a rule.

A third component of coping strategies is the associated expressive behaviors. The range of expressive behaviors that can be triggered from coping span the agent's gaze, facial expression, gesture and posture. The coping strategy also impacts the agent's sentence planning and surface realization. These various forms of behavioral expression are an integral part of the specification of the coping strategy.

The rationale for this integration of coping with behavioral expression stems from two concerns. First, the details of these expressive behaviors, as well as how they relate to the agents underlying emotions, beliefs and intentions, differ markedly across coping strategies. For instance, an agent that is making amends might freely express their true appraisal-based feelings of guilt and concern, for example through facial expressions, gestures, posture and dialog. However, an agent who is shifting responsibility might suppress expressions of guilt and rather express anger at the person they are shifting responsibility to, prior to any feeling of anger that might arise due to changes in beliefs about who is responsible. Thus the expression serves a more deliberate communicative role designed to persuade others. For such reasons, we allow decisions about emotional expression to be part of the coping mechanism even though the changed beliefs and intentions will also engender subsequent emotion.

Second, the wide range of expressive behaviors that an agent can use must be managed in a coherent fashion. The coping strategies allow us to address this concern in a focused way. The agents we design incorporate a wide range of outward behaviors in order to interact believably with the environment as well as other agents and humans. They interact within scenarios that would be very stressful in real life. The agent bodies have fully articulated limbs, facial expressions, and sensory apparatus. They can move in the environment, manipulate objects and direct their gaze in appropriate ways. They are capable of rich, multi-modal communication that incorporates both verbal behaviors as well as non-verbal behaviors. In addition, they have facial expressions, body postures and the ability to perform various kinds of gestures. A key challenge for the agent design is to manage this flexibility in the agent's physical presence in a way that conveys *consistent* emotional state and individual differences. The coping strategy provides a framework, a focus, for achieving this consistency across modalities.

6.3 Dynamics and Expression

A key assumption of our work is that emotions, and the responses they engender, are part of an unfolding dynamic process. At one level, appraisals are forming emotional assessments from the various goals and events represented within the agent. Simultaneously, the agent is responding to external events of various kinds, including dialog from other agents and the learner. Within the agent, beliefs are updated, plans and dialog generated, etc. These processes reference and change aspects of the causal interpretation which, via the focus mechanism, drive the coping process. Coping may then lead to further changes in beliefs and intentions. And as the interpretation changes, so do the underlying appraisals.

We believe this cycle of appraisal, coping and re-appraisal creates an emotional dynamic essential for creating engaging, believable characters that convey a rich mental life (Marsella, Gratch & Rickel, 2001). In this, we are heavily influenced by advances in understanding the dynamics underlying human emotion. For example, Scherer (1984) and Smith & Lazarus (1990) argue for a model of human emotion processes that are in constant flux, with the feedback of coping and re-appraisal processes being integral to that flux. Such theories provide us with principled guidance in developing the emotional dynamics of our agents, as well as how those dynamics are revealed behaviorally.

7. STATUS

To date, our model has been implemented within the context of an Army peacekeeping scenario. The same emotional mechanisms are used across the various intelligent agents in the scenario, the sergeant, the mother and the medic, differing only in the agent's specific plans, beliefs and personality. The system has been viewed by hundreds of people, coming from a range of expertise, including artificial intelligence, psychology, entertainment as well as domain experts from the military. Early versions of the system have been favorably viewed (see Swartout et al, 2001). Lewis Johnson has begun a formal evaluation of how people interpret the agents' verbal and non-verbal behavior (personal communication). One key question that needs to be addressed is the extent to which people can consistently recognize the coping strategy employed by the agent and whether these strategies alter the student's interpretation of events.

8. CONCLUSION

Emotions have a pervasive influence on human behavior. Modeling this influence in virtual humans is a difficult challenge. The interplay between emotions and behavior is not static or unidirectional. A person's response to emotion may subsequently impact their emotional state via reappraisals of their emotional reactions or other strategies for coping with emotional stress. A key component of this dynamic interplay is the impact of emotions on beliefs.

In the work reported here, we have set out a preliminary model of the impact of emotions on beliefs, using an agent's coping mechanisms to tie together changes in the agent's planning representations to the emotional appraisal mechanisms that reason over those representations. This has allowed us to model a wider range of coping strategies that span both emotion and problem focused coping in a general fashion. We have also discussed several enhancements to the appraisal model and how coping mechanisms integrate within the agent's overall behavior.

However, work remains in further developing these coping mechanisms and the underlying representations on which they operate. For example, the modeling of responsibility takes into account degrees of responsibility. It does not take into account the degree to which the responsible party intended to cause harm or to be helpful. This determination is critical to both appraisal and coping. Further, the modeling of personality's impact on coping behavior needs to be extended. Nevertheless, our overall approach looks very promising. We are now seeing unexpected, emergent coping behavior from our agents that at times can be disturbingly lifelike, even though it is not always by our design.

9. ACKNOWLEDGMENTS

10. REFERENCES

[1] Cassell, J, Vilhjalmsson, H. and Bickmore, T. BEAT: the Behavior Expression Animation Toolkit. In *Proceedings of ACM SIGGRAPH* 2001.

[2] Fleischman, M. and Hovy, E. Emotional variation in speech-based natural language generation. In *Proceedings of the International Natural Language Generation Conference*. Arden House, NY. 2002.

[3] Frijda, N. 1987. Emotion, cognitive structure, and action tendency. *Cognition and Emotion, 1*, 115-143.

[4] Gratch, J. Émile: marshalling passions in thraining and education. Proceedings of the Fourth International Conference on Intelligent Agents, Barcelona, SPAIN. 2000.

[5] Gratch, J. & Marsella, S. Tears and Fears: Modeling Emotions and Emotional Behaviors in Synthetic Agents. *Proceedings of the Fifth International Conference on Autonomous Agents*, Montreal, CANADA, 2001, Pp. 278-285.

[6] Hovy, E. H., *Generating Natural Language under Pragmatic Constraints*. Lawrence Erlbaum Associates, Hillsdale, NJ: 1988.

[7] Lazarus, R.S. Emotion and Adaptation. Oxford Press. 1991.

[8] Marsella, S. Johnson, W.L. & LaBore, C. 2000. Interactive Peda-gogical Drama. In *Proceedings of the Fourth International Conference on Autonomous Agents*, 2000, Pp 301-308.

[9] Marsella, S., Gratch, J. & Rickel, J. The Effect of Affect: Modeling the Impact of Emotional State on the Behavior of Interactive Virtual Humans. In *Agents 2001 Workshop on Representing, Annotating, and Evaluating Non-Verbal and Verbal Communicative Acts to Achieve Contextual Embodied Agents*, 2001.

[10] Neal Reilly, W.S., 1996. *Believable Social and Emotional Agents*. Ph.D Thesis CMU-CS-96-138. Carnegie Mellon Univ.

[11] Ortony A., Clore, G. L., & Collins, A. 1988. *The Cognitive Structure of Emotions*. Cambridge University Press.

[12] Parkes, K.R. Locus of control, cognitive appraisal and coping in stressful episodes. *Journal of Personality and Social Psychology*, 46, 1984, Pp. 655-668.

[13] Pearl, J. 2002. Reasoning with cause and effect. *AI Magazine, v23 (1)* pp. 95-112.

[14] Poggi, I. & Pelachaud, C. Emotional meaning and expression in performative faces. In International Workshop on Affect in Interactions: Towards a New Generation of Interfaces, Siena, Italy, 1999.

[15] Rickel, J. & Johnson, L. 1999. Animated agents for procedural training in virtual reality: perception, cognition, and motor control. *Applied Artificial Intelligence*, v13:343-382.

[16] Scherer, K. R. 1984. On the nature and function of emotion: A component process approach. In K.R. Scherer & P. Ekman (Eds.), *Approaches to emotion*, pp 293-317.

[17] Smith, C.A. & Lazarus, R.S. Emotion and Adaptation. In Pervin (ed), *Handbook of Personality: theory & research*, Guilford Press, NY, 1990, 609-637.

[18] Swartout, B., Hill, R., Gratch, J., Johnson, W.L., Kyriakakis, C., LaBore, C., Lindheim, R., Marsella, S., Miraglia, D., Moore, B., Morie, J. Rickel, J., Thiebaux, M., Tuch, L., Whitney, R. & Douglas, J. Toward the Holodeck: Integrating graphis, sound, character and story. In *Proceedings of the Fifth International Conference on Autonomous Agents*, Montreal, CANADA, 2001, Pp. 409-416.

[19] Thagard, P. Why wasn't O. J. convicted: emotional coherence in legal inference. *Cognition and Emotion* (forthcoming).

[20] Traum, D. & Rickel, J. Embodied Agents for Multi-party Dialogue in Immersive Virtual Worlds. In Proceedings of the *International Conference on Autonomous Agents and Multi-agent Systems*. 2002.

[21] Wells, A., and Matthews, G. *Attention and emotion: a clinical perspective*. Lawrence Erlbaum, NJ, 1994.