From: AAAI Technical Report SS-93-05. Compilation copyright © 1993, AAAI (www.aaai.org). All rights reserved.

Variables Influencing the Intensity of Simulated Affective States*

Clark Elliott

The Institute for the Learning Sciences
Northwestern University
1890 Maple Avenue
Evanston, Illinois, 60201

Institute for Applied Artificial Intelligence
DePaul University
243 South Wabash Ave
Chicago, Illinois 60604

email: elliott@ils.nwu.edu

Greg Siegle

The Institute for the Learning Sciences
email: siegle@ils.nwu.edu

Abstract

An important, yet minimally explored, aspect of emotion simulation is the way in which changes in emotion eliciting situations can give rise to different intensities in the resulting emotion instances. Using the work of Ortony, et al. [Ortony *et al.*, 1988] as a guide, we propose a set of *emotion intensity variables* to be used in modeling the causes of varying emotion intensity, and discuss their implementation within the coarse-grained simulation environment of the Affective Reasoner [Elliott, 1992], a program that reasons about emotion. These variables, our motivation for selecting them, and portions of two functions which use them in computing simulated emotion intensities, are presented in this paper.

Motivation

Affect plays an important role in the way humans respond to situations. The affective processes may, from the perspective of a reasoning system, be viewed as a heuristic mechanism for efficiently assessing one's circumstances and acting upon them [Oatley, 1987]. An important, but often ignored component of the simulation of such processes is the intensities of the affective states (e.g. moods, emotions) that arise.

In this article we discuss an approach to reasoning about emotion intensity within the context of the Affective Reasoner [Elliott, 1992] (hereafter, AR) a computer simulation that reasons about emotions in a multi-agent system. The AR is designed around the constraining hypothesis that there are twenty-four distinct categories of emotion, each based on a different set of eliciting conditions (c.f., [Ortony *et al.*, 1988]). In this context we have postulated the existence of functions that map interpretations of simulated situations into scalar intensity ratings for emotion instances within those categories. Working abductively from descriptions of emotion-generating situations, we used *emotion intensity variables* to explain the ways in which changes in various aspects of the simulation might cause changes in emotion intensity. These variables, which represent both situations external to the agent, and dispositions and moods internal to the agent, together with some preliminary functions that illustrate their use in computing emotion intensities, are described in this paper.

Background

In our current research, we simulate simple worlds populated with with agents capable of responding "emotionally" as a function of their concerns. Agents are given unique pseudo-personalities modeled as both a set of *appraisal frames* representing their individual goals, principles, preferences, and moods, and as a set of *channels* for the expression of emotions. Combinations of appraisal frames are used to create agents' interpretations of situations that unfold in the simulation. These interpretations, in turn, can be characterized by the simulator in terms of the eliciting conditions for emotions. As a result, in some cases agents "have emotions," which then may be expressed in ways that are observable by other agents, and as new simula-

*Preparation of this article was supported in part by Andersen Consulting through Northwestern University's Institute for the Learning Sciences.

tion events which might perturb future situations. Additionally, agents use a case-based heuristic classification system to reason about the emotions other agents are presumed to be having, and to form representations of those other agents' personalities that will help them to predict and explain future emotion episodes involving the observed agent [Elliott and Ortony, 1992].

Ortony, et al. [Ortony *et al.*, 1988] discuss twenty-two emotion types based on valenced reactions to situations being construed as goal-relevant events, acts of accountable agents, or attractive or unattractive objects (including agents interpreted as objects). This theory has been extended to include the two additional emotion types of *love* and *hate* [Elliott, 1992]. A summary of these emotion types using groupings based on the associated eliciting conditions appears in figure 1.

Previous implementations of the AR allowed the mapping of situations into the twenty-four emotion types by reasoning about simple eliciting conditions, but they did not provide for the determination of emotion intensity. Using the work of Ortony, et al. [Ortony *et al.*, 1988] as a guide, we analyzed a set of descriptions of such situations and created a set of *emotion intensity variables* to explain the causes of varying emotion intensity, within a coarse-grained simulation paradigm. We reduced the resulting set of variables to a computable formalism, and represented sample situations in the AR. We then isolated three areas of the simulation where variables in either the short-term *state* of an agent, the long-term *disposition* of an agent, or the *emotion-eliciting situation* itself helped to determine the intensity of the agent's subsequent affective state. These three areas, and the intensity-relevant variables embodied in them, are discussed in the next three sections.

Three Categories of Emotion Intensity Variables

In the following analysis we will limit our consideration of emotion intensity to variables that pertain to what Frijda et al. [Frijda *et al.*, 1992] refer to as the *overall felt intensity* of an emotion. They describe overall felt intensity as comprising "whatever would go into the generation of a response to a global question such as this: 'How intense was your emotional reaction to situation S?"'

We have segregated variables thought to affect the intensity of emotion into three groups. The first group, *simulation-event variables*, comprises variables whose values change independently of situation interpretation mechanisms. The second group, *stable disposition variables*, consists of variables that are involved in an agent's interpretation of situations, tend to be constant, and help to determine an agent's personality and role in the simulation. The last group, *mood-relevant variables*, contains those variables that contribute to an agent's mood state.

The values in the *simulation-event variables* group change independently of an agent's interpretation of them. In other words, although these variables are still subject to appraisal by the agent, the *value changes* themselves are external to the appraising agents. Another way to consider these variables is that a single change in the value of one of them may simultaneously, and differentially, affect several appraising agents. Such variables capture "objective" features of the world such as the loudness of a noise or the brightness of a light. Thus, for example, the objective degree of loudness of a piece of music might relate to its *appealingness* for one agent, but to its *repulsiveness* for another. Nonetheless, the volume of the music, a variable *within* the emotion-eliciting situation, in both cases contributes to the calculation of the intensity for the resulting emotions.

By contrast, the *stable disposition variables* group contains variables that help to determine an agent's bias toward interpreting an emotion-eliciting situation one way or another. For example, it would be these which would allow us to specify that, for a given volume level, one agent experiences intense dislike for the music, whereas another experiences only mild dislike. That these variables are considered *stable* does not mean that they can never change but rather that such changes are rather slow, and tend to be unidirectional. For example, someone who finds rap music repulsive, may, over time, come to dislike it less, or even, ultimately, to find it appealing. This group also includes some *relationship variables* that help to define the strength of *friendship* or *animosity* between agents, and their *emotional distance*.

The last group, the *mood-relevant variables*, contains variables which (1) alter an agent's interpretation of situations, (2) are volatile, (3) are bipolar in nature, (4) are not dispositional, so that they naturally return to (agent-specific) default values over time, and (5) may be the result of prior affective experience. For example, if an agent is represented as "feeling unwell or depressed," then hearing rap music might be simulated as being particularly uplifting, "taking the agent's mind off her pain," or as particularly intolerable, "driving her to distraction."

Our current simulation work is based on the proposals of Ortony, et al. [Ortony *et al.*, 1988], and Frijda et al. [Frijda *et al.*, 1992], and most of our intensity variables derive in one form or anther from their work.[1] The main question that arises in trying to implement these and other variables in a com-

[1] In some cases we wished to incorporate affective phenomena not specifically addressed in these works (e.g., "valence bias") and have therefore created computable formalisms for these phenomena based on relevant psychological literature. We do not claim that our formalisms for these phenomena reflect actual cognitive processes at a low level.

Figure 1: Emotion types (Table based on [O'Rorke and Ortony, 1992] and [Elliott, 1992])

Group	Specification	Name and Emotion Type
Well-Being	appraisal of a situation as an *event*	**joy**: pleased about an *event* **distress**: displeased about an *event*
Fortunes of Others	presumed value of a a situation as an *event* affecting another	**happy-for**: pleased about an *event* desirable for another **gloating**: pleased about an *event* undesirable for another **resentment**: displeased about an *event* desirable for another **sorry-for**: displeased about an *event* undesirable for another
Prospect based	appraisal of a situation as a prospective *event*	**hope**: pleased about a prospective desirable *event* **fear**: displeased about a prospective undesirable *event*
Confirmation	appraisal of a situation as confirming or disconfirming an expectation	**satisfaction**: pleased about a confirmed desirable *event* **relief**: pleased about a disconfirmed undesirable *event* **fears-confirmed**: displeased about a confirmed undesirable *event* **disappointment**: displeased about a disconfirmed desirable *event*
Attribution	appraisal of a situation as an accountable *act* of some agent	**pride**: approving of one's own *act* **admiration**: approving of another's *act* **shame**: disapproving of one's own *act* **reproach**: disapproving of another's *act*
Attraction	appraisal of a situation as containing an attractive or unattractive *object*	**liking**: finding an *object* appealing **disliking**: finding an *object* unappealing
Well-being / Attribution	compound emotions	**gratitude**: admiration + joy **anger**: reproach + distress **gratification**: pride + joy **remorse**: shame + distress
Attraction / Attribution	compound emotion extensions	**love**: admiration + liking **hate**: reproach + disliking

puter program that generates simulated emotions is, what features are present (or might be present) in the simulation events, and how can they be represented in ways that lead to human-like emotions on the parts of our automated agents?

The three central variables described by Ortony, et al. [Ortony *et al.*, 1988] (i.e., the *desirability* of an event, the *blameworthiness* or *praiseworthiness* of an act, or the *attractiveness* of an object, or agent construed as an object) determine the *valence* of the resulting emotion. In addition, these variables, each given here as a ***simulation event variable / stable disposition variable*** pair representing the actual feature values within the simulation, and their importance to the agent, are also the primary determinants of emotional intensity. Secondary to these are variables, such as that representing physiological arousal, which modify the base intensity levels generated by the central variables. Some variables only admit values that can reduce or maintain the intensity of an emotion. For example, events are ordinarily assumed to be perceived as *real*. However, if explicitly represented otherwise, the intensity of an emotion resulting from the (at least partially perceived as unreal) situation will be lessened. Lastly, some variables are represented as pairs containing a bias factor and a strength, and will act differently on the intensities of differently valenced emotions.

Ranges and defaults

The ranges we propose for each variable are arbitrary, and are intended for use in functions that simply treat the variables as multiplicative factors. However, an attempt has been made to partially order the effects the variables have on intensity calculations by differentially limiting the values over which each variable may range. No claim is made for psychological fidelity on this account. This is not as critical as it might seem, since each of the values within the ranges must, in all cases, be given as an assessment, either direct or indirect, of some feature value. Once a scale has been established it is used as a reference for interpreting events and internal states of the agent. If a range proves, in practice, to be too high, it tends to be used more conservatively than if it is too low, and vice versa.

Moreover, the ranges we have specified for variables are not important to the design of our intensity architecture. Instead we have merely provided arbitrary first designations for these ranges, based loosely on introspective analysis. These ranges, default values for variables, and the intensity functions that use them, have been isolated in the implementation, and are treated essentially as data, so that they may be easily altered to reflect the current cognitive theory.

In the basic intensity functions given at the end of this article we use values for the intensity variables as factors: those values below 1 reduce the strength of the emotion, those above 1 increase it. In the current model, primary determinants of emotion intensity range from zero to 10, with a default of 3. Factors which are treated as weaker modifiers of intensity range from (approximately) zero to 3. Modifiers which, within our model, can only *reduce* intensity, range from zero to 1. Variables whose effects on intensity calculations are determined by the valence of an emotion (such as a variable which heightens the intensity of negatively valenced emotions but lessens the intensity of positively valenced emotions) are given both a *bias value*, and a strength ranging from 1 to 3.

It is not desirable to have to specify a value for each of the intensity variables in each emotion eliciting situation that arises in the simulation. In addition, some variables (such as that representing the concept of surprise), which we might desire to represent in a particular situation, are sometimes difficult, if not impossible, to calculate.[2] Intensity functions that are designed to use the missing values, must therefore either use an alternate calculation method, or must use defaults, as is done in the current implementation. Because in this implementation intensities are calculated by multiplying the intensities of component factors together, we use a default value of 1 for most variables since this then gives them the convenient property of affecting the intensity calculation only when otherwise specified.

The intensity variables

Simulation-event variables

- **Goal realisation/blockage.** The degree to which an agent interprets a situation as having contributed to the achievement or blocking of one of his or her goals. This is derived from a simulation feature value, or set of values, and represents a desirability continuum.[3] Example: "How big of a tip the waiter got for his services." Range: -10 to 10, where -10 represents complete blockage of a goal, and 10 complete achievement. No default, this must be specified by calculation from some feature value(s).
- **Blameworthiness-praiseworthiness.** The degree to which an observing agent interprets an observed agent as having upheld or violated one of the observing agent's principles, in some situation. This is derived from a set of simulation values, which might include values for the amount of effort expected in a given situation, the accountability of an agent as determined by role, and so forth. Range: -10 to 10, where -10 represents maximal blameworthiness, etc. No default.
- **Appealingness.** The degree to which an agent interprets a situation as containing an object (or agent construed as an object) that is seen as appealing. Example: how good the concert was.[4] Range: zero to 10. No default.
- **Repulsiveness.** The degree to which an agent interprets an object as being repulsive. Example: how bad-smelling the person was. Range: -10 to zero. No default
- **Certainty.** The degree to which the appraising agent is certain that the event or act has, or will, actually come about. For example, if the simulation event is represented as hearsay, the certainty is likely to be less. Range: zero to 1, with 1 being complete certainty. Default 1.
- **Sense-of-reality.** The degree to which a situation is perceived as real by an agent. For example, a client's suspension of disbelief is likely to depend on the degree to which a salesperson is able to make a hypothetical event seem *real*. Range: zero to 1. Default 1.
- **Temporal proximity.** The distance, in time, of a situation that is being appraised. More recent events, and expected events that are closer to coming about, generate more intense emotions. This is particularly useful in the AR as a way of creating recurrent instances in of an emotion (i.e., rumination). Range: zero to 1. Default 1.
- **Surprisingness.** The degree to which the agent is likely to be surprised by the simulation event. This represents deviation from norms, and in particular, roles. This is, in general, difficult to represent. Note that this is *not* a measure of expectation (see the footnote above). Range: 1 to 3, with 3 being intensely surprising. Default 1.

[2] For example, for a person to be surprised when a jet airplane crashes into their building they must know that this is something that does not normally occur. By contrast, a person is not surprised when the phone rings in their office, because this is something that does normally occur. This is a true default reasoning problem because though one event is surprising and the other is not, the person need not have active expectations regarding planes or telephones (see also [Ortony and Partridge, 1987]). We have included these difficult-to-calculate variables because they are important theoretically, and because in some simulation paradigms we have a pragmatic way of making such values available to the affective reasoning machinery.

[3] The continuum for this variable, and the next three, must be measured along the correct dimension. This is not always a straightforward problem. For example, in the Biblical story of the two harlots that are both claiming to be the mother of a newborn (Kings 1:3), Solomon threatens to divide the child in half and give a piece to each of the women who claim him. The real mother desires first that she get full possession of the child, second that she have possession of no part of him, and least of all that she have half of him.

[4] A more straightforward example would be, "how fresh the flowers were." This example was chosen because it illustrates the concept that the *judgment* of appealingness, and so forth, is internal to the agent, while the *degree* of that appealingness may be set by an external, albeit difficult to measure, value.

- **Effort.** The degree to which an agent has invested resources in achieving the goal, or in upholding the standard. This is not a factor for preferences. Range: zero to 3. Default 1.
- **Deservingness.** The degree to which the agent believes that the subject agent is deserving of the good or bad fortune (for the subject agent) that appears to stem from the simulation event. The intensity of a fortunes-of-others emotion is dependent only upon whether the experiencing agent is happy or unhappy about the other's fortune and the perception of their deservingness, and not the perceived goodness or badness of the event for the other agent. If the agent is perceived as being deserving, then the emotion moves in a positive direction. If the agent is perceived as being undeserving, then the emotion moves in a negative direction. For example, gloating over an *adversary*'s misfortune (a positive emotion), becomes more positive (i.e., more intense) as he or she is perceived as being more deserving, and becomes less positive (i.e., less intense) as he or she is perceived as being less deserving. On the other hand, pity will decrease if a *friend* is seen as being more deserving of the bad fortune, and will increase if he or she is seen as being less deserving of the bad fortune. Represented as a bias either toward deservingness or underservingness, and a strength, 1 to 3. Default 1.

Stable disposition variables

The specification *stable* refers to the position these variables have in determining the relatively stable personalities and roles of the automated agents in the simulation. This is not to say that these values cannot change over time, but rather that such changes will be considered moderately permanent, with no tendency to return to the original state. Although we include the *friendship* and *animosity* variables in this category, this is an arbitrary decision since the relationships they represent, as we define them, can be affected by mood.

Appraisal bias variables These variables, based on the central intensity variables of Ortony, et al. [Ortony *et al.*, 1988], appear as part of the agents' appraisal mechanisms and represent the degrees to which various events, actions of agents, and objects are important to them. For example, two agents may like music, but for one it is *very* important and has the potential for generating ecstasy, whereas for the other it may lead, at most, to a moment of mild pleasure.

The importance of situations for which goal-based and standard-based appraisals are made may be assessed differently depending on whether a goal is achieved (standard upheld) or blocked (standard violated). This is because some concerns can only lead to negative or neutral outcomes, whereas others can only lead to neutral or positive outcomes. For example, one is not ordinarily happy about "not getting cancer," but might be very distressed if such a *preservation* health goal were blocked. Similarly, *serendipitously* winning in the lottery is an important event, whereas losing is not. The same is true of standards, where, for example, it is not normally considered praiseworthy to refrain from committing crimes, but it *is* considered blameworthy to commit them. Preferences and non-preferences are represented separately, and so in each instance a single variable suffices. For discussion see also [Schank and Abelson, 1977], and [Elliott, 1992]).

The default values for the variables in this section are meaningful only in that they highlight our assumption that the degree to which a goal, standard, or preference is important to an agent makes a relatively significant contribution to calculations of emotion intensity. However, to avoid such absurdities as having the breaking of a shoelace and the theft of one's car lead to similarly intense emotions, these values must be specified.

- **Importance to agent of achieving goal.** How desirable the achievement of a particular goal is for the agent. Range zero to 10. Default is 3.
- **Importance to agent of not having goal blocked.** How undesirable the blockage of a particular goal is for the agent. Range zero to 10. Default is 3.
- **Importance to agent of having standard upheld.** How praiseworthy the act of upholding a particular standard is for the agent. Range zero to 10. Default is 3.
- **Importance to agent of not having standard violated.** How blameworthy the act of violating a particular standard is for the agent. Range zero to 10. Default is 3.
- **Influence of preference on agent.** The degree to which an agent likes or dislikes an object, or agent construed as a object, irrespective of the agent's goals and principles. Range: zero to 10. Default is 3.

Stable relationship variables

- **Friendship-animosity.** The degree to which an agent is in a unidirectional *friendship* or *animosity* relationship with an agent (including itself). Represented as a bias, either toward friendship or animosity, and a strength, zero to 3. Default strength is 1. There is no default bias, but it must be specified for fortunes-of-others emotions.
- **Emotional interrelatedness of agents.** This is unidirectional. It is akin to intimacy, but only in the sense that an agent can perceive this intimacy in relationship to an enemy as well as to a friend. This variable attempts to capture the concept that events which affect, and actions which are performed by, agents who are perceived as being more emotionally intertwined with the experi-

encing agent, are likely to generate stronger emotional responses in that agent. Range: zero to 3, where 3 represents a high degree of perceived emotional interrelatedness. Default 1.[5]

Mood variables

Mood representation in the AR is in the early stages of development. The variables given here represent a first attempt at accounting for some of the most salient cases in which prior mood contributes to emotion intensity. Discussion follows at the end of this section.

Non-relationship mood variables

- **Arousal.** The degree to which the agent is physiologically aroused. In future work it may be useful to discriminate multiple dimensions for this variable (such as how tense or relaxed the agent is, how active or drowsy, and so forth). Our model is not so sophisticated. In general we consider the physiological processes to be beyond the scope of this work. Range: 0.1 to 3 where 0.1 represents a comatose state and 3 represents maximum arousal.
- **Physical well-being.** The degree to which the agent is feeling (physically) ill or well. Range: 0.1 to 3, where 0.1 represents intense discomfort and 3 intense vitality.
- **Valence bias.** The degree to which an agent will have a bias towards negative or positive emotions. This is useful in the simulation for the coarse-grained representation of the, essentially, good and bad moods that may result from the emotional experiences of the automated agents. This variable is distinct from the two variables following in that, rather than change the intensities of specific classes of emotions, it sets a base starting point for the calculation of all valenced reactions. Represented as a bias, positive or negative, and a strength 1 to 3, where 3 represents a maximal tendency to appraise the world in a negative light, or in a positive light. Default strength is 1. There is no default bias.

  There are a number of psychological phenomena which may be thought of as contributing to an agent's valence bias, notably including mood disorders. Though such phenomena are not well understood and rigorous models of their determination are scarce [MacLeod and Mathews, 1989], there is a body of literature pertaining to mood disorders and more common, less serious, analogs of these disorders, experienced by non-disordered people in everyday situations (e.g., mild, possibly transient depression). We have therefore included variables to represent mild degrees of two acknowledged mood-biasing phenomena, namely depression and anxiety.[6]
- **Depression-ecstasy.** From generalized sadness and feeling of hopelessness to generalized euphoria and feeling of well-being. This variable is modeled after the clinical notions of depressive and manic states which seem to have a number of cognitive correlates. Depression is generally characterised as a predisposition towards sadness and negative life-construals while mania is characterised by overestimation of abilities, ignoring negative repercussions of events, etc. [Association, 1987]. Interestingly, non-depressive subjects seem to be biased towards positive construals of themselves and others (e.g., having a bias towards viewing themselves as praiseworthy) while depressives seem to lack this bias, yielding even-handedly mixed positive and negative self-construals [Greenberg and Alloy, 1989]. Represented as a bias, positive or negative, and a strength 1 to 3, where 3 represents a the highest degree of depression, or the highest degree of ecstasy. Default strength is 1.5. (See discussion below.) There is no default bias.
- **Anxiety-invincibility.** From generalized fear and feeling of loss of control, to generalized hope and feeling of powerfulness. This variable is modeled after the clinical notion of general anxiety which seems to have cognitive correlates which are in some cases similar to, and in other ways distinct from, depression. For example, anxious subjects seem to ascribe more negative anxiety-relevant traits (e.g., nervousness) to themselves and to unknown others than positive anxiety-relevant traits (e.g., assurance) [Greenberg and Alloy, 1989]. Represented as a bias, positive or negative, and a strength 1 to 3, where 3 represents a the highest degree of anxiety, or the highest degree of perceived invincibility. Default strength is 1. There is no default bias.
- **Importance of all Goals, Standards, and Preferences.**

  The degree to which all of the goals, standards, and preferences of an agent are important to that agent. As with valence-bias, this is a coarse-grained representation of a phenomenon that affects the intensity of all emotions, namely that when an agent's emotional resources are diminished or enhanced, that agent's goals, standards,

[5] It is not clear what the relationship between this variable and the friendship-animosity variable should be. We have created this as a distinct variable to handle such instances as animosities that arise in the course of family life, and so forth. To wit: if a family member cheats you of your inheritance it may be a much more intense experience than if "the government" does, regardless of whether the relationship to that family member is one of friendship or animosity.

[6] We treat depression and anxiety as mood-biasing phenomena rather than as moods themselves, in keeping with much cognitive psychology research (e.g., [Beck, 1967]).

and preferences may all become more important, or less important, respectively. For example, when an agent has had a series of bad things happen, the agent may have an increased *desire* for something good to happen. When the agent's goals are blocked in this state, the agent is more intensely unhappy about it; when the agent encounters some appealing object, it is seen as all the more appealing. Range: 0.3 to 3, where 3 represents the maximally *increased* importance of goals, standards and preferences.

Relationship mood variables

- **Liability-creditableness.** The degree to which an agent will bias judgments of the liability or creditableness of another agent. For example, when an agent is in an *angry* or *hating* mood towards another agent they will tend to hold them more liable for blameworthy actions and less creditable for praiseworthy ones; when they are in a *grateful* or *loving* mood toward that other agent then the opposite is true. Represented as a bias, *liable* or *creditable*, and a strength, 1 to 3. Default is 1.

Discussion Rudimentary moods for agents in the AR can result from an emotion, or series of emotions. To make use of this mechanism, it has been necessary to specify several granularities of mood. General mood biasing factors (such as *valence-bias*) tend to be automatically generated by the simulation as a result of emotions, whereas more specific mood factors (such as *depression-anxiety*) tend to be hand-coded as a result of specific situations. When values for more specific mood-biasing variables are present, they are preferred over those for more general mood-biasing variables.

In our model, *anxiety* and *depression* interact to form distinct valence biases. For example a *non-depressed*, *non-anxious*, agent will react in a way which is biased towards positive anxiety-relevant traits (e.g., "assurance") while an *anxious* agent will have a bias towards negative anxiety-relevant traits (e.g., "tension"), and a *depressed* agent will have the bias removed entirely. As such, *depression* may have the effect, in the model, of dulling an agent's usual anxious nervousness (e.g., a depression-induced representation of "it just doesn't matter").

To illustrate, let us suppose that in a simulation an agent is "fired from her job." This agent's concerns lead to both sadness over the loss of comfortable circumstances and companions (stemming from the blockage of some career and social goals), and fear over the prospects of future financial hardship (stemming from the *prospect* of a financial goal being blocked). If we raise the prior value of her "anxiety" (by giving *anxiety-invincibility* a negative bias and increasing its value), the predominant emotion becomes fear, ("How am I going to pay the rent?"). Raising her prior level of "depression" (by giving *depression-ecstasy* a negative bias and increasing its value) mediates her fear but increases her sadness, ("Everything just seems to be going wrong!").

Similarly, if a *non-depressed*, *non-anxious*, simulated agent is "involved in a successful business deal," the agent might ordinarily be predisposed to attain an "undue" intensity of pride from the event (as non-depressives often have an inflated sense of self-involvement in positive causal interactions [Vasques, 1987]), while *depression* would temper this effect, and *anxiety* would provoke the agent to fear the deal's consequences.

Intensity Function Illustrations

The intensity calculation functions that make use of these variables within the simulation are still under consideration. Our initial goal is simply to establish some reasonably monotonic relationship between the feature values which we are able to represent within the simulator and the intensity of the simulated emotions resulting from the specified eliciting conditions. There are many such simple relationships that can be captured. However, to model the subtle effects of some of the variables, including their inter-dependence with one another, and their non-linear effects, more complex functions are needed.[7]

As with the determination of ranges and defaults, the particular intensity-calculation functions used are not important to the reasoning architecture. These can be changed to accommodate different kinds of psychologically motivated experiments. The only fixed link from the intensity functions to the rest of the design is that they must return the scalar values which are used to set the various thresholds for different intensities in the generation of emotion instances, and that they return values for the different categories of emotion that are consistent with one another. Also important to note is that values for variables must be used consistently in both the elicitation of emotions, and their manifestation, since the appraisal mechanism passes variable bindings to the emotion manifestation mechanism.

Lastly, one final implication of our approach is that an agent might appraise a situation many different ways such that all the appraisals lead to the same intensity level in the resulting emotion, even with respect to the same principles and goals. An agent might, for example, be particularly angry because she is already not feeling well, or alternatively because she is already annoyed with the one responsible for her anger. In our architecture, such differences are important because, as just noted, the bind-

[7] Even the notion of emotional intensity itself is underspecified since a full treatment would have to account for such parameters as the amplitudes of multiple intensity peaks, their durations, and so forth. (c.f., [Frijda *et al.*, 1992]).

ings for the values of the different emotion-intensity variables are passed to the action generation component of the AR, and might influence the reasoner's choice of actions for the experiencing agent.

Normalization variables are included in each function so that the relative strengths of emotions will be correct. Normalization is important for two reasons. First, different emotion types have different numbers of variables, which would affect their overall intensity range. Second, the intensities for compound emotions (e.g., *anger*) are based on the intensities bindings of *two* construals rather than one, and this can only be done in a consistent manner if the constituent construals are first normalized.

Figure 2 shows the basic outline of a function that calculates an emotion intensity for simple goal-relevant appraisals. Note that the defaults for *mood-variables* (e.g., *depression-ecstasy*) are those for the function only, and may be superseded by a different default for a particular agent.

Figure 3 shows a portion of a similar function used to calculate the intensities of fortunes-of-others-emotions. In this case the variables represent the *supposed* values for the other agent. We add variables for *friendship-animosity* and *emotional-interrelatedness*. As these values increase so do the strength of the resulting fortunes-of-other emotions. In addition we add a calculation for the effect of perceived deservingness or undeservingness. For example, if one is resentful over the good fortune of an adversary, but the adversary is seen as deserving, then the intensity of the resentment is lowered according to the degree of deservingness.

Results

Our task was to define a set of variables that would allow us to calculate coarse emotion intensities for all emotion episodes that the Affective Reasoner was capable of representing. We constrained our task by specifying three discrete levels of intensity for each of the twenty-four emotion classes, giving us seventy-two different emotion/intensity duples that were to be represented. To test our representation we analysed a number of different emotion episodes, as well as complete scenes, for their affective content, and used the emotion-intensity variables to represent the perceived causes of the different intensities, within the context of our emotion theory. Episodes were analysed with respect to the way in which the intensity variables allowed us to map simulation-event features into values for the eliciting conditions of the different emotion types (see figure 1). Complete scenes were additionally analysed for the ways in which affective states could be mapped to mood changes for the automated agents, using the intensity variables, so that the intensities of subsequent emotions would be correct for those agents. The present set of variables was sufficient to represent all of these episodes and scenes, at the Affective Reasoner's level of granularity.

The intensity model has not been implemented for all emotion types. Unsolved representational problems exist for some of the variables, such as *surprisingness* and *deservingness*. The functions that use the emotion intensity variables are in the early stages of development. Moreover, the cognitive correlates of the variables we have discussed interact, and they have cognitive import beyond that of emotion intensity. For example, a disturbance in what we have termed *sense of reality* might not only lower the intensity of person's emotions but could also affect which emotions are generated, given the same eliciting conditions (e.g., when audience members laugh at gruesome scenes in a horror movie). Because we represent intensity variables and emotion intensities as scalars our model is not able to address these phenomena. A more cognitively faithful representation would allow multidimensional intensity variables to affect each other in context dependent ways, and would use such variables in not only the determination of emotion intensity but also in the basic construal process responsible for the selection of the emotion types themselves. We consider our work to be only a preliminary effort.

```
(defun goal-relevant-intensity
    (achieved-blocked
     &key importance-achieving importance-not-blocking certainty
     sense-of-reality temporal-proximity surprisingness
     effort arousal physical-well-being importance-gsps
     valence-bias valence-bias-strength
     anxiety-invincibility anxiety-invincibility-strength
     depression-ecstasy depression-ecstasy-strength)
  (let ((emo-valence (if (> achieved-blocked 0) 'positive 'negative)))
    (*
     *goal-relevant-normalization*
     achieved-blocked
     (if (eql emo-valence 'positive)
         (or importance-achieving *ia-default*)
         (or importance-not-blocking *inb-default*))
     (or certainty *certainty-default*)
     (or sense-of-reality *sense-of-reality-default*)
     (or temporal-proximity *temporal-proximity-default*)
     (or surprisingness *surprisingness-default*)
     (or effort *effort-default*)
     (or arousal *arousal-default*)
     (or physical-well-being *physical-well-being-default*)
     (or importance-gsps *importance-gsps-default*)
     (if valence-bias
         (if (eql valence-bias emo-valence)
             valence-bias-strength
             (protect-divide 1 valence-bias-strength))
         *valence-bias-default*)
     (if anxiety-invincibility
         (if (eql anxiety-invincibility emo-valence)
             anxiety-invincibility-strength
             (protect-divide 1 anxiety-invincibility-strength))
         *anxiety-invincibility-default*)
     (if depression-ecstasy
         (if (eql depression-ecstasy emo-valence)
             depression-ecstasy-strength
             (protect-divide 1 depression-ecstasy-strength))
         *depression-ecstasy-default*))))
```

Figure 2.

```
(defun fortunes-of-other-intensity
     ...
     *fortunes-of-others-normalization*
     friendship-animosity-strength
     (or emotional-interrelatedness *emotional-interrelatedness-default*)
     ...
     (if deservingness-undeservingness
         (if (eql deservingness-undeservingness 'deserving)
             (if (eql emo-valence 'positive)
                 deservingness-undeservingness-strength
                 (protect-divide 1 deservingness-undeservingness-strength))
             (if (eql emo-valence 'negative)  ;;; note: d/u = 'undeserving
                 deservingness-undeservingness-strength
                 (protect-divide 1 deservingness-undeservingness-strength))
             )
         *deservingness-undeservingness-default*)))
     ...
```

Figure 3.

References

[Association, 1987] American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders (DSM IIIR)*. American Psychiatric Association, Washington, D.C., 1987.

[Beck, 1967] Aaron Beck. *Depression: Clinical, Experimental, and Theoretical Aspects*. Hoeber, New York, 1967.

[Elliott and Ortony, 1992] Clark Elliott and Andrew Ortony. Point of view: Reasoning about the concerns of others. In *Proceedings of the Fourteenth Annual Conference of the Cognitive Science Society*, Bloomington, IN, August 1992. Cognitive Science Society.

[Elliott, 1992] Clark Elliott. *The Affective Reasoner: A Process Model of Emotions in a Multiagent System*. PhD thesis, Northwestern University, May 1992. The Institute for the Learning Sciences, Technical Report No. 32.

[Frijda *et al.*, 1992] Nico Frijda, Andrew Ortony, Jaap Swagerman, and Gerald Clore. The complexity of intensity. *Review of Personality and Social Psychology*, pages 60–89, 1992.

[Greenberg and Alloy, 1989] Michael Greenberg and Lauren Alloy. Depression versus anxiety: Processing of self- and other-referent information. *Cognition and Emotion*, 3(3):207–223, 1989.

[MacLeod and Mathews, 1989] MacLeod and Mathews. Cognitive-experimental approaches to the emotional disorders. *Handbook of Behavior Therapy*, pages 116–150, 1989.

[Oatley, 1987] Kieth Oatley. Editorial: Cognitive science and the understanding of emotions. *Cognition and Emotion*, 3(1):209–216, September 1987.

[O'Rorke and Ortony, 1992] Paul O'Rorke and Andrew Ortony. Explaining emotions. Submitted for publication, 1992.

[Ortony and Partridge, 1987] Andrew Ortony and Derek Partridge. Suprisingness and expectation failure: what's the difference? In *Proceedings of the Tenth International Joint Conference on Artificial Intelligence*, Milan, Italy, August 1987. IJCAI.

[Ortony *et al.*, 1988] Andrew Ortony, Gerald L. Clore, and Allan Collins. *The Cognitive Structure of Emotions*. Cambridge University Press, 1988.

[Schank and Abelson, 1977] Roger C. Schank and Robert Abelson. *Scripts, Plans, Goals and Understanding*. Lawrence Erlbaum Associates, Hillsdale, New Jersey, 1977.

[Vasques, 1987] Carmelo Vasques. Judgment of contingency: Cognitive biases in depressed and nondepressed subjects. *Journal of Personality and Social Psychology*, (52):419–431, 1987.