Cathexis: A Computational Model for the Generation of Emotions and their Influence in the Behavior of Autonomous Agents

by

Juan David Velásquez

Submitted to the Department of Electrical Engineering
and Computer Science in partial fulfillment of the
requirements for the degree of

Master of Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September, 1996

Signature of Author ....
Department of Electrical Engineering and Computer Science
September, 1996

Certified by ....
Pattie Maes
Associate Professor Media Arts and Sciences
Thesis Supervisor

Certified by ....
Marvin Minsky
Toshiba Professor Media Arts and Sciences
Thesis Supervisor

Accepted by ....
F.R. Morgenthaler

Cathexis: A Computational Model for the Generation of Emotions and their Influence in the Behavior of Autonomous Agents

by

Juan David Velásquez

Submitted to the Department of Electrical Engineering and Computer Science on August 9, 1996, in partial fulfillment of the requirements for the degree of Master of Science in Computer Science and Engineering

Abstract

This thesis presents *Cathexis*, a distributed, computational model of emotions that addresses a number of issues and limitations of models proposed to date including the need to consider and model the dynamic nature of different types of emotions, such as basic emotions, emotion blends, and mixed emotions, the need to consider both cognitive and noncognitive elicitors of emotion, the need for differentiating emotions from other affective phenomena, such as moods, and most important, the need for a flexible way of modeling the influence and effects of emotions on the motivations and behavior of the agents. The computational model is described, and its current implementation as a framework to create emotional agents and model emotional phenomena is presented.

Thesis Supervisor: Pattie Maes
Title: Associate Professor Media Arts of Sciences

Thesis Supervisor: Marvin Minsky
Title: Toshiba Professor Media Arts and Sciences

Acknowledgments

There are many people to thank for contributing to this work and to three great years I have had at MIT.

I want to specially thank Pattie Maes and Marvin Minsky for being great advisors, excellent motivators, and role models. I was, without any doubt, very fortunate to be able to work with both of them. I thank Pattie for her confidence, and for extending me her hand in the moment I needed it the most. I am very grateful for her advice, her patience (specially with my abstract ideas), her continuing support, and most specially, for helping me keep a balance between my cognitive and emotional sides! I thank Marvin for his support, his numerous ideas, and for being an incredible source of inspiration.

Unlimited thanks to Roz Picard for her motivation, and for invaluable discussions, comments, and suggestions in relation to this work. Numerous thanks as well to Bruce Blumberg for great discussions, and for his enthusiasm with the ideas on this thesis.

Thanks to Push Singh for taking the time to listen and comment on this work. Thanks to the great people of the Autonomous Agents group, and to all people in the Affective Computing group for sharing their ideas and for always being willing to discuss on a variety of topics. Numerous thanks to Karen Modrak, Betty Lou McClanahan, and Marylin Pierce for their help and patience during this time.

Many thanks to Randy Davis for being a great academic advisor, and for his invaluable support, patience, and confidence during the first hard years at MIT. I am also grateful to Lynn Stein for her advice and help while finding the topic of this thesis.

Many thanks to the people at CECI for their tremendous support and friendship. Thanks to Steve Lerman for the opportunity he gave me three years ago when I first came to MIT as a Visiting Engineer, and for his support and good academic and life-related advice during this time. Special thanks to Jud Harward for being a great project leader and most important, an incredible friend and role model.

Infinite and deepest thanks to Claudia Urrea, for many years of love and support. She made every minute more fun, easier to go through when tough, more memorable, and certainly more worthwhile.

Incomparable and unlimited thanks to my family for their unconditional and continuous support and love. Their perseverance and energy throughout life have influenced who I am and have inspired everything I have accomplished.

To the many other people who contributed to this work and who have not been listed here, my warmest thanks, and I ask that you take it on my cognitive side and not on my emotional one. For I have just momentarily forgotten, but you always remain present in my heart!

To my family, and to the memory of Carlos.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

> *"The question is not whether intelligent machines can have any emotions, but whether machines can be intelligent without any emotions. I suspect that once we give machines the ability to alter their own abilities we'll have to provide them with all sorts of complex checks and balances."*
> -- Marvin Minsky

Emotions are certainly a fundamental part of our lives. They influence the way we behave, and how we communicate with others, yet when we look at current research in Artificial Intelligence (AI) we see that not too much work is being done on understanding and modeling emotions.

We usually think of emotions as being nonrational. The word "emotional" by itself most often carries along a negative connotation as we know that people might act like "crazy" when emotions surge out of control. Knowing this, why would we want to create emotional agents? Why would we want to give emotions to computers? Do we really want to create computers that can be depressed or that get annoyed or angry at us and decide not to do their tasks? In the case of synthetic animated agents in interactive storytelling systems, the answers are rather obvious since, as Bates has pointed out, emotion is one of the primary means to achieve the illusion of life and create "believable agents", thus helping the users to suspend disbelief [Bates 1994]. In regards to other kinds of autonomous agents, such as software agents, and physical robots, these answers are less obvious and perhaps many people would argue against providing agents with emotions or emotion-like mechanisms. Nevertheless, as we will describe shortly, by designing computational models of emotions and creating agents that are based on these models, we can find new ways of improving the interaction between humans and computers, and at the same time, we might obtain a better understanding of what emotions are and what is their role in human cognition and behavior.

Recently, neurological evidence has indicated, contrary to what most people believe, that emotions are essential and that they serve a substantial function in human intelligence. In [Damasio 1994], for instance, arguments are made based on neurological evidence that not having "enough" emotion can bring disaster into our behavior and actions. In his work, Damasio describes how some of his patients, which have damage in the prefrontal-amygdala circuit, are terribly impaired in regards to their decision-making, and although there is no deterioration in their cognitive abilities and their intelligence seems to be intact, they make terrible decisions in their jobs and personal lives, and they can endlessly obsess about simple things such as making a simple appointment.

Damasio argues that this happens because they have lost access to their emotional learning -- memories of the likes and dislikes we acquire during our lives, and so, no emotional reactions are triggered which would otherwise play a biasing role in the decision-making process by pointing in the right direction.

In related work, Salovey and Mayer [Salovey and Mayer 1990], and later Goleman, have suggested that emotions are a crucial and fundamental part of human intelligence, and that "*emotional intelligence*: abilities such as being able to motivate oneself and persist in the face of frustrations; to control impulse and delay gratification; to regulate one's moods and keep distress from swamping the ability to think; to empathize and to hope", is far more important in predicting success in life than the traditional accounts of intelligence based on IQ tests [Goleman 1995].

This kind of evidence suggests that emotions are indispensable even for rational decisions and therefore we should study them in order to better understand human cognition. Furthermore, this implies that if we want to create agents that are to be effective at decision-making and perhaps truly intelligent, we will need to provide them with emotions or at least emotion-like mechanisms that work as a complement to the rest of the system.

Of no less importance is the fact that emotions play a crucial role in communication and interpersonal relationships. Recent research suggests that even though computers have always been regarded as not being able to have or express emotions, emotional factors do play a role in human-computer interaction [Nass , Steuer, and Tauber 1994]. In this work, a number of classical studies about human interaction were conducted where humans were substituted for computers making the study one where tests for human-human interaction where applied to human-computer interaction. The results showed that computer users (even experienced ones) do apply social rules to their interaction with computers, even though they report that such attributions are inappropriate. Therefore, by providing computers with mechanisms to express and perceive emotions, we can greatly improve human-computer interaction.

All these studies have motivated and generated increased interest in the field of emotions and have also caused the emergence of new areas of research such as *Affective Computing*, which as defined by Picard is "computing that relates to, arises from, or deliberately influences emotions" [Picard 1995].

Recent developments in the area of synthetic agents, such as [Maes et al. 1995; Blumberg 1994; Loyall & Bates 1993; Elliot 1992; Hayes-Roth, Brownston & Sincoff 1995], have also promoted the study of emotions and their influences in behavior. Many of these influences have been studied and theoretical models that account for such influences have been designed [Bower and Cohen 1982]. Nevertheless, up to date, relatively few computational models of emotion have been proposed. Some of these models include emotion reasoners such as BORIS [Dyer 1983], Elliott's Affective Reasoner [Elliott 1992], and the Em architecture [Reilly and Bates 1992]. Other systems include psychological models such as Abelson's model of attitude change [Abelson 1963], PARRY, a model of artificial paranoia [Colby 1975], Swagerman's ACRES system [Swagerman 1987], and Pfeifer's FEELER [Pfeifer 1988]. Chapter 2 reviews some of these models.[1]

This thesis describes Cathexis, a computational model of emotions which addresses a number of issues and limitations of models proposed to date including the need for models of different kinds of affective phenomena, such as emotions and moods, the need to consider different systems for emotion activation, and the need for a flexible way of modeling the influences of emotion on the motivations and behavior of agents. The following sections discuss the nature of the problem in more detail.

1. For a more detailed review on the history of models of emotions the reader is referred to: Pfeifer, Rolf. Artificial Intelligence Models of Emotion. In: Hamilton, V., Bower, G. H., and Frijda, N. eds. *Cognitive Perspectives on Emotion and Motivation*, 287-320. Netherlands: Kluwer.

1.1 Modeling Emotions

Before describing a model of emotions it would be appropriate to mention something about what it is we are modeling. This brings us to the difficult issue of defining emotion. A large number of definitions for the concept can be found in the literature, yet very limited consensus has been reached. There is some agreement, however, on some of the basic characteristics of emotion. Most researchers agree that one of the emotion components corresponds to the elicitors or activators of emotion. Also, there is agreement in that emotions include an expressive or motor component, and that particular expressive movements (e.g. smiling) help define emotion (e.g. indication of Happiness). Some of the aspects involved in this expressive component include central nervous system efferent activity, prototypical facial expressions, body posture, head and eye movements, vocal expression, and muscle action potentials. Finally, most researchers would agree that once an emotion is generated, it registers in consciousness. This is what is sometimes referred to as the emotion experience or feeling. As Izard suggests, "it seems reasonable to stay with the notion that emotion experience is a motivational condition or process in consciousness that manifests itself as action readiness, action tendency, a biasing of perception, or a feeling state" [Izard 1993].

Having some notion of what an emotion is, or at least what are its basic components, we can start discussing some of the issues that require careful consideration while designing a computational model of emotion and which constitute some of the issues for which this thesis makes specific contributions to.

1.1.1 Elicitors of Emotion

One of the basic issues in modeling emotions is that of emotion generation or activation. Given a specific situation or the occurrence of an event, what emotion should be activated, if any? Why do people in a particular situation experience one particular emotion and not another? Why do different people in the same situation, might experience different emotions? Many theorists have dedicated much of their time to find the answers to these questions. Results from this work can be seen in the several theories and taxonomies that have been proposed for analyzing the conditions that lead to emotions [Oatley & Johnson-Laird 1987; Ortony, Clore, & Collins 1998; Roseman, Spindel, and Jose 1990; Weiner 1986]. Most of these theories, however, concentrate mostly (or only) on defining and understanding the cognitive processes that elicit emotion, and even though the role of emotions in evolution and adaptation suggests that there might be more than one system or mechanism for generating them, the idea of alternative activation systems (different from the cognitive ones) has remained as an underresearched issue.

The consideration of both cognitive and noncognitive elicitors of emotion is one of the contributions of the computational model of emotions presented in this thesis.

1.1.2 Dynamic Nature of Emotions

Once an emotion is generated, it does not remain active forever. After some period of time, unless there is some sort of sustaining activity, it disappears. This seems like a trivial issue, but nevertheless, since in most computational models of emotion there is no real notion of time, it is one that is frequently not dealt with. To appropriately model the generation of emotions and their interaction with other parts of the system, computational models should take into account the dynamic nature of emotions.

In this thesis, emotions are modeled as dynamic systems that once they become active, unless there is some continuous eliciting stimuli, their intensity decays according to some specific function until they become inactive again.

1.1.3 Emotions and Other Affective Phenomena

Emotions differ from other kinds of affective phenomena. The term *mood,* for instance, refers to a different but related affective phenomenon which carries (and sometimes is saturated by) emotions. Moods can be distinguished from emotions in terms of their time course. As Ekman suggests [Ekman and Davidson 1994], what we call moods last much longer than emotions. Although there is no basic agreement on how long emotions usually last, most people would agree that moods last longer. Ekman has also maintained that emotions can be very brief, typically just a couple of seconds and at most minutes but sometimes people speak of emotions lasting hours, however, he argues that what actually happens is that people experience a series of repeated but discrete emotion episodes [Ekman 1992].

Most models proposed to date do not consider other kinds of affective phenomena, such as moods. Yet evidence suggests that moods and emotions dynamically interact in important ways which should be considered in a comprehensive model of emotions. The computational model presented in this thesis provides models for different kinds of affective phenomena including basic and mixed emotions, moods and temperaments.

1.1.4 Influence of Emotion

This issue has to do with deciding and modeling how, once an emotion has been activated, it influences further behavior of the system. Evidence suggests that the emotion experience has powerful influences in memory processes, goal generation, action selection, further generation of emotions, reasoning style, and learning, among other things.

Although modeling all of the possible different influences of emotion would be quite an enormous and difficult task, a good model of emotion should at least acknowledge the existence of such influences and provide a flexible way of modeling some of them, while leaving open space so that further developments can be added at a later time. This is in fact the approach taken within this thesis to address this issue. As it will become more clear in Chapter 3 and Chapter 4, the computational model of emotions we propose can be extended at a later time to include systems that model different influences of emotion not considered here.

1.2 Overview of this Thesis

The remaining chapters of this thesis are organized as follows: Chapter 2 reviews some of the relevant work in models and cognitive theories of emotion. In Chapter 3 we present Cathexis, a computational model which provides an alternative approach to modeling emotions and addresses the issues described in this chapter. Chapter 4 describes an implementation based on the ideas presented in Chapter 3. Chapter 5 describes the results obtained from the creation of a test-bed environment that incorporates our model. And finally, Chapter 6 presents our concluding remarks and some directions for future work.

Chapter 2

Related Work

As we discussed in Chapter 1, relatively few computational models of emotion have been proposed up to date. The purpose of this chapter is to review some of these models and the theories that have been used as the basis for them.

2.1 Architectures and Models of Emotion

Most of the models of emotion proposed to date fall in the category of *reasoners*. That is, systems that reason about emotions. Some of these models have been integrated into architectures that model agent behavior. However, among these, generally the model of emotions is not consider as fundamental in order to select the appropriate behaviors for the agents. This section briefly describes some of these different models.

2.1.1 BORIS

BORIS is a story understanding system designed to analyze and comprehend short but complex narratives [Dyer 1983; Dyer 1987]. These narratives often contain descriptions of the emotional states and the emotional reactions of characters to the different situations they encounter. BORIS includes a component designed to understand the significance of these emotional reactions with the idea that doing this will help in the analysis and understanding of the narrative. The system takes as its input some text representing a story and then reasons about the possible emotional states of the characters in that story. For example, in a story that describes a robbery scenario, BORIS may reason about one of the characters being afraid after a thief has stolen his watch.

Since BORIS was not specifically designed to model emotions, but rather to understand narratives, its affect component is very simple and it is limited to *reasoning* about the meaning of emotional reactions and emotion terms when used or described in language, but it does not model active emotional states, nor does it consider many other aspects of the emotional experience, which, as discussed in Chapter 1, are essential to produce emotional behavior.

2.1.2 The Society of Mind

Marvin Minsky's Society of Mind [Minsky 1986], is a collection of many small, incredible ideas that provide partial models which attempt to explain how human minds work. As a general theory of the mind, the Society of Mind certainly address many issues which suggest what emotions are and how they come about.

In particular, Minsky argues that minds are made of many little components, called *agents*, specialized in some domain of thought but which have no mind of themselves. In the same spirit, emotions are seen as varieties of thoughts which are based on specialized agents called *proto-specialists*. In children, these agents start out as separate components that do not know much about each other and therefore do not make use of each other's activities. Later, as our mind develops, these agents grow together learning how to use one another, and thus, forming complicated networks of exploitation which cause the complexity of adult emotions [Minsky 1986].

As it will become more evident later, many of the ideas described in the Society of Mind, including the concepts of proto-specialists, exploitation, and cross-exclusion, had significant influences on the computational model presented in this thesis.

2.1.3 The Cognition and Affect Project

The main goal of the Cognition and Affect project at the University of Birmingham is to understand the types of architectures that are capable of accounting for not only intelligent behavior but also moods, emotions, and other affective phenomena [Sloman 1991].

Sloman's approach follows a design-based model which claims that in order to understand many mental states, including affective states, they need to be distinguished relative to a whole architecture. Therefore, Sloman is mainly concerned with specifying and studying architectures for modeling intelligence and trying to discover and understand the mechanisms that produce emotions and how they fit into the overall architecture.

In some ways, this work is similar to the Oz project at Carnegie Mellon University (See 2.1.4) in the sense that both projects attempt to design broad agent architectures. One of the main differences between these projects, though, is that the Cognition and Affect project claims that certain states and processes for which others usually have provided explicit models, can instead emerge from complex motivational systems.

2.1.4 Em

Em is a set of tools designed to assist artists in the creation of believable emotional agents [Reilly 1996; Reilly and Bates 1992]. Em handles most of the emotional and social aspects of agents created as part of the Oz project at Carnegie Mellon University whose main goal is to create believable agents for use in interactive fiction and virtual realities

[Bates, Loyall, and Reilly 1992].

Em's model of emotions is primarily based on the OCC model (See 2.2.3) but it is also influenced by ideas from Dyer [Dyer 1983]. Like in the OCC model, Em's generation of emotions is based on the comparisons of external events with goals, actions with standards, and objects with attitudes.

Em has been integrated into the Tok architecture [Bates, Loyall, and Reilly 1992] and tested on different text-based simulation systems and one graphic system which includes three agents called *The Woggles*, which are animated cartoon-like blobs with semi-expressive faces that coexist in an animated world [Loyall and Bates 1993].

As an emotion architecture Em includes several of the features described in Chapter 1, including models for different emotion types, consideration of emotion decay, and mapping to behavioral features based on the emotional state of the agent. However, given that the system is designed for artistic purposes, some of these features are contemplated only superficially and may be oversimplified in comparison to real life emotional systems.

Although on the surface Em may seem similar to the work described in this thesis, the goals contemplated on both projects, and the approaches taken to modeling emotions are fundamentally different. Em's design has been specifically adjusted to meet an artistic end. Therefore, including a model of emotions is just one of the means to aid artists in the creation of believable agents. In contrast, the work described in this thesis is primarily concerned with designing a comprehensive, integrated model of emotions that considers the many different aspects of real life emotional systems, and it is secondarily concerned with the different applications of such model.

2.1.5 The Affective Reasoner

The Affective Reasoner is a real-time simulation platform for *reasoning* about emotions that supports multiple agents [Elliott 1992; Elliot 1994]. The model of emotions included in this work contains representations for twenty-four distinct emotion types, based on the work of Ortony, Clore, and Collins (See 2.2.3). It also includes a simple mapping of emotions to several different expressions which provide agents with the ability to display emotional reactions to different situations that arise in the simulations.

Additionally, agents use a case-based classification system to reason about the emotions of other agents and to build representations of their personalities which can be used later to predict and explain future emotion episodes involving those particular agents.

One interesting aspect of the whole model is the integration of different multimedia elements such as music, text-to-speech, and animated faces to aid in the expression of emotion and increase the user's ability to suspend disbelief when interacting with the agents.

2.2 Cognitive Appraisal Theories

According to cognitive appraisal theorists, what determines whether an emotion will be felt and what emotion it will be, depends on the evaluation and appraisal of specific situations, rather than on the situations themselves.

Although several emotion theorists have addressed the role of cognitive appraisals in emotion, only some of these theorists have provided specific relationships between appraisal configurations and specific emotions. This section reviews some of the theories that explain emotional states based on specific configurations of appraisals.

2.2.1 Weiner's Theory of Causal Attributions

Weiner proposed that many emotions arise from explanations for events and outcomes. He claimed that attributions arise from achievement, affiliative, and power concerns. His theory focuses on achievement concerns.

Following the outcome in achievement situations, Weiner claimed that a two-step process occurs. First, a person will experience positive affect if the outcome of the situation is viewed as successful, and negative affect if the outcome is regarded to be a failure. In the second step, the person will interpret the outcome of the event based on several dimensions, such as the locus of causality which determines if the event was caused by the self (internal) or by others (external), and control which determines if the situation was controllable or not.

Table 2-1 illustrates how these dimensions relate to the different emotions discussed by Weiner. For instance, anger results from attributions to an event that was the product of the controllable actions of others (external locus and control), whereas pride, results from attributing a positive outcome to the actions of self (internal locus), when self had control.

Attribution Dimensions	*Positive Outcome*	*Negative Outcome*
Locus: Internal		
Controllable	pride	guilt
Uncontrollable		shame
Locus: External		
Controllable	gratitude	anger
Uncontrollable		pity

Table 2-1 Weiner's Proposed Relationships Between Dimensions and Emotions

2.2.2 Scherer

Scherer proposed that emotional processes involve five functionally defined systems:

1. An information processing system that evaluates the stimulus through perception, use of memory, prediction, and evaluation of available information.
2. A support system that regulates the internal state through control over neuroendocrine, somatic, and autonomic states.
3. An executive system that plans, prepares actions, and chooses among competing motives.
4. An action subsystem that controls motor expression and overt behavior
5. A monitoring system that controls attention devoted to current states and channels the feedback that results from the attention to other subsystems.

The information processing system is based on appraisals, called stimulus evaluation checks (SECs). Scherer proposed five major SECs, four of which have subchecks [Scherer 1988]. These appraisals or checks are described below:

1. *A novelty* appraisal which decides whether the pattern of external or internal stimulation has changed. This appraisal includes three other subappraisals: *suddenness*, *predictability*, and *familiarity*, which determines the extent to which the occurring event is sudden, predictable and familiar, respectively.
2. An appraisal of *intrinsic pleasantness* which determines whether the stimulus is pleasant, bringing about approach tendencies, or unpleasant, creating avoidance tendencies.
3. A *goal-significance* appraisal which assesses whether the event is conducive to or obstructive of the person's goals. This appraisal includes several subappraisals:

- *concern relevance*: determines the relevance of the event in relation to the person's goals or needs
- *outcome probability*: determines what is the probability of a specific event happening
- *expectation*: evaluates the consistency of the outcome of an event with the persons expectations for the present time
- *conduciveness*: determines the extent to which the event obstructs the person's goals
- *urgency*: determines the urgency of a needed response to the occurrence of an event.

4. An appraisal of *coping potential*, evaluates the degree to which the person believes they can handle the situation. Subappraisals for this check include:
   - *agent*: determines the cause of the event
   - *motive*: determines the motive underlying the event
   - *control*: evaluates the extent to which the person perceives that he or she has control over the situation and its consequences
   - *power*: determines the sense of power experienced by the person in dealing with the event
   - *adjustment*: evaluates the potential for adjustment to the outcome of the event
5. The last appraisal proposed by Scherer is that of *compatibility standards*. This appraisal determines whether the event is compatible with external and internal standards such as social norms, cultural conventions, or the expectations of others.

Table 2-2 presents the relationships between the major appraisal configurations and emotions proposed by Scherer.

	Emotions					
Appraisal/ SEC	*Joy*	*Fear*	*Anger*	*Sadness*	*Disgust*	*Shame*
1. Novelty						
Suddenness	high/med	high	high	open	open	low
Familiarity	open	open	low	open	low	open
Predictability	low	low	low	open	low	open
2. Pleasantness	open	low	open	open	very low	open
3. Goal Significance						
Concern Relevance	self/rel	body	order	rel/ord	body	self
Outcome Probability	very high	high	very high	high	very high	very high
Expectation	open	dissonant	dissonant	open	open	open
Conduciveness	conducive	obstructive	obstructive	open	open	open
Urgency	very low	very high	high	low	medium	high
4. Coping Potential						
Cause: Agent	open	other/nat	other	other	open	self
Cause: Motive	cha/int	open	intent	intent	open	int/neg
Control	open	open	high	high	open	open
Power	open	very low	high	low	open	ope
Adjustment	medium	low	high	high	open	medium
5. Compatibility						
External	open	open	low	very low	open	very low
Internal	open	open	low	very low	open	very low

Table 2-2 Scherer's Proposed Relationships Between SECs and Emotions

2.2.3 Ortony, Clore, and Collins

Ortony, Clore, and Collins propose a model (from now on referred as the OCC model) which provides a framework for the relationship between emotions, beliefs, goals, and standards [Ortony, Clore, and Collins 1988]. They claim that people focus their attention on one of the following three aspects: events, agents, and objects. In this model, emotions are seen as valenced reactions to any of these perspectives, with their nature being determined by the way in which the eliciting situation is interpreted by whoever experiences them. By valenced they mean either positive or negative. Distress for example, is a negatively valenced emotion, whereas joy is a positively valenced one.

Instead of describing every possible emotion term, the OCC model describes emotions in terms of families, which they call emotion types. Emotions within each family share similar causes. For instance, the joy type describes all emotions caused by pleasing desirable events. Within this category, several emotion tokens are included, such as joyful, cheerful, happy, jubilant, and pleased, which have similar characteristics and might only differ in small things such as their intensity.

This is perhaps one of the most influential models that has been proposed. In fact, many other models and architectures (including some described above) are based on it. One of the limitations of this model is that although it proposes a comprehensive theory about the structure and causes of emotions, it does not address other very important aspects, such as the physiological, expressive, and behavioral elements which are crucial in understanding what emotions are and what their role is in human cognition.

Figure 2-1 illustrates the entire structure proposed in this model.

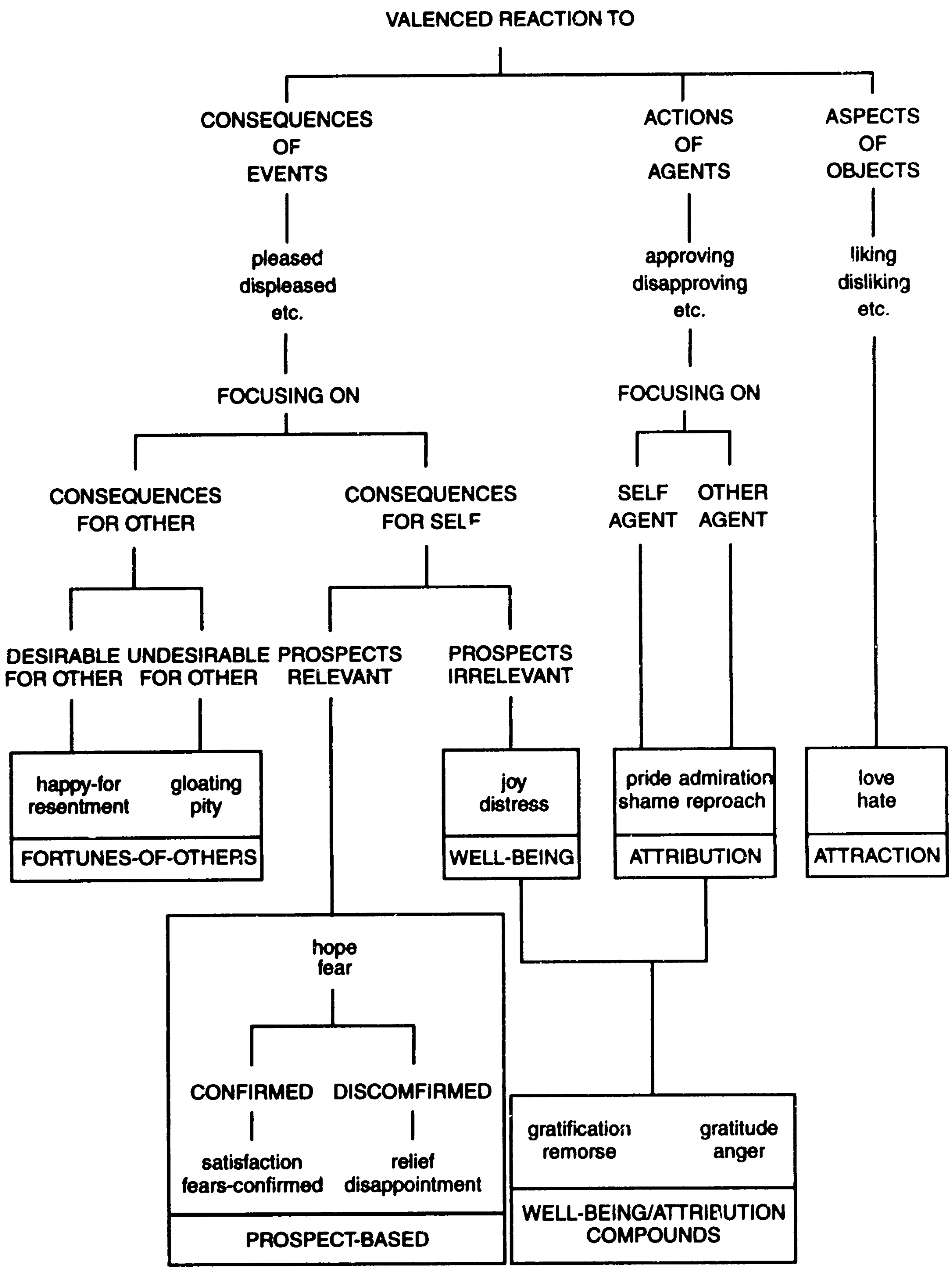

Figure 2-1 Ortony, Clore, and Collins' structure of emotion types.

2.2.4 Roseman

Roseman's theory, as all other appraisal theories, tries to determine the particular appraisals of events that elicit emotions. In this theory, appraisal configurations for 16 discrete emotions are analyzed. Roseman proposes that five different appraisals influence emotions. These appraisals are:

1. *Motivational State*: An appraisal determining whether an individual's motive for a particular situation is *aversive* (a punishment that the individual wants to avoid) or *appetitive* (a reward that the individual seeks to obtain).
2. *Situational state*: An appraisal assessing whether the events are *consistent* or *inconsistent* with an individual's motives.
3. *Probability*: This dimension determines whether the occurrence of an outcome is *certain* or *uncertain*.
4. *Power*: An appraisal determining if the individual considers himself or herself *weak* or *strong* in a given situation.
5. *Agency*: This dimension evaluates whether the outcome is caused by impersonal *circumstances*, some *other* person, or the *self*.

Roseman claims that different combinations of these appraisals will elicit different emotions. Table 2-3 presents the appraisal configurations hypothesized to elicit 16 discrete emotions according to him. For example, he predicts that a person will experience sadness (sorrow), when an event caused by circumstances has occurred (it is certain) and it is evaluated by the person as being inconsistent with a motive to obtain reward.

<table>
<tr><th rowspan="3"></th><th colspan="2">Positive</th><th colspan="2">Negative</th></tr>
<tr><th colspan="2">Motive-Consistent</th><th colspan="2">Motive-Inconsistent</th></tr>
<tr><th>Appetitive</th><th>Aversive</th><th>Appetitive</th><th>Aversive</th></tr>
<tr><td>A. Caused by circumstances</td><td colspan="4" rowspan="2">Surprise</td></tr>
<tr><td>1. Unknown</td></tr>
<tr><td>2. Weak</td><td colspan="2" rowspan="2">Hope</td><td colspan="2" rowspan="2">Fear</td></tr>
<tr><td>a. Uncertain</td></tr>
<tr><td>b. Certain</td><td>Joy</td><td>Relief</td><td>Sorrow</td><td>Disgust</td></tr>
<tr><td>3. Strong</td><td colspan="2" rowspan="2">Hope</td><td colspan="2" rowspan="3">Frustration</td></tr>
<tr><td>a. Uncertain</td></tr>
<tr><td>b. Certain</td><td>Joy</td><td>Relief</td></tr>
<tr><td>B. Caused by others</td><td colspan="2" rowspan="7">Liking</td><td colspan="2" rowspan="5">Disliking</td></tr>
<tr><td>1. Weak</td></tr>
<tr><td>a. Uncertain</td></tr>
<tr><td>b. Certain</td></tr>
<tr><td>2. Strong</td></tr>
<tr><td>a. Uncertain</td><td colspan="2" rowspan="2">Anger</td></tr>
<tr><td>b. Certain</td></tr>
<tr><td>C. Caused by self</td><td colspan="2" rowspan="7">Pride</td><td colspan="2" rowspan="4">Shame, Guilt</td></tr>
<tr><td>1. Weak</td></tr>
<tr><td>a. Uncertain</td></tr>
<tr><td>b. Certain</td></tr>
<tr><td>2. Strong</td><td colspan="2" rowspan="3">Regret</td></tr>
<tr><td>a. Uncertain</td></tr>
<tr><td>b. Certain</td></tr>
</table>

Table 2-3 Roseman's Appraisal Theory

2.3 Discussion

The work described in this section includes, or is related to, in one way or another to models of emotions. These models, however, are limited in the sense that they either do not address important issues like those described in Chapter 2, or they address them superficially. Specifically, most of these models do not differentiate between different kinds of affective phenomena, such as emotions, moods, and temperaments. Also, given that most of them are based on some of the taxonomies and appraisal theories described above, they only consider the cognitive aspects that contribute to the activation of emotional systems. Furthermore, the majority of these models fall in the category of *reasoners* about emotions and thus do not consider the influences of emotions in other systems. The following chapter describes Cathexis, a computational model of emotions which attempts to address some of these issues and limitations.

Chapter 3

The Cathexis Model

We have developed a distributed model for the generation of emotions and their influence in the behavior of autonomous agents. The model is called Cathexis[1] and has been inspired by work in different fields including among others, Psychology, Ethology and Neurobiology. The main contributions of this model are:

• It provides models for different affective phenomena, including basic emotions, emotion blends, mixed emotions, moods, and temperaments.

• An approach to emotion generation that takes into consideration both cognitive and noncognitive elicitors of emotion.

• A general model which considers the dynamic nature of emotions.

• A flexible, incremental way to model the influences and effects of emotions in the agent's behavior which considers both the Expressive and Experiential components of emotion (i.e. It considers influences in the agent's expression -facial expression and body posture- as well as the influences in the agent's motivational and internal state).

Figure 3-1 provides a high level view of the model's architecture. The following sections describe this architecture in more detail.

---

1. From the greek *kathexis* meaning holding or retention. The term was introduced into the literature as a translation of Freud's term "besetzung", which connotes a concentration of emotional energy on an object or idea.

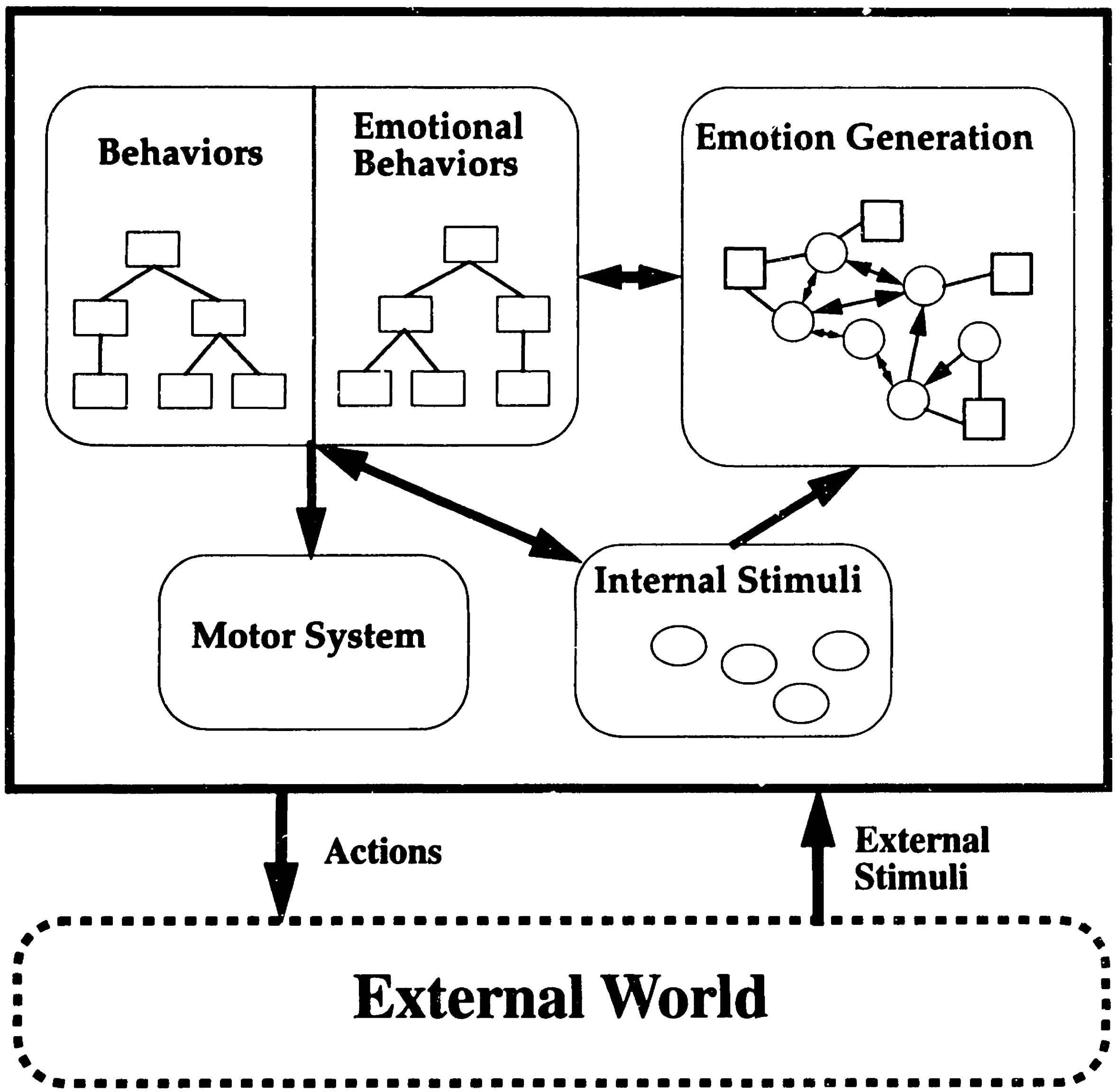

Figure 3-1 Cathexis Architecture.

3.1 Emotion Generation System

As discussed before, one of the most basic issues in modeling emotions is that of emotion generation. This issue has to do with the very basic question of what emotions should a computational model include? Furthermore, given a specific situation, or the occurrence of an event, what are the emotions elicited by that situation or event? What elicits them? With what intensity are they activated? Should other affective phenomena be modeled as well? All of these questions should be considered while designing a model of emotions. This section describes what approaches are taken to address these issues within the Cathexis model.

3.1.1 Basic Emotions

The proposed computational model, drawing on ideas from Ekman, Izard, Plutchik, and several other theorists [Ekman 1992; Izard 1991; Plutchik 1994; Johnson-Laird and Oatley 1992], contains explicit models for the so-called *basic* or *primary* emotions. The expression *basic emotions* has been used by researchers in many different ways. In this work the term basic, as in [Ekman 1992], is used to emphasize how evolution has played a significant role in forming the unique and common characteristics that emotions exhibit, as well as their current function. In other words, it is used to claim that there are a number of separate discrete emotions which differ from one another in important ways, and which have evolved to prepare us to deal with fundamental life tasks, such as, falling in love, fighting, avoiding danger, adjusting to losses, and so on.

There is still much debate over precisely which emotions can be considered basic or even if there are basic emotions at all. Part of the problem with defining what the basic emotions are is due to our use of language. The rich language of emotion does not correspond very well with the whole emotional experience or with the different components of emotion. Most languages provide a great variety of words that can be applied to many situations to describe how we feel. Sometimes a single word can be used to refer to different emotional experiences, and other times we can use different words to refer to the same experience. Take for instance the word pity. We use this word to refer to a situation in which we are saddened by the misfortune of others. In this context, the word pity refers in part to the particular variation of sadness, but it also refers to the relationship we have with the person who is suffering, and it even refers to the problem this person is facing. The feeling of sadness may be one, but the word pity to describe it may be replaced with another word, such as sympathy, rue or compassion depending on the particular situation. Because of the many different possibilities offered by the language of emotion, we fail to recognize that many of the terms we use to describe an emotion refer in fact to variations of a single basic emotion, or to no emotion at all but rather to other kinds of affective phenomena such as moods, traits, or even emotional disorders [Shaver et al. 1987; Johnson-Laird and Oatley 1989].

In pursuing the concept of basic emotions and knowing that some of the labels given to the basic emotions might still cause confusion and disagreement, the proposed model follows Ekman and others in thinking of emotions in terms of groups or families [Goleman 1995; Ekman 1992]. Within Cathexis, each emotion is not a single affective state but rather a family of related affective states. Each member of an emotion family shares certain characteristics, such as similarities in antecedent events, expression, likely behavioral response, physiological activity, etc. These characteristics differ between emotion families, distinguishing one from another. Cathexis includes explicit models for the following basic emotions: *Anger*, *Fear*, *Distress/Sadness*, *Enjoyment/Happiness*, *Disgust*, and *Surprise*.

Providing explicit models of these emotions does not mean to imply that there are no other basic emotions. It simply means that for other emotions there is not enough evidence that suggests that they are basic in the sense described here. Nevertheless, there are other emotion families, such as *Embarrassment/Shame* and *Interest/Excitement* which may be basic emotions as well, but unless some evidence is found that suggests this, they will not be included in the list of basic emotions for which Cathexis provides explicit models.

Recently, Pfeifer indicated some of the problems with explicit models of emotion and discussed how the lack of consensus and the difficulties in defining what an emotion is, lead to suspect that emotions are not clearly defined "things" or "faculties", but rather emergent [Pfeifer 1995]. We do not believe this is the case for the basic emotions, but we do agree that not all affective phenomena, including some which we might commonly label as emotions, can be clearly delineated and studied as such, and that instead they are, as Sloman suggests, well defined, natural consequences of the design of complex motivational systems [Sloman 1991].

The following sections describe how Cathexis deals with other affective phenomena, including emotion blends, mixed emotions, and moods.

3.1.2 Emotion Blends and Mixed Emotions

Certainly, some affective states considered to be emotions are not included in the list of basic emotions defined above. Let us remember though, that the concept of emotion families allows the inclusion of some of these states as variations (in intensity, expression, etc.) of a basic emotion. This of course does not include all of these different affective states. Does this mean that the remaining states are not emotions then? Not necessarily. Even though there are some theorists who claim that all emotions are basic, there is some evidence that suggests the existence of *blends*, a term which describes the co-occurrence of two or more basic emotions at a time. Examples of blends might be emotions such as *Grief* and *Jealousy*. *Grief*, for example, is certainly more specific than *Sadness*. Besides a variation of *Sadness*, people who experience *Grief* are also likely to feel *Anger* and *Fear*, and perhaps even *Surprise* depending on the situation. A similar thing happens with *Jealousy*, in which the person who is jealous feels a variation of *Anger* but *Fear* and *Sadness* may also be felt.

Another example of emotion blends is seen in what is normally referred to as *mixed emotions*. Consider the following scenario: Suppose you placed a bet on a basketball game in which your best friend is a member of one of the two opponent teams. Knowing that the team in which your friend plays is really bad and has lost six games in a row, you decide to place your bet against the team where your best friend plays (and you call yourself a friend?). Now, suppose your best friend's team wins. In a situation like that, you might feel happy while sharing your friend's happiness, but at the same time, you might feel sad or even angry because you lost your money (well deserved!). Or in a more real life situation, when Uta Pippig, winner of the 100th Boston Marathon was interviewed some time after crossing the finish line, she expressed feeling tremendously happy for winning the race, surprised for she believed she would not win, somewhat sad that the race was over, and a bit fearful because during the race she was having acute abdomen pain. As it can be seen in these examples, it is not uncommon to experience mixed, conflicting emotions at one time.

Cathexis allows for the co-occurrence of two or more basic emotions at the same time. It is in this way that the model is able to deal with this kind of emotions. So, even though there are no explicit models to represent emotion blends or mixed emotions, they exist as part of the interaction between two or more active basic emotions. The intensity level and the influences (both expressive and experiential) of each of the active emotions, give emergence to these secondary emotions. Note that these emotions are not just attributions of observers of the system, because in fact, the agent's emotional state at that particular moment is composed of several active basic emotions.

3.1.3 Moods and Other Affective Phenomena

A number of words considered to be emotion terms actually refer to moods rather than emotions (e.g. nervousness, irritation, euphoria). Sometimes, even the words emotion and mood are used interchangeably by some psychologists and laypeople alike, to refer to certain aspects of affect. However, as we mentioned before in Chapter 1, moods and emotions differ in several aspects including their function. It is therefore important to differentiate them within a model of emotions.

Moods and emotions interact in important ways. Emotions can induce particular moods and moods can change the probability that certain emotions will be triggered. Moods appear to lower the activation threshold for those emotions which occur most frequently during a particular mood. For instance, when you are in an bad mood, you become angry more easily than usual. Simple events that might ordinarily not make you angry can now provoke anger easily. As Ekman suggests [Ekman 1994], moods seem to be caused in two different ways. The first cause of moods can be due to changes in one's biochemical state which might be in turn caused by changes such as lack of sleep or lack of food as we have seen in children and also in adults. The second cause of moods can be due to dense emotional episodes in which a specific emotion is generated at a very high intensity, again and again, with little time between each occurrence. For instance, repeated occurrences of joyful emotions usually modify an individual's emotional state and induce him or her to be in a good or happy mood.

Following a psychobiological perspective, Cathexis differentiates between moods and emotions in terms of levels of arousal. Emotions may consist of high arousal of specific types of brain systems. These systems, however, may be tonically activated at low levels by a variety of internal stimuli, as well as relatively weak perceptual inputs. Thus, moods may be explained as low tonic levels of arousal within emotional systems, while emotions would be explained as high levels of arousal in these same systems [Panksepp 1994]. The concurrent activation of several of these systems will lead to the possibility of an enormous display of mood states, some of which could be described with common labels, such as *cheerful*, *irritable*, *melancholic*, *anxious*, and so on.

As it will become more evident later, by allowing changes in the parameters that control the function of emotional systems, Cathexis differentiates temperament from moods and emotions, which gives developers and other users of the model, the ability to specify the agent's affective characteristics at different levels.

3.1.4 Cognitive and Noncognitive Elicitors of Emotion

In 2.2, we described some of the proposed models or taxonomies for analyzing the conditions leading to emotions. Appraisal theorists [Frijda 1986; Oatley and Johnson-Laird 1987; Ortony, Clore, & Collins 1998; Roseman, Spindel, and Jose 1990; Weiner 1986] claim that what determines whether an emotion will be felt and what emotion that will be, depends on the evaluation and interpretation of events, rather than on the events per se. Although only a few of these authors take an extreme position in which cognition is the exclusive determinant of the emotions, most theories on the generation of emotion are concerned only with cognitive elicitors.

In contrast, and with a strong influence from Izard's multi-system for emotion activation [Izard 1993], and also drawing on ideas from Tomkins [Tomkins 1962], the Emotion Generation System in Cathexis takes into account the existence of both cognitive and non-cognitive elicitors of emotions. These elicitors correspond to one of the four activation systems defined by Izard and which are described as follows:

- *Neural*: This system includes neuroactive agents that can lead to emotion (independent of cognition], such as, neurotransmitters, drugs and electrical or chemical stimulation of the brain. Evidence has shown that physiological mechanisms, such as, hormones, sleep, diet, etc., as well as radical changes of temperature, can trigger changes in neurochemical processes, and thus elicit emotions. It may also be the case (and actually in Cathexis it is), that neural systems run continuously and periodically to activate certain emotions or alter emotion thresholds in the absence of effective stimuli for the other emotion activation systems. This independent activation, according to Izard, may explain some of the individual differences in positive and negative emotionality, sometimes regarded as emotional traits or temperament.

- *Sensorimotor*: This system covers sensorimotor processes, such as facial expressions, body posture, muscle action potentials, and central efferent activity, that not only regulate ongoing emotion experiences but are also capable of eliciting emotion. In particular, several theorists have argued that sensory feedback from facial expressions and body posture are causes of emotion [Duclos et al 1989]. Consider, for example, a person who is angry at another person and has a disposition to fight. The feedback from the response of the sensorimotor system, which includes changing the facial expression and clenching a fist, may intensify the experienced emotion.

- *Motivational*: This system includes all motivations that lead to emotion. In this thesis, and following both Tomkins and Izard [Tomkins 1962; Izard 1993], motivations include drives (e.g. *Thirst* and *Hunger*), emotions (e.g. *Anger*, and *Happiness*), and pain. Note also that in contrast with some other theories, emotions are distinguished from drives. As Tomkins suggested, drives are cyclical in nature, and are associated with and satisfied by a relatively restricted range of stimuli. Emotions, on the other hand, are not cyclical, can be related to an enormous variety of phenomena, and can motivate an equally wide range of cognition and actions [Tomkins 1962]. Some examples of elicitors in this system include the innate response to foul odors or tastes producing disgust, pain or aversive stimulation causing anger, and emotions like sadness eliciting others, such as anger, and vice versa. For a more complete review on this kind of elicitors see [Tomkins 1962; Tomkins 1963].

- *Cognitive*: This system includes all type of cognitions that activate emotion, such as appraisal and interpretation of events, comparisons, attributions, beliefs and desires, memory, and so on. Several of the theories that analyze different sets of cognitive appraisals or attributions which elicit, or are associated with, some specific emotions were previously described in 2.2.

In Cathexis, cognitive elicitors for the different basic emotions are based on an adapted version of Roseman's revised theory [Roseman 1990]. Several reasons influenced the decision for choosing Roseman's theory and not one of the others. First, Roseman's theory is broad enough in scope. That is, it attempts to address all emotion-eliciting stimuli that people normally face. Second, it presents clear testable predictions for the relationships between specific appraisal configurations and the emotional states they produce. Furthermore, these predictions are supported by empirical research as can be seen in [Roseman 1990]. And third, it fits well within the multi-system activation model described here since Roseman assumes that all emotions have a motivational basis, and that in some cases these motivations may be noncognitive.

The first change we made to Roseman's theory consisted on including a new appraisal. Namely, that of *unexpectedness* to reflect the results obtained in [Roseman 1990] which indicate that an appraisal of unexpectedness or novelty, rather than uncertainty, leads to surprise. Also, Roseman's theory considers sixteen discrete emotions, but some of these emotions are in the author's opinion, not emotions but *emotional attitudes* as in the case of like and dislike, or emotions accompanied of some sort of cognition as in the case of pride and hope. The second change consisted then, of simply considering those appraisal configurations which lead to the basic emotions for which Cathexis provides explicit models.

It should be mentioned, however, that even though Roseman's theory was the one selected as a basis for the cognitive elicitors, Cathexis' design is open and any of the other theories may very well be integrated or used as part of the model. In particular, the work of Scherer [Scherer 1988] may also be integrated in future work.

By considering both cognitive and noncognitive elicitors of emotion, the proposed model is more robust and gives developers more flexibility in defining the agent's affective characteristics.

3.1.5 Modeling Emotions, Moods, and Temperament in Cathexis

Most computational models of emotions do not differentiate between affective states, but rather they concentrate only on modeling emotion. To do this, most previous models of emotion include production systems as part of their architectures. In these systems, knowledge about the generation of emotions is represented in production rules with the traditional IF-THEN format. The left hand side of the production rules identifies the conditions that should be met to elicit some emotion. The right hand side of the rules labels the particular emotion. Figure 3-2 shows an example of a production rule taken from FEELER, a system developed by Pfeifer [Pfeifer 1988].

IF	current_state is negative for self and was caused by person p1 and p1 caused it intentionally and emotional_target is p1
THEN	ANGER at p1

Figure 3-2 Typical emotion generation rule. Example for the generation of anger.

Emotions, moods, and temperaments are modeled in Cathexis as a network composed of special emotional systems comparable to Minsky's "proto-specialist" agents [Minsky 1986][2]. Each of these proto-specialists represents a specific emotion from the list of basic emotions defined above (See 3.1.1). Within each proto-specialist, different sensors are monitoring both external (e.g. events in the environment) and internal sensory stimuli (e.g. drive levels, feedback from sensorimotor processes) for the existence of the appropriate

2. The idea of proto-specialists was influenced in part by Tinbergen's Central Excitatory Mechanism-Innate Releasing Mechanism (CEM-IRM) systems [Tinbergen 1969].

conditions that would elicit the emotion represented by that particular proto-specialist. These sensors are arranged into four different groups that correspond to the four different kinds of emotion activation systems described before (*Neural, Sensorimotor, Motivational,* and *Cognitive*). Input from these different sensors either increases or decreases the intensity of the emotion proto-specialist to which they belong. Associated with each proto-specialist are two threshold values. The first threshold, α, controls the activation of the emotion. That is, once the intensity goes above this threshold, the emotion proto-specialist becomes "active"[3] and releases its output signal to other emotion proto-specialists, and to the Emotional Behavior system which selects an appropriate behavior according to the state of these emotional systems (See 3.2). Figure 3-3 illustrates these ideas. The second threshold, ω, specifies the level of saturation for that emotion proto-specialist. That is, the maximum level the intensity of that particular proto-specialist can reach. This is consistent with real life emotional systems in which levels of arousal will not exceed certain limits.[4] Another important element associated with an emotion proto-specialist is a decay function, $\Psi()$, which controls the duration of the emotion episode as it will be explained below (See 3.1.7).

3. The word active is used here to represent that an emotion proto-specialist has released its intensity value to other emotional systems (Emotional Behavior system, other proto-specialists), which could be considered as the occurrence of an emotion episode. However, emotion proto-specialists are always active even if they are aroused at low levels as it is the case with moods.
4. I am grateful to Roz Picard for her comments and suggestions about thresholds and emotion intensities.

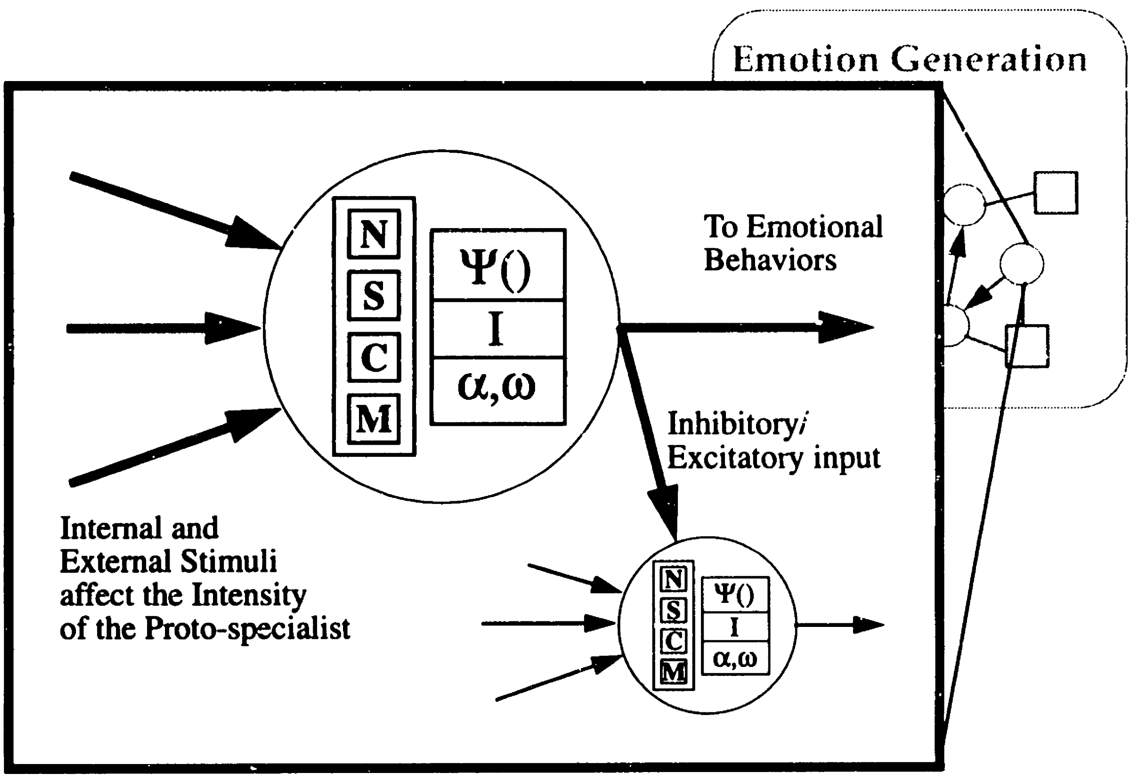

Figure 3-3 Emotion proto-specialists. Each basic emotion is modeled by one emotion proto-specialist. Associated with each emotion proto-specialist are several parameters: sensors arranged in four different groups (*Neural, Sensorimotor, Motivational, and Cognitive* activation systems), activation and clipping thresholds (α and ω), an intensity value (I), and a decay function Ψ().

All of these emotion proto-specialists run in parallel and are constantly updating their intensities. No one particular proto-specialist is in control of the system. In fact, as it was mentioned before, in order to model emotion blends and mixed emotions Cathexis allows for the co-occurrence of two or more basic emotions at the same time, which means that more than one emotion proto-specialist may have its intensity with values different than zero. Besides this kind of interaction, a more explicit one exists in which an emotion proto-specialist is connected to one or more other proto-specialists. In this way, Cathexis models emotion-elicited emotions as described in our discussion of motivational elicitors of emotion (See 3.1.4). This can certainly cause some problems in the system. Consider the case in which a proto-specialist arouses one or more different proto-specialists. These in turn might arouse some others, and so on and so forth, causing a spread of activity which would make them interfere with one another, thus blocking the normal functioning of the system. This is known as the "avalanche effect" [Minsky 1986] and to deal with this issue, the model uses different mechanisms in which active emotion proto-specialists suppress the activity of certain others (cross-exclusion), or send an inhibitory signal whose strength is proportional to their intensity (negative feedback). This type of interaction is consistent with real life emotional systems, in which high arousals in one or more systems will tend to inhibit other emotional systems.

It should be mentioned, however, that even though this inhibition is typical of certain emotions (e.g Fear inhibits Happiness, Happiness and Sadness inhibit each other, etc.), Cathexis does not dictate the specific emotions that should inhibit/excite each other, nor does it enforce the mechanism used (cross-exclusion or negative feedback). Here again, the model was designed to be open and flexible enough in such way that the mechanism is supported, but how it is used depends on the user of the model. This allows an agent developer or whoever uses Cathexis, to model different kinds of emotional phenomena, including perhaps emotional disorders.

So far it has been described how Cathexis models the basic emotions as well as emotion blends and mixed emotions. How are moods and temperaments modeled? Following the psychobiological perspective suggested in 3.1.3, moods are modeled as low levels of arousal within the same emotion proto-specialists. So, while high arousal of emotion proto-specialists will tend to inhibit other proto-specialists, mild arousal may very well allow several systems to be concurrently active, leading to the chance of an enormous display of mood states compared to the limited number of basic emotional states. This representation is consistent with the enormous subtleties of human moods and feelings, since the possible combinatorial states of the basic emotion proto-specialists (taking into account their overall intensities, time courses of activity, and the interactions within the elicitor systems) would be almost incalculable [Panksepp 1994]. It is also consistent with the common observation that moods seem to lower the threshold for arousing certain emotions because emotional systems (proto-specialists) that are aroused, as it happens in the representation of moods, are already providing some potential for the activation of an emotion. Finally, it is consistent with the observation that the duration of moods appears to be longer than that of the emotions, since at low levels of arousal, the intensity of the emotion proto-specialists will decay more slowly.

Temperaments, as defined in 3.1.3, are simply modeled through the different values the activation threshold α, can have for the different proto-specialists. Thus, an individual who has propensities to be in fearful moods might have a lower threshold for the emotion of fear in comparison to other individuals who do not.

3.1.6 Thresholds and Emotion Intensity

The preceding section mentioned the idea of thresholds, and levels of arousal. But what is a low or high level of arousal? How are these calculated? In other words, what determines the intensity of an emotion? This question has hardly been a subject of research. Besides the analyses by Ortony, Clore, and Collins [Ortony, Clore, and Collins 1988], Frijda [Frijda 1996], and Bower [Bower and Cohen 1982], emotion intensity as such, has not received a lot of attention.

Understanding how the intensity of an emotion is affected, amounts to specifying the precise nature of each of the factors that contribute to it. Thus, in Cathexis, the intensity of an emotion is affected by several factors including the previous level of arousal for that emotion (which takes into account the mood), the contributions of each of the emotion elicitors for that particular emotion, and the interaction with other emotions (inhibitory and excitatory inputs). All of this is modeled as described in equation (3.1):

$$EI_{it} = \chi\left(\left(\Psi(EI_{i(t-1)}) + \sum_{k} EL_{ki} + \sum_{l} (EX_{li} \cdot EI_{lt}) - \sum_{m} (IN_{mi} \cdot EI_{mt})\right),\ 0,\ \omega_i\right) \tag{3.1}$$

Where EI_{it} is the value of the intensity for Emotion i at time t; $EI_{i(t-1)}$ is its value at the previous time step; $\Psi()$ is the function that represents how Emotion i decays; EL_{ki} is the value of Emotion Elicitor k, where k ranges over the Emotion Elicitors for Emotion i; EX_{li} is the Excitatory Gain that Emotion l applies to Emotion i; EI_{lt} is the intensity of Emotion l at time t; IN_{mi} is the Inhibitory Gain that Emotion m applies to Emotion i; EI_{mt} is the intensity of Emotion m at time t; and $X()$ is the function used to constrain the intensity of Emotion i to be between 0, and the clipping threshold ω for Emotion i.

In turn, the value of each emotion elicitor depends on each of the conditions that contribute to the activation of that particular elicitor. Note that these conditions are elicitor-specific. For instance, the value of a cognitive elicitor, depends on the values for each of its eliciting conditions which in this case include motivational state, situational state, unexpectedness, probability, controllability, and agency as described in 2.2.4, whereas the value of a motivational elicitor depends on the level of a particular motivation (e.g. level of hunger). Equation (3.2) describes how this is modeled:

$$EL_{it} = \sum_k EC_{ki} \tag{3.2}$$

Where EL_{it} is the value of Emotion Elicitor i at time t, and EC_{kj} is the value of Eliciting Condition k, where k ranges over the Eliciting Conditions for Emotion Elicitor j.

Eliciting Conditions have a numeric value and weight associated with them. These weights vary depending on the eliciting condition and the emotion elicitor to which the condition belongs. The overall value of an eliciting condition is determined by equation (3.3):

$$EC_{it} = v_i \cdot w_i \tag{3.3}$$

Where EC_{it} is the value of Eliciting Condition i, and v and w are the value and weight for Eliciting Condition i.

Finally, a brief discussion on the role of thresholds is appropriate. The role of the clipping threshold ω, was described in equation (3.1). As for the activation threshold, we mentioned that an emotion proto-specialist is activated (i.e. it releases its intensity value to the Emotional Behavior System and inhibits/excites other proto-specialists) if and only if its intensity goes above a specified threshold α. This is modeled by equation (3.4):

$$O_{it} = \begin{cases} EI_{it}, & \text{if} \quad (EI_{it} > \alpha_i) \\ 0, & \text{otherwise} \end{cases} \tag{3.4}$$

Where O_{it} is the output for Emotion i at time t; EI_{it} is the Intensity of Emotion i at time t; and α_i is the activation threshold for Emotion i.

3.1.7 Emotion Decay

If the concept of emotion intensity has received little attention, the concept of emotion decay has received much less. Perhaps this is why most previous models of emotion do not consider this issue. With the current evidence on this matter one can only speculate as to what causes emotions to decay and what are the mechanisms and the systems involved in this process. Nevertheless, the fact still remains that emotions, once generated, do not remain active forever and a model of emotions should be able to handle this situation.

In Cathexis, each emotion proto-specialist has been provided with a built-in decay function (depicted in Figure 3-3 as $\Psi()$) which controls the duration of the emotion once it has become active. This function is specific to each emotion proto-specialist which allows for different models of emotion decay. It should also be noted that the model does not enforce any specific implementation of these functions. That is, the decay function for an emotion can be implemented as a constant of time or as a more complex function of the different elicitors for that particular emotion (e.g. in terms of resolution of goal-centered issues). Regardless of their implementation, in every update cycle the decay functions of the active emotions are evaluated. These functions make sure that unless there is some excitatory input for the emotion, its intensity is lowered and after a few cycles, once its intensity goes below the specified threshold, it becomes inactive.[5]

3.2 Emotional Behavior System

The previous section was concerned with how emotions, moods, and temperaments were modeled and generated. This section discusses how these emotional systems influence and affect behavior. Researchers such as Frijda, suggest that emotional systems may elicit physiological changes, and they may elicit feelings, but unless they imply or involve changes in the readiness of specific behaviors they are of no consequence for either the individual or for others [Frijda 1996]. These changes are modeled in the Emotional Behavior System. In other words, the Emotional Behavior System decides what Emotional Behavior is appropriate for the agent to display, given its emotional state at some point in time.

The Emotional Behavior system is a distributed system composed of a network of behaviors, such as "engage in fight", "avoid anger", or "smile". Each of these behaviors competes for the control of the agent. The decision of which behavior should be active is based on the value of each of the behaviors, which is recalculated on every cycle as described below. Each Emotional Behavior contains two major components: the Expressive or Motor Component and the Experiential Component, as shown in Figure 3-4.

5. Note again the use of the words active and inactive here, which correspond to the occurrence of an emotion episode and its later decay. However, remember that if the intensity of an emotion proto-specialist is below the activation threshold, this only means, unless the intensity is zero, that a low level of arousal exists and therefore a mood.

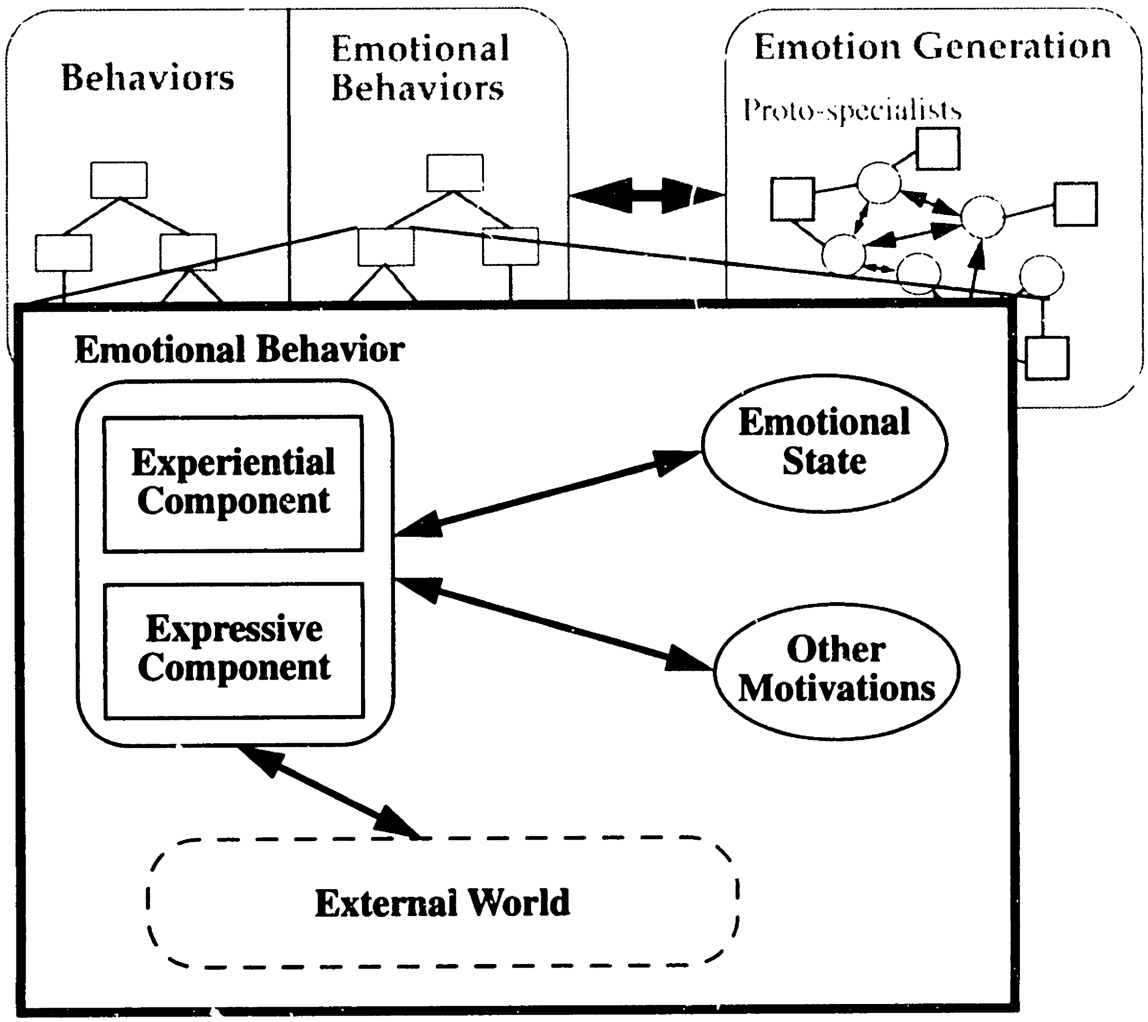

Figure 3-4 Emotional Behaviors, when active, issue motor commands to modify the agent's facial expression and body posture, and may also modify the agent's motivational state (emotions, moods, drives).

3.2.1 Expressive Component

Several researchers agree that emotions and moods include an expressive or motor component, or at least, an efferent activity in the central nervous system [Plutchik 1994; Izard 1993]. Certainly, the expression of emotion is one of the most important aspects in Emotional Behavior. In Cathexis, the expressive component of an Emotional Behavior contemplates the following aspects:

- *Prototypical facial expression*: Motor commands to alter the facial expression in accordance to the specific Emotional behavior.

- *Body posture*: Series of motor commands to alter the agent's body posture in accordance to the specific Emotional Behavior.

- *Vocal expression*: May consist of non-language sounds such as laughing, yawning, grunting, etc., or it may include more complex elements to control speech, such as loudness, pitch, and temporal variability.

3.2.2 Experiential Component

The other component of an Emotional Behavior represents the emotion experience. As described in Chapter 1, the experience of emotions and moods bias action and cognition. Several researchers agree that the experiential component of emotions and moods can be identified as motivational processes that manifest themselves as action readiness, perceptual biases, selective filters for memory and learning, and a feeling state [Bower and Cohen 1982; Izard 1993; Niedenthal and Kitayama 1994]. The main aspects considered in this model include:

- *Motivation*: As part of an Emotional Behavior, motivations are specified as the main causes of activation or release of that behavior. They are also specified within the experiential component because an Emotional Behavior, when active, may influence different motivational systems. Behaviors affect the levels of drives, such as hunger or thirst, but they may also affect emotions and moods as it was discussed with sensorimotor elicitors of emotion (See 3.1.4).

• *Action tendency*: According to many researchers, among which Frijda is perhaps the main supporter of the concept, changes in action readiness or action tendencies are a major, if not the major, aspect of the response to an event of emotional significance [Frijda 1986]. In Cathexis, action tendency and readiness are modeled by the Emotional behavior itself. Thus, an Emotional Behavior, such as aggression, which may be represented as several different behaviors including perhaps fighting, insulting, biting, etc., is modeled as a direct response to the emotion of anger, or an irritable mood.

Other influences of emotions and moods, such as *perceptual biases and selective filtering* include complex interactions with cognitive systems and constitute by themselves whole fields of active research which are beyond the scope of this work. Nevertheless, we believe it is important to acknowledge their existence and we have tried to design Cathexis as an open model that allows incremental additions of systems or components that model these influences. Chapter 6 discusses some of these possibilities as future research.

3.2.3 Selection of Emotional Behavior

As mentioned before, Emotional Behaviors compete with each other to obtain control of the agent. Competition is based on the values of each emotional behavior. That is, the behavior with the highest value at some point in time, becomes the active behavior. This value is updated every cycle, and as seen in Figure 3-4, it depends on several factors, called the *releasers* for that behavior, which may include motivations (emotions, moods, drives, pain), as well as a wide variety of external stimuli.[6] The value of an Emotional

---

6. The idea of releasers or releasing mechanisms is taken from the field of Ethology. Primarily, from the work of Tinbergen [Tinbergen 1969].

Behavior is summarized in equation (3.5):

$$EB_{it} = \sum_k BR_{ki} \tag{3.5}$$

Where EB_{it} is the value of Emotional Behavior i at time t; and BR_{ki} is the value of Behavior Releaser k, where k ranges over the different releasers relevant to Emotional Behavior i.

The value of each of these Behavior Releasers depends on the specific nature of the releaser. For instance, if the releaser is an internal motivation, such as the drive of hunger, or a particular emotional state, its value is calculated in terms of the intensity of that motivation. On the other hand, if the releaser consists of some sort of external stimuli, then its value depends on sensory input which determines if the stimulus of interest is present and matches some specific criteria.

This model of Emotional Behavior is certainly simple. A more complex model may include inhibitions between behaviors in a similar fashion to that of inhibitory inputs for emotional systems (See 3.1.5), and it may also include models of Behavior depletion or fatigue as those described by Ludlow [Ludlow 1980]. It should be noted, however, that this part of the model has intentionally been made simple and general enough so that Cathexis can be used and integrated into several different architectures of action-selection, such as those of Maes [Maes 1991], and Blumberg [Blumberg and Galyean 1995].

Chapter 4

Implementation

This chapter describes an implementation for the computational model described in Chapter 3. This implementation is described with fair detail not only to discuss and explain some of the intricate concepts of the model, but also to give agent developers ideas and suggestions on how to use it and extend it. A fundamental knowledge on Object-Oriented Programming (OOP) concepts and design notation is assumed.[1]

Readers not interested in this level of detail may skip this chapter without losing the sense of later chapters.

4.1 Implementation

Cathexis has been implemented in its totality as part of an object-oriented framework (a collection of classes designed to form a cohesive whole aimed at a particular problem domain, in our case, that of modeling emotions) that allows agent developers to create emotional agents. This framework has been implemented in the C++ programming language and another implementation in the Java programming language is under way.

The design of the framework is guided by the following principles:

- classes as designed as small as possible

1. For more information on these concepts and the notation used in this chapter, the reader is referred to: Booch, Grady. *Object-oriented Analysis and Design.* Redwood City, CA: Benjamin Cummings, 1994

- classes depend as little as possible on other classes of the framework
- the power of derivation, multiple inheritance and polymorphism has been used to a great extent to add functionality to the framework and provide simple mechanisms that allow developers to extend the framework

These principles ensure that the framework is small enough that can be learned quickly and flexible enough that it can be used at the convenience of the developer. That is, each of the components can be used separately, say to model emotional phenomena (using only the emotion generation module), or to create emotional agents (using the emotion generation and emotional behavior modules).

As seen in Figure 3-1, the main two modules in Cathexis are those of Emotion Generation and Emotional Behavior. Each of these modules contains representations of several important concepts including emotion proto-specialists, emotion elicitors, behavior releasers, emotional behaviors, and so on. The following sections describe how each of these concepts are implemented within the framework.

4.1.1 Motivations: Drive and Emotion Proto-specialists

A very important concept in Cathexis is that of a *motivation*. In both Chapter 1 and Chapter 3, we discussed the idea of emotions as motivational systems. Furthermore, we mentioned that emotions were not the only kind of motivations in the system, but that drives, such as hunger or thirst were also important motivations that interacted in significant ways with these emotional systems.

The relationship between drives, emotions, and the concept of motivations suggests the first use of inheritance in the framework. In order to define all the characteristics and protocol interface that drives and emotions share in common as motivational systems, we have created an abstraction represented by the class `Motivation`. The definition for this class is illustrated in Figure 4-1. This class is designed as an abstract class because by itself, an instance of it would have no useful meaning. Instead, it is expected that its concrete subclasses will add structure and behavior by implementing its abstract operations.

```
class Motivation: public Dependent {
protected:
    RWCString mName;
    float mIntensity;
public:
    Motivation(const RWCString & name);
    virtual ~Motivation();

    RWCString GetName() const;
    float GetIntensity() const;

    void SetIntensity(float intensity);
    virtual int operator==(const Motivation & aMotivation);

    virtual void UpdateIntensity() = 0;

    virtual void Trigger(const RWCString & level);
    virtual void Activate();
};
```

Figure 4-1 Definition of class Motivation

The code excerpt shows that an instance of this class, a specific motivation, has a name and an intensity data members associated with it. Note that the method `UpdateIntensity()`, has no implementation and it only serves as a definition of common interface for all motivations.

Abstractions for drives and emotions are written as derived classes of the `Motivation` class as illustrated in Figure 4-2.

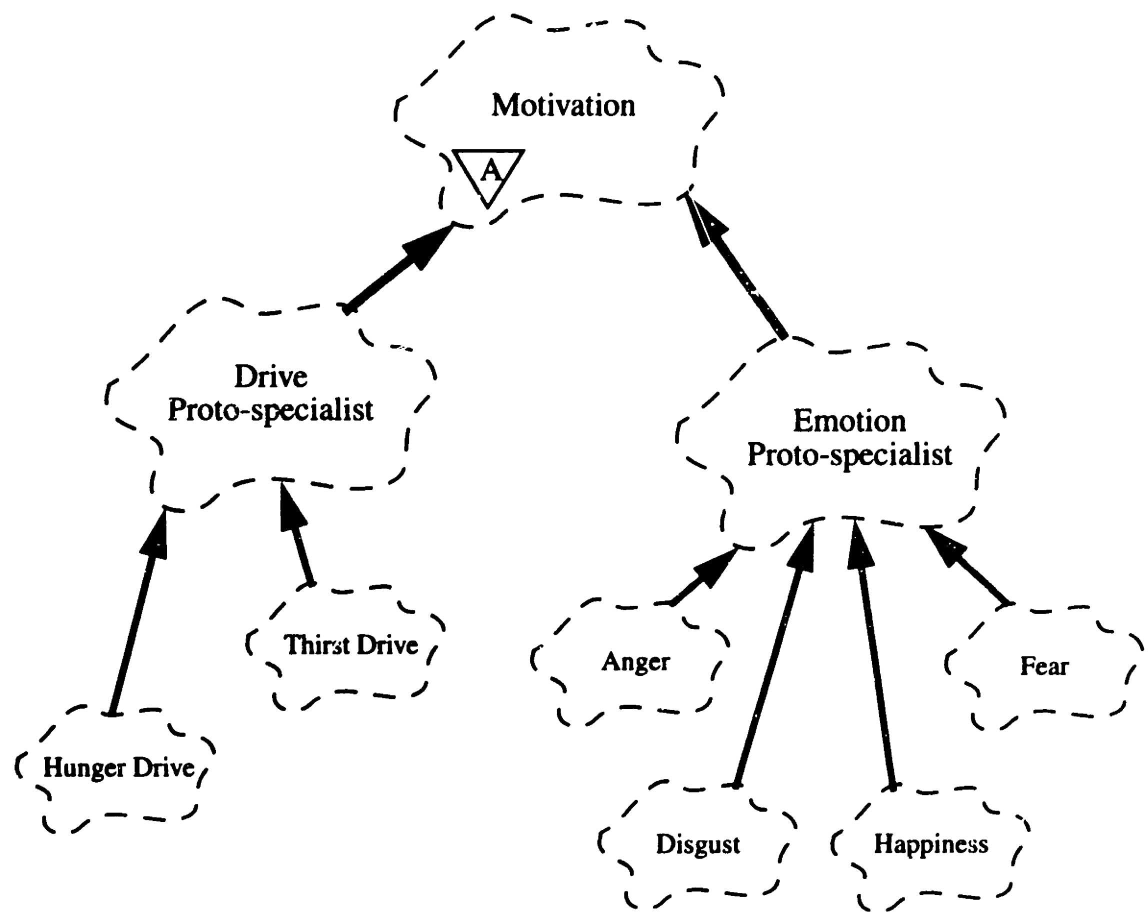

Figure 4-2 Motivations Class Diagram

Drives are modeled as reactive machines, similar to the emotion proto-specialists defined in 3.1.5. The differences between drive proto-specialists and emotion proto-specialists are captured within each class as seen in Figure 4-3 and Figure 4-4. Given that our main concern relies on the model of emotions, some of the details on the class definition of drive proto-specialists have been omitted. It should be mentioned, however, that drive proto-specialists have two important data members which correspond to the growth and damp rates. The values of these rates are used in conjunction with the value returned by the method `getFeedbackEffect()`, which represents the effects of behaviors on a particular drive, to update the level of intensity for a particular drive on every update cycle. These values are specified as part of the constructor method for a specific drive proto-specialist. This means that a developer can create different instances of the class `DriveProtoSpecialist` and specify different growth and damp rates for each specific instance. However, instead of creating different instances that represent different drives, we encourage the developer to use inheritance and derive specific subclasses that represent more specialized drives. As a sample implementation, we have provided subclasses that represent the drive of Hunger and the drive of Thirst.

```
class DriveProtoSpecialist: public Motivation {
    float mGrowthRate;
    float mDampRate;
public:
    DriveProtoSpecialist(const RWCString & name, float growth,
                                 float damp);
    virtual ~DriveProtoSpecialist();

    virtual void UpdateIntensity();
    virtual void Trigger();
    virtual void Activate();
protected:
    virtual float getFeedbackEffect();
};
```

Figure 4-3 Definition of class DriveProtoSpecialist

Emotional systems are the core of the framework. These systems are represented with the class `EmotionProtoSpecialist` as illustrated in Figure 4-4. This class definition deserves, of course, a more detailed explanation.

```
class EmotionProtoSpecialist: public Motivation {
    ...
protected:
    float mAlphaThreshold;
    float mOmegaThreshold;

    Dictionary<RWCString,Elicitor> * mpElicitors;
    List<EmotionProtoSpecialist> * mpInhibitoryList;
    List<EmotionProtoSpecialist> * mpExcitatoryList;

public:
    EmotionProtoSpecialist(const RWCString & name, float alphaT,
                           float omegaT);
    virtual ~EmotionProtoSpecialist();
    ...
    float GetAlphaThreshold() const;
    float GetOmegaThreshold() const;
    void SetAlphaThreshold(float threshold);
    void SetOmegaThreshold(float threshold);

    void AddElicitor(const RWCString & type, Elicitor * elicitor);
    Elicitor * RemoveElicitor(Elicitor * anElicitor);

    void InhibitEmotion(EmotionProtoSpecialist * anEmotion,
                        float inhibitoryGain);
    void ExciteEmotion(EmotionProtoSpecialist * anEmotion,
                        float excitatoryGain);
    ...
    virtual float DecayFunction();
    void InhibitPeers();
    void ExcitePeers();
    ...
    virtual void UpdateIntensity();
    void Trigger(const RWCString & level);
    virtual void Activate();
};
```

Figure 4-4 Definition of class EmotionProtoSpecialist

All of the concepts described in 3.1.5 for emotion proto-specialists are implemented in this class. The activation and clipping thresholds, α, and ω, are implemented as floating point values that are passed to the constructor method and can be accessed and modified at any moment in time.

Each instance of an emotion proto-specialist has a set of emotion elicitors corresponding to the different emotion activation systems defined in 3.1.4. This set of elicitors is implemented with a dictionary class that allows for the fast indexing of different emotion elicitors within each activation system.

The different kinds of emotion elicitors are modeled with several different abstractions as seen in Figure 4-5. In general, instances of emotion proto-specialists will have several emotion elicitors within each system. These elicitors can be added or removed to an instance of an emotion proto-specialist, by calling the corresponding `AddElicitor()` and `RemoveElicitor()` methods, both of which rely on inheritance and polymorphism to provide a common interface for all different kinds of elicitors.

An emotion proto-specialist may inhibit or excite other emotional systems as it was discussed in 3.1.5. Two different data members are provided within the class to keep track of which emotion proto-specialists should be sent inhibitory or excitatory inputs. Different emotion proto-specialists can be added to these lists by calling the corresponding `InhibitEmotion()` and `ExciteEmotion()` methods.

The intensity of a particular emotion proto-specialist is calculated in the `UpdateIntensity()` method based on the equations described in 3.1.6. How this intensity decays depends on the implementation of the `DecayFunction()` method which is defined as `virtual` so that it can be redefined for the different basic emotions.

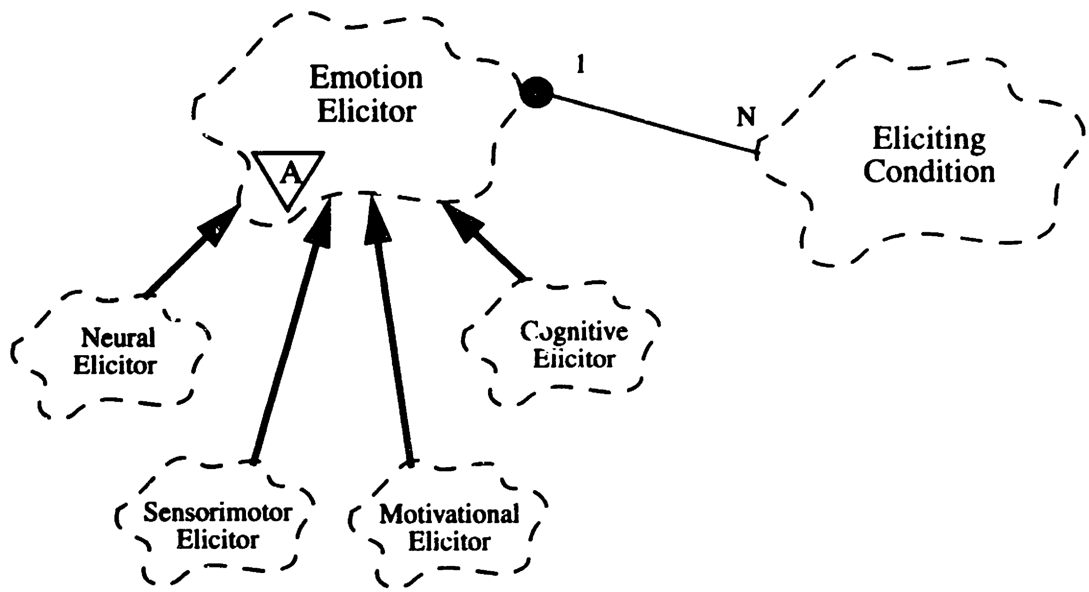

Figure 4-5 Emotion Elicitors Class Diagram

Other methods that perform management functions for the `EmotionProtoSpecialist` class, have been omitted from this discussion.

As it can be seen, this class provides all the elements and mechanisms necessary to model emotional systems as those described in Chapter 3. Although some default values are suggested for some of these elements, the model and its implementation do not enforce any particular settings.

Developers can create direct instances of the `EmotionProtoSpecialist` class to represent different emotions such as anger, fear and so on. However, as it is the case with drive proto-specialists, we encourage the use of inheritance to achieve this same purpose. As a matter of fact, the framework already includes sample implementations for each of the basic emotions described in Chapter 3 which are readily available to developers.

Although other abstractions exist in the emotion generation module, we have described the ones we believe are the most important and will be most likely used by developers to model emotion phenomena and or create emotional agents.

4.1.2 Emotional Behaviors

This section will briefly describe some of the implementation details for another important abstraction in the system: an emotional behavior.

Emotional Behaviors are modeled as subclasses of the abstract class `Behavior`. The definition for this class is illustrated in Figure 4-6.

```
class Behavior: public Dependent {
    RWCString mName;
    float mValue;
    List<Releaser> * mpReleasers;
public:
    Behavior(const RWCString & name);
    virtual ~Behavior();

    RWCString GetName() const;
    void SetName(const RWCString & name);
    float GetValue() const;
    void SetValue(float value);
    ...
    void AddReleaser(Releaser * releaser);
    Releaser * RemoveReleaser(Releaser * releaser);

    virtual void Trigger(const RWCString & level);
    virtual void UpdateValue();
    virtual void Activate() = 0;
    ...
};
```

Figure 4-6 Definition of class Behavior

Behaviors, as defined in 3.2, compete to obtain control of the agent. This competition is based on their value which is here implemented as a floating point number. The value of a particular behavior depends on the contributions of the different releasers for that behavior. Releasers are modeled as a separate abstraction to allow different kinds of releasers (external stimuli, motivations, etc.) to be used with a single common interface. Releasers can be added to or removed from a behavior list of releaser by calling the corresponding `AddReleaser()` and `RemoveReleaser()` methods defined in the `Behavior` class.

The definition of the more specific `EmotionalBehavior` class is shown in Figure 4-7.

```
class EmotionalBehavior: public Behavior {
protected:
    List<EmotionalExpression> * mpExpressiveComponent;
    List<EmotionInfluence> * mpExperientialComponent;
    ...
public:
    EmotionalBehavior(const RWCString & name);
    virtual ~EmotionalBehavior();

    void AddExpression(EmotionalExpression * facialExp);
    void AddInfluence(EmotionInfluence * influence);

    virtual void Trigger(RWCString & level);
    virtual void Activate();
};
```

Figure 4-7 Definition of class Emotional Behavior

This class implements the different concepts discussed in section 3.2. Specifically, every instance of an Emotional Behavior has an expressive component and an experiential component. These two components are implemented as lists containing instances of the classes `EmotionalExpression` and `EmotionInfluence`, respectively. The class `Emotion-`

`alExpression` defines the characteristics and functionality shared by all emotional expression. Inheritance is used again to specify concrete expressions such as facial expressions, body posture, and so on.

Similarly, the class `EmotionInfluence` defines the characteristics and functionality shared by emotion influences, such as influences in motivations, perceptual biases, and the like.

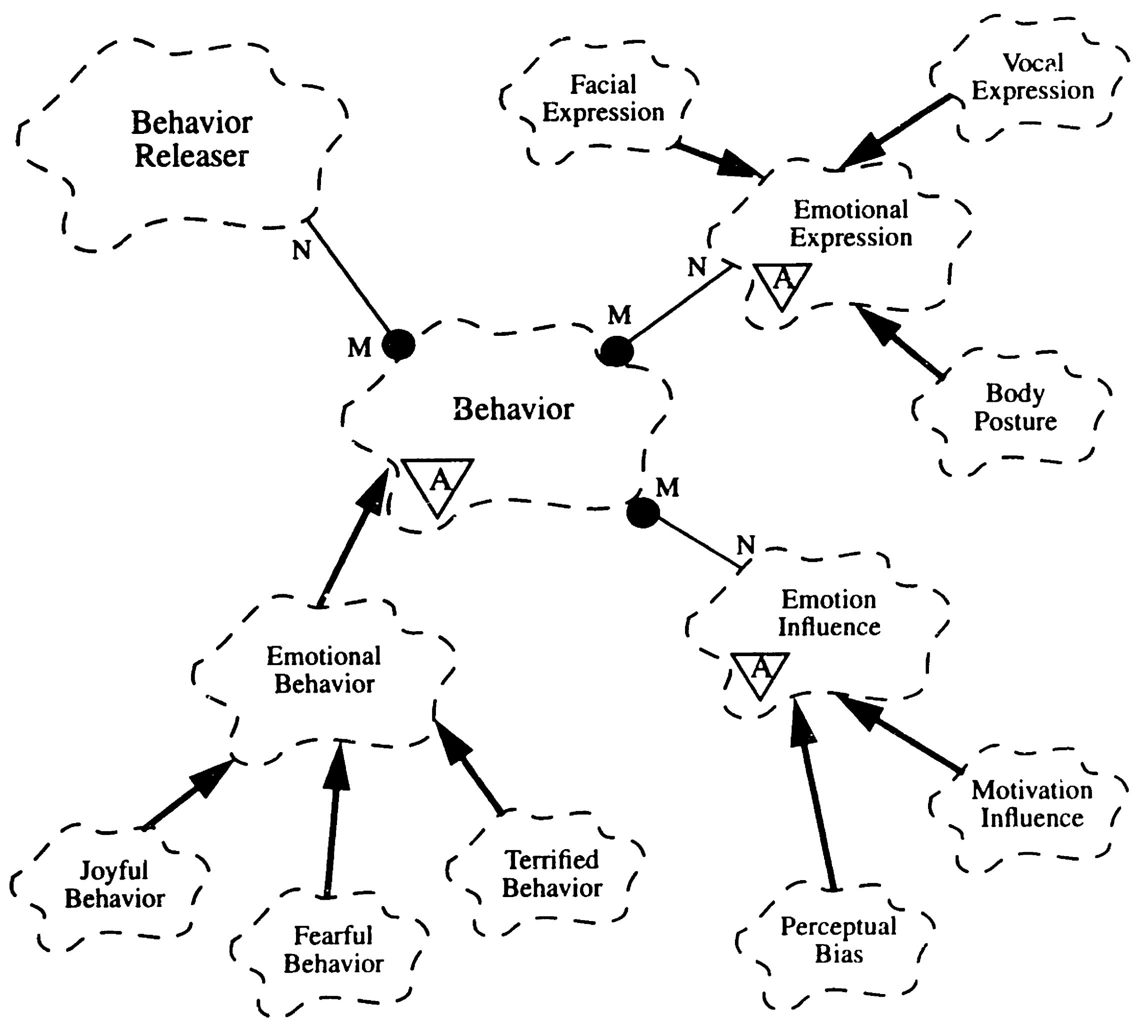

Figure 4-8 Emotional Behaviors Class Diagram

A more comprehensive view of the relationships between these classes can be seen in Figure 4-8.

The framework provides sample implementations of several kinds of Emotional Behaviors as appropriate reactions to the occurrence of different emotion episodes and moods. However, it is expected that developers will create their own Emotional Behaviors depending on their specific interests.

As we mentioned in 3.2.2, so far Cathexis only considers influences in motivations and action readiness as part of the experiential component of an Emotional Behavior. However, the framework has been designed so that developers can integrate other models of emotion influences if desired or needed, through the use of inheritance.

4.1.3 Communication Between Modules

The Emotion Generation module can be used separately to model different sorts of emotion phenomena. However, if the developer wants to create agents that display Emotional Behaviors appropriate to different emotional situations that rise in the environment, the Emotional Behavior module should be used as well. This implies that both modules need to communicate in some way. To deal with this issue, the framework includes another important abstraction that serves the purpose of current working memory for the agent and at the same time provides a shared memory space that can be used for communication purposes. This abstraction is modeled as the `Blackboard` class, which as its name implies, is an implementation of a traditional blackboard structure.

All the different kinds of events that might happen in relation to the agent can be posted on the blackboard so that all modules have access to them. These events may be caused by external stimuli coming from the environment, or they may be the product of cognitions and internal stimuli caused by changes in the different motivational systems of the agent (drives, and emotional systems). For instance, something that happens in the environment will be posted as an environment event in the blackboard. This event may trigger some systems to become active which may in turn generate different kinds of events, such as cognitive events, motivational events, and so on.

Figure 4-9 illustrates the different kinds of events already implemented as part of the framework, and their relationship with other abstractions.

Following the same principles of design for the framework, the developer can design and implement different kinds of events if needed without having to alter any of the internal mechanisms.

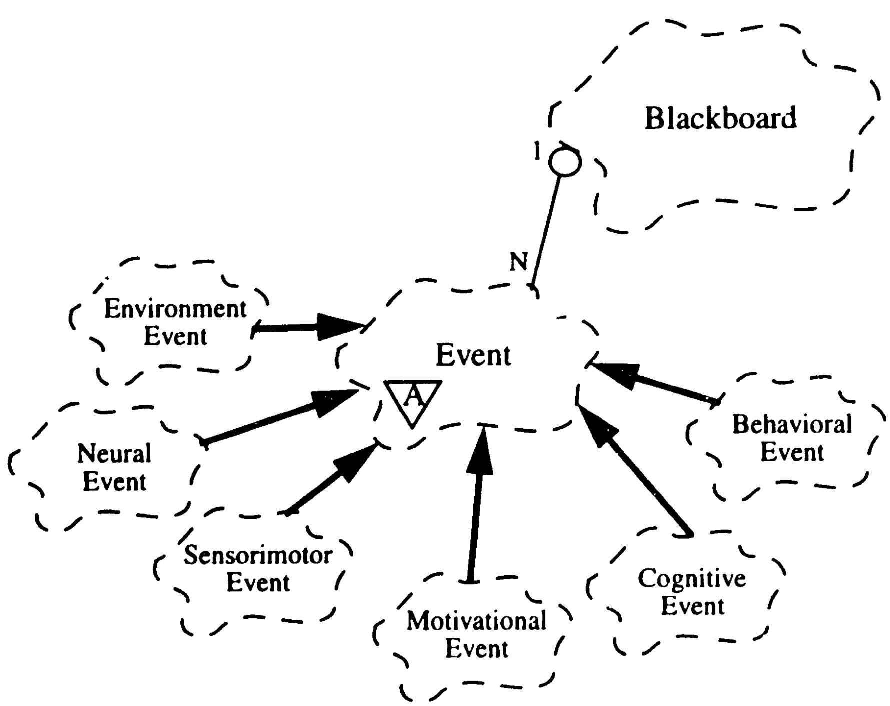

Figure 4-9 Blackboard and Events Class Diagram

4.1.4 Agents

Certainly, an agent developer is interested in designing an implementing agents. An abstraction designed to facilitate the creation of emotional agents is provided within the framework. This abstraction is illustrated in Figure 4-10. Many of the implementation details have been omitted and only those relevant to this discussion are showed.

```
class Agent{
    Dictionary<RWCString, Motivation> * mpMotivations;
    List<Behavior> * mpBehaviors;
    Blackboard * mpBlackboard;
    Behavior * mpActiveBehavior;
    ...
public:
    Agent();
    virtual ~Agent();
    ...
    Motivation * GetMotivation(const RWCString & motivation);

    void AddMotivation(Motivation * motivation);
    Motivation * RemoveMotivation(Motivation * motivation);

    void AddBehavior(Behavior * behavior);
    Behavior * RemoveBehavior(Behavior * behavior);

    void UpdateMotivations();
    void UpdateDrives();
    void UpdateEmotions();
    void SelectBehavior();
private:
    void initBlackboard();
    void initEmotions();
    void initDrives();
    void initBehaviors();
};
```

Figure 4-10 Definition of class Agent

The class `Agent` already implements all the relationships between the different modules as it is illustrated in Figure 4-11. It also provides an implementation of an algorithm for emotion generation and action-selection which is described below in 4.1.5.

To create emotional agents using this framework, developers need only specify what are the motivations (drives and emotions) and behaviors for a specific agent. To aid in this purpose, the class `Agent` provides several methods to add, remove, and manage both motivations and behaviors. These methods rely heavily on inheritance and polymorphism so that different subclasses of these abstractions can be used, which provides the developer with a lot of flexibility and power.

As it was mentioned before, several implementations of Drives, Emotions, and Emotional Behaviors are readily available in the framework for the developer to use in addition to any other abstractions he or she designs.

Again, instead of creating specific instances of this class, the developer might want to design specific agents as derived classes.

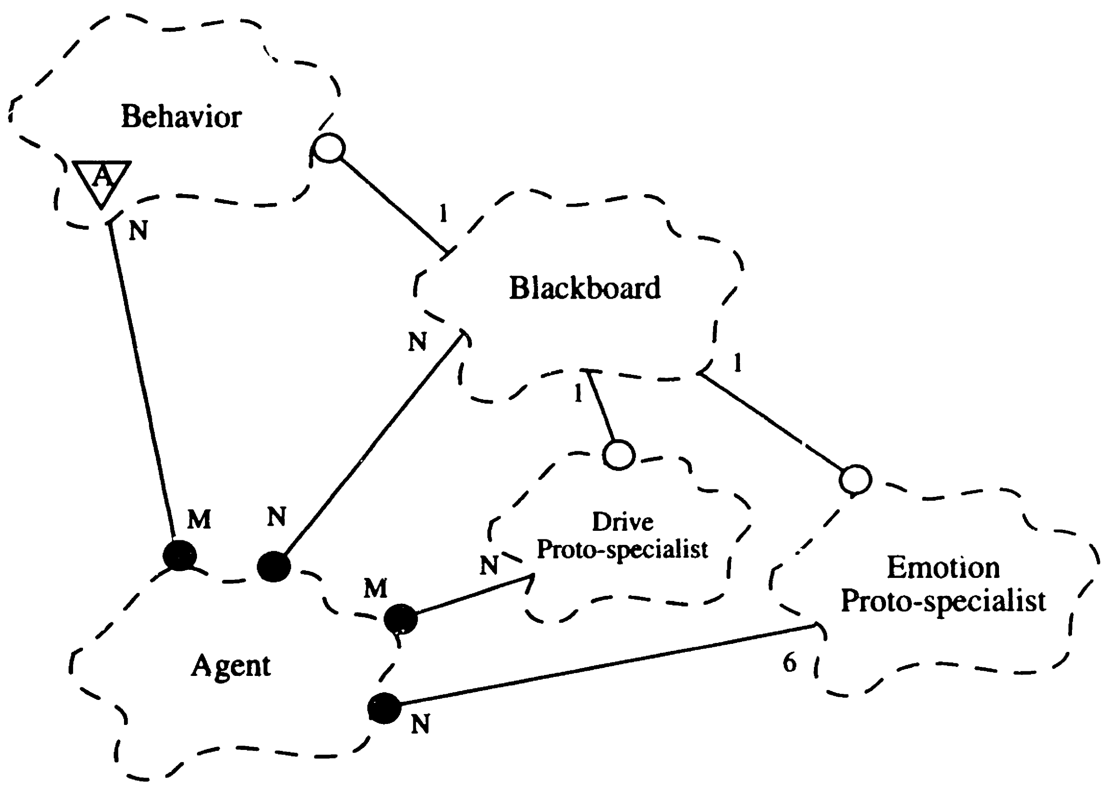

Figure 4-11 Agent Class Diagram

4.1.5 Main Algorithm

The algorithm for emotion generation and action-selection included in the implementation of the class `Agent` is described below. The interaction diagram for the main instances involved in the process is illustrated in Figure 4-12.

On each update cycle:

1. Any event in the external environment is posted as an environment event on the blackboard.
2. The values for all of the agent's motivations are updated. This is done as a two-step process as follows:
   a. All drive proto-specialists update their values. This is done as follows:

- The specific growth and damp rates for this drive proto-specialist are applied to its current intensity.
- The effects from any relevant behavioral activity are considered and applied to its intensity.
- The change of intensity value is posted as a motivational event on the blackboard.

b. All emotion proto-specialists update their values. This is done as follows:

- If the emotion proto-specialist is active, its decay function is evaluated.
- The different elicitors for this emotion proto-specialist are evaluated against the different events posted in the blackboard.
- The inhibitory and excitatory inputs from other emotion proto-specialists are subtracted and added to the intensity of this emotion proto-specialist.
- If necessary, the intensity of this emotion proto-specialist is constrained to remain between 0 and the clipping threshold ω.
- If the intensity of this emotion proto-specialist is above its defined activation threshold α, its intensity value is released, inhibiting and/or exciting the appropriate emotion proto-specialists.
- The change of intensity value is posted as a motivational event on the blackboard.

3. The values of all Emotional Behaviors are updated based on the current sensory stimuli (external stimuli and internal motivations).
4. The Emotional Behavior with the highest value becomes the active behavior. This event is posted as a behavioral event on the blackboard.
5. If the active Emotional Behavior has an expressive component attached, it sends the appropriate motor commands to modify the agent's expression.
6. If the active Emotional Behavior has an experiential component, it evaluates it and updates all appropriate motivations.

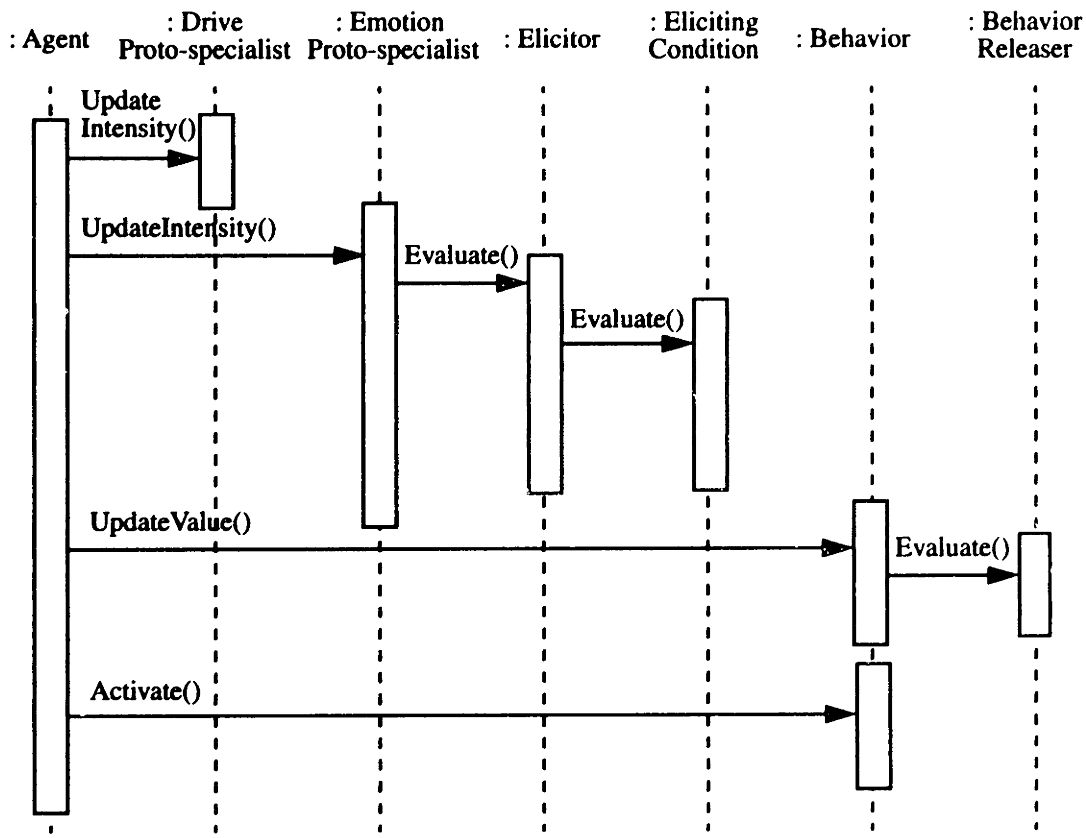

Figure 4-12 Interaction Diagram for thc Scenario for Emotion Generation and Action-selection.

Chapter 5

Results

5.1 Simon the Toddler: A Test-bed Environment

The framework described in the previous chapter has been used to build an environment in which the user interacts with "Simon", a synthetic agent representing a young child. The main goal of creating Simon was not to create an emotional, believable agent, but rather to build an environment that would serve the purpose of a test-bed in which we could experiment and test the internals of the Cathexis model, while at the same time, evaluate how useful the framework is as a tool to create models of emotional phenomena and emotional agents.

5.1.1 Simon's Motivations

In order to test the relationships between motivations (drives, emotions, and pain), we designed Simon to have several instances of different motivational systems. Specifically, we created four drive proto-specialists with different growth and damp rates as explained in 4.1.1, to represent the drives of *Hunger*, *Thirst*, *Fatigue* (the need to rest and sleep), and *Interest* (the need to explore and play with things). Additionally, we used the existing emotion proto-specialists for the six different basic emotions (*Enjoyment/Happiness*, *Distress/Sadness*, *Fear*, *Anger*, *Disgust*, and *Surprise*).

5.1.2 Simon's Behaviors

We gave Simon a repertoire of behaviors according to the different proto-specialists that we included as part of his motivational system. For each of the six basic emotion proto-specialists we created three (in some cases only two) Emotional Behaviors that correspond to different intensities of that particular emotion. For instance, for the case of *Anger*, we created three different Emotional Behaviors which represent Simon's actions when he becomes *Angry*, *Furious*, and *Raged*. Each of these Emotional Behaviors have different expressive and experiential components as described in 3.2. The expressive component includes different facial expressions similar to those shown in Figure 5-1, and for some behaviors, it also includes vocal expressions in the form of non-language sounds such as a cry, or a laugh (implemented as system sounds on a Power Macintosh). The experiential component includes influences in the motivations of the agent, such as lowering the intensity of the *Hunger* drive and increasing the intensity of *Fatigue* when the *Sorrowful* behavior has been active for some time, and specific actions such as biting, laughing, and crying which in this environment are implemented as messages that get displayed once the behavior is active.

Other behaviors include *Sleep*, *PlayWithToy*, *Eat*, and *Drink*, which as their names imply, were created to satisfy each of the drive proto-specialists. Figure 5-2 shows an example of an Emotional behavior.

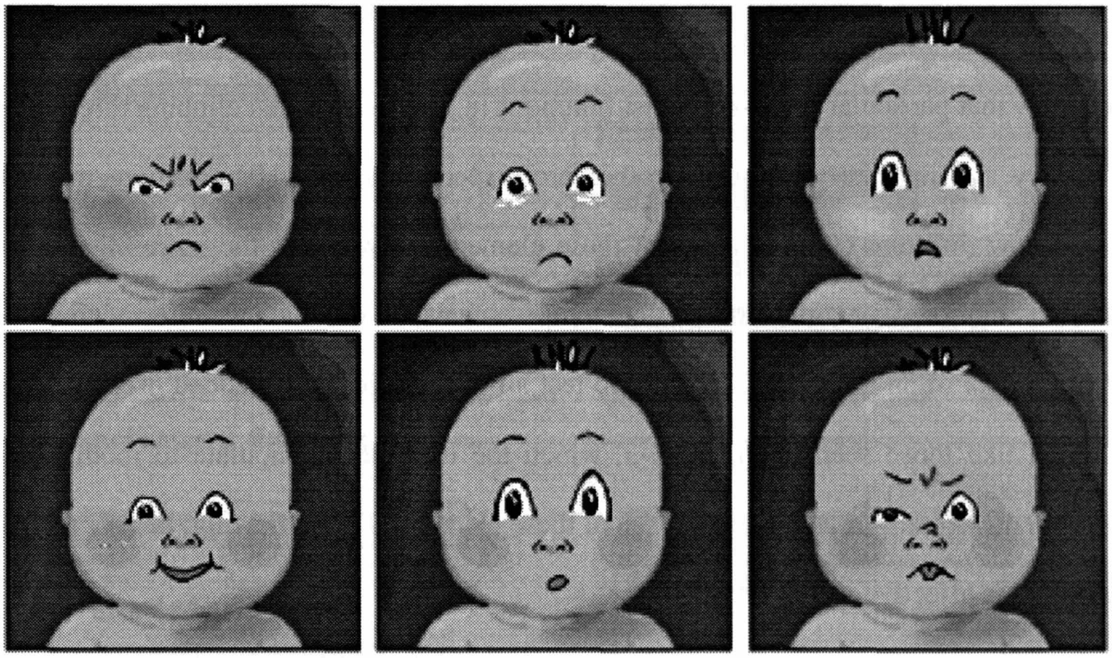

Figure 5-1 Facial Expressions for Different Emotional Behaviors

```
Emotional Behavior: Sorrowful
Releasers: Sadness Emotion present, intensity above 1.66
Expressive Component
  Facial Expression: PICT ID = 'sorrow'
  Vocal Expression: SND ID = 'cry'
  Body Posture: NIL
Experiential Component
  Action: "Simon is pouting and crying"
  Influence Motivations: Decrease Hunger by (BV * HungerGain)
                         Increase Fatigue by (BV * FatigueGain)
```

Figure 5-2 Description of the Sorrowful Emotional Behavior

5.1.3 Simon's Affective Style

As we discussed in 3.1.5, each emotion proto-specialist has different elements that determine how that particular proto-specialist functions (e.g. activation and clipping thresholds, excitatory and inhibitory outputs to other proto-specialists, etc.). Furthermore, we discussed how different configurations of these elements allowed for the representation of different emotional reactions, moods, and temperaments, all of which we referred to as the *affective style* of an individual. In order to test all of these ideas, we added user interface controls, like those seen in Figure 5-3, which the user can manipulate to modify and experiment with different values for each of these elements.Thus, allowing the user to control Simon's affective style.

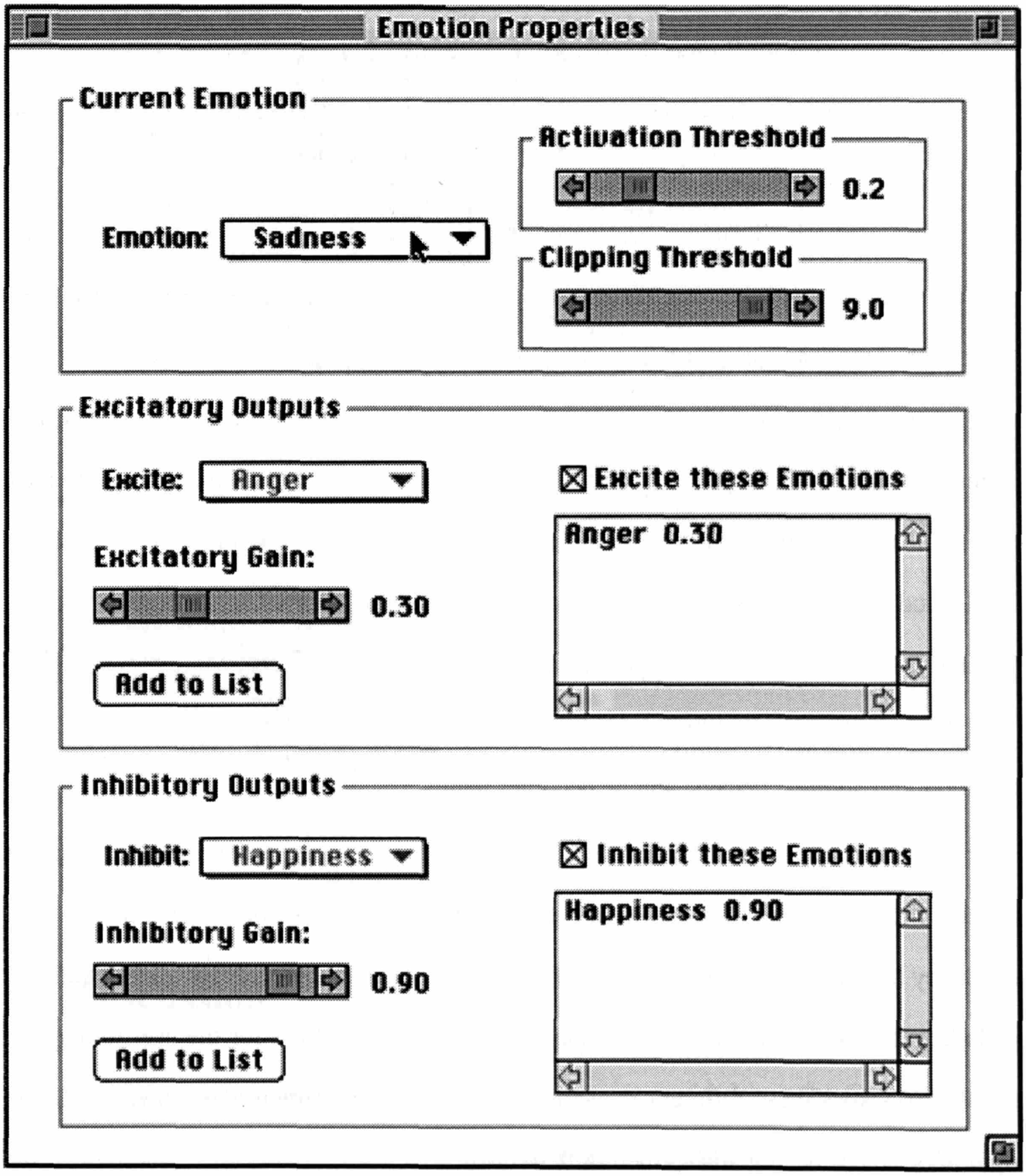

Figure 5-3 Controlling Emotion Proto-specialists Components

5.1.4 Interacting with Simon

The user interacts with Simon by performing actions in the environment and by giving objects to him. Both actions and objects are represented with icons which the user can control using the mouse and the keyboard. For instance, the user can move objects around the environment, give or take away the child's toys, turn on and off the light in Simon's room, turn the fan (or the heater) on and off to change the room's temperature, and perform several other actions, thus, providing external stimuli to Simon which, along with the internal stimuli generated by motivational systems will cause him to react emotionally. For example, when his desire to play is high, and the user makes toys available for him to play with, Simon becomes happy. When the lights of the environment are turned off or a (simulated) loud sound is played, he becomes afraid and in fact may remain fearful for a while. Similarly, when the level of *Hunger* is high and there is no food around, his level of *Distress* increases and eventually becomes sad and angry. Once this need is satisfied, his level of *Anger* and *Sadness* decays, and he becomes happy.

Feedback to the user indicating Simon's emotional state and active behavior is given through different views. A window with a 2-D cartoon-like image of Simon's face is used to convey his current emotional state. Additionally, the intensity levels for all motivations, including drives and emotions are outputted on separate windows for the user to inspect. Figure 5-4 shows these different views in a situation where Simon is displaying *Disgust* right after he has been given spoiled milk to drink.

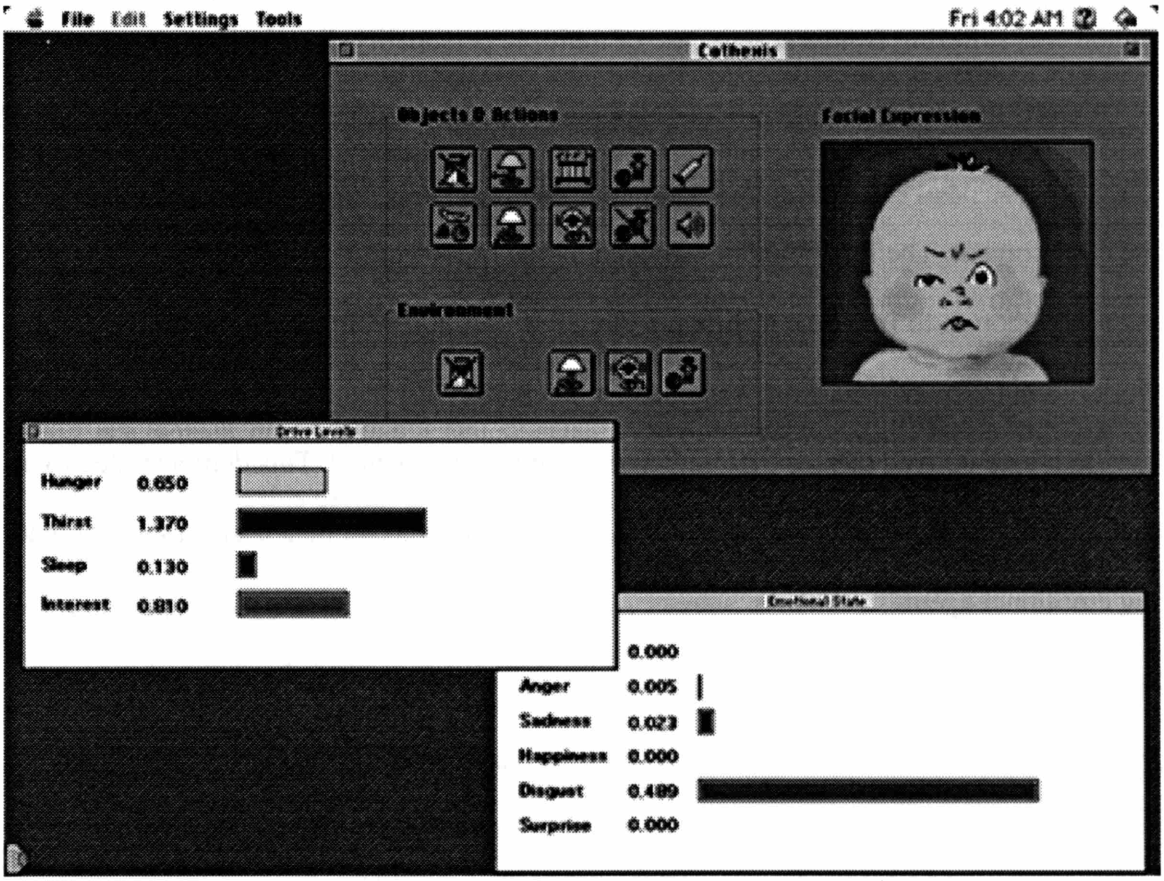

Figure 5-4 An expression of Disgust

5.2 Discussion

We created the environment described above so that we could debug the Cathexis model and experiment with many of the concepts involved in it. Interactions and experiments with Simon have shown that the different systems in Cathexis indeed work as they were described in Chapter 3. Additionally, by manipulating the values of the different components of emotion proto-specialists we were able to model different kinds of emotional phenomena and create situations in which, depending on these values, Simon would react in completely different ways. From an observer's point of view, it would have appeared that Simon was suffering from a severe case of multiple personality disorder!

A word should be said also about the implementation. It turned out that creating Simon was easier than expected. In fact, most of the time was spent writing the user interface for the environment. The functionality provided by the framework described in Chapter 4 was certainly enough to implement Simon's model of emotions. In the cases were the framework did not provide what was wanted, as it happened with specific behaviors, it was very easy to create derived classes that implemented what was needed. This demonstrated to us that not only as a model, but also as a framework, Cathexis can be a useful tool for creating emotional agents.

Chapter 6

Conclusions

6.1 Future Work

There are several improvements and extensions that can be made to the computational model presented in this thesis. We describe some of them in this section.

6.1.1 Schema-Activated Emotion

In the current implementation of the model, cognitive elicitors of emotion include only those that are based on appraisals and attributions. However, as we have mentioned before, some other kinds of cognition elicit emotions as well. Memories of past experiences, for instance, are a significant cause of emotions. It would be very interesting to extend the model in such a way that it could create associations between situations and the emotions elicited in those situations. These associations could be used later in the model to activate specific emotions whenever similar situations occur. A model of schema-activated emotions, such as this one, would resemble the fast and highly adaptive emotion activation system that follows a thalamoamygdaloid pathway as discovered by LeDoux [LeDoux 1993].

6.1.2 Exploitation of Emotion Proto-specialists

An interesting relationship to explore between cognition and emotions would be that of *exploitation*. This concept refers to how an agent (or agency) makes use of the activity of another agent without understanding how it works [Minsky 1986]. Consider the following scenario from "everyday" life:

A graduate student is working late trying to meet the deadline to turn his thesis in. Although this student needs as much time as he can get to finish his work, he has not been able to wake up early in the mornings, partly because he has been working late the night before. Wanting to get the "extra" time in the mornings, he tries several tricks, such as putting the alarm clock far away from the bed so that when it rings he has to stand up hoping that in the process he will wake up. Unfortunately, none of these tricks work. Days have passed and the deadline has become very real. Unexpectedly, very early that morning, he starts waking up and the very first thought that comes to his mind is "I only have n days left! (n being a ridiculously small number)" and this though is accompanied by a strong feeling of fear. Before he is able to understand what is happening, he is already taking a shower and getting ready to work!

In this scenario, according to Minsky's society of mind, the *Wake-up* agent managed to get its job done by *exploiting* the *fear* agent. In Cathexis, exploitations like this exist at the motivational level. Recall that in 3.1.4 we described how drive proto-specialists elicit (exploit) emotion proto-specialists to get their jobs done. Also, the mechanism of inhibition and excitation described in 3.1.5 allows for the mutual exploitation of emotion proto-specialists. However, modeling complex exploitations between emotion proto-specialists and agents such as *Wake-up* has not been considered and would certainly be very interesting work to do.

6.1.3 Influences of Emotion

Another area of future work includes extending the experiential component of Emotional Behaviors to consider other influences of emotions. For instance, it would be very interesting to extend this research and study how emotions influence the agent's perception of the world. In relation to the model of schema-activated emotions suggested above, Cathexis could be extended to consider some of the influences of emotions in memory and learning, such as mood-consistent learning, and mood-dependent retrieval, as described in [Bower and Cohen 1982].

Also, as discussed in Chapter 1, emotions might be indispensable in order to make rational decisions. Therefore, providing agents with emotion-like mechanisms might be necessary if we want them to be effective at decision-making and perhaps truly intelligent. Studying how particular emotions would influence these kinds of decision-making processes would be a very interesting area for future work.

6.2 Applications

As a computational model Cathexis has not been specifically designed to meet the needs of any particular application. However, its design was meant to be as open and flexible as possible so that it could be used in several different application domains. This section explores some of these possible applications.

6.2.1 Entertainment

The area of entertainment is certainly one in which the model presented here would have some obvious applications. Many kinds of entertainment, such as simulations, video games, virtual reality, interactive drama, animations, theater, and so on, employ synthetic characters that act and exist in some environment [Maes 1995]. In Chapter 1 we discussed how emotions were a fundamental characteristic of interactive, believable agents. Cathexis can be used to create synthetic agents that display emotional behaviors which could be then used in any of these entertainment applications.

6.2.2 Education

Another area which could benefit from the use of a model like Cathexis is education. As with some applications in entertainment, simulations and training applications may use synthetic agents that interact with students, this time with educational purposes. The model could also be used in a related but different way. Instead of using it to model emotional behavior (as it is the case with synthetic agents), it could be used to *reason* about emotions. This could indeed prove to be very useful as a user-modeling tool. Consider for instance the area of Intelligent Tutoring Systems (ITSs). An ITS could incorporate Cathexis within its architecture and use it as a tool to model the student. Thus, reasoning about the different sorts of things (situations, topics, type of examples, etc.) that elicit certain emotions on a student, could help an ITS decide what are the appropriate tutoring strategies for that student.

6.2.3 Autonomous Agents

In the area of software agents, Cathexis could be used to provide better and more effective user interfaces. For instance, the model could be used to create interface agents that understand about emotions and that would use this knowledge to improve the communication with the user. Also, the task of assisting the user in some specific domains (e.g. shopping for music on the web) could certainly be improved if the agent is able to create a comprehensive and robust model of the user's likes and dislikes. In my opinion obtaining knowledge about the user's emotions and affective preferences is of crucial importance in achieving this. Therefore, using Cathexis in some of these applications may be appropriate and useful.

6.2.4 Emotion Theory

One of the secondary goals of this work was to obtain a better understanding of human emotions and their relationship with cognition. This computational model can be used to test and experiment with some of the different models of emotions proposed by theorists in the field.

6.3 Final Discussion

We have developed a computational model for the generation of emotions and their influence in the behavior of autonomous agents. This model, by drawing on ideas from several different fields offers an alternative approach to model different types of emotions, such as basic emotions, emotion blends, and mixed emotions. The model considers the dynamic nature of emotions, and, in contrast to other models proposed to date, it considers both cognitive and noncognitive elicitors of emotions, and it differentiates emotions from other affective phenomena, such as moods. Finally, it provides a flexible way of modeling the influence of emotion on the behavior of the agent, which takes into account aspects of both the Expressive and Experiential components of emotion.

The model has been implemented as part of a flexible and extensible object-oriented framework that provides enough functionality for developers to create models of emotional phenomena and emotional agents.

Bibliography

[Abelson 1963] Abelson, R.P. Computer Simulation of "hot cognition". In: Tomkins, S.S., and Messick, S. eds. *Computer Simulation of Personality.* New York: Wiley, 1963.

[Bates 1994] Bates, J. The Role of Emotion in Believable Agents. *Communications of the ACM* 37(7):122-125, 1994.

[Bates, Loyall, and Reilly 1992] Bates, J., Loyall B., and Reilly, S. An Architecture for Action, Emotion, and Social Behavior, Technical Report, CMU-CS-92-144, School of Computer Science, Carnegie Mellon University, 1992.

[Blumberg and Galyean 1995] Blumberg, B., and Galyean, T. Multi-Level Direction of Autonomous Creatures for Real-Time Virtual Environments. In Proceedings of *SIGGRAPH 95*, Los Angeles CA, 1995.

[Blumberg 1994] Blumberg, B. Action-Selection in Hamsterdam: Lessons from Ethology. In Proceedings of the *Third International Conference on the Simulation of Adaptive Behavior*, 108-117. Brighton, England: MIT Press, 1994.

[Bower and Cohen 1982] Bower, G.H., and Cohen, P.R. Emotional Influences in Memory and Thinking: Data and Theory. In: Clark, M.S. and Fiske, S.T. eds. *Affect and Cognition: The 17th Annual Carnegie Symposium on Cognition*, 291-331. Hillsdale, NJ: Erlbaum, 1982.

[Colby 1975] Colby, K.M. *Artificial Paranoia.* New York: Pergamon, 1975.

[Cytowic 1993] Cytowic, R. *The Man Who Tasted Shapes.* New York, NY: Putnam, 1993.

[Damasio 1994] Damasio, A. *Descartes' Error: Emotion, Reason, and the Human Brain.* New York: Gosset/Putnam, 1994.

[Darwin 1965] Darwin, C. The Expression of Emotions in Man and Animals. Chicago: University of Chicago Press, 1965 (Original work published in 1872).

[Davidson 1994] Davidson, R. On Emotion, Mood, and Related Affective Constructs. Ekman, P., and Davidson, R. J. eds. *The Nature of Emotion: Fundamental Questions.* New York: Oxford University Press, 1994.

[Duclos et al 1989] Duclos, S. et al. Emotion-specific Effects of Facial Expressions and Postures of Emotional Experience. *Journal of Personality and Social Psychology, 57,* 100-108, 1989.

[Dyer 1987] Dyer, M. Emotions and Their Computations: Three Computer Models. In: *Cognition and Emotion*, 1(3):323-347, 1987.

[Dyer 1983] Dyer, M. *In-Depth Understanding*. Cambridge, MA: MIT Press, 1983.

[Ekman 1992] Ekman, P. An Argument for Basic Emotions. In: Stein, N. L., and Oatley, K. eds. *Basic Emotions*, 169-200. Hove, UK: Lawrence Erlbaum, 1992.

[Ekman 1994] Ekman, P. Moods, Emotions, and Traits. In: Ekman, P., and Davidson, R. J. eds. *The Nature of Emotion: Fundamental Questions*. New York: Oxford University Press, 1994.

[Ekman and Davidson 1994] Ekman, P., and Davidson, R. J. eds. *The Nature of Emotion: Fundamental Questions*. New York: Oxford University Press, 1994.

[Elliot 1992] Elliot, C.D. The Affective Reasoner: A Process Model of Emotions in a Multi-Agent System. Ph.D. Thesis, Institute for the Learning Sciences, Northwestern University, 1992.

[Elliot 1994] Elliot, C.D. Multi-Media Communication with Emotion-driven "Believable Agents". In Proceedings of the *AAAI Spring Symposium on Believable Agents*, 1994.

[Elliot 1995] Elliot, C.D. Getting to the point: Emotion as a necessary and sufficient element of story construction. In Proceedings of the *AAAI Spring Symposium on Interactive Story Systems*, 1995.

[Forgas and Bower 1987] Forgas, J., and Bower, G. Mood Effects on Person-Perception Judgements. *Journal of Personality and Social Psychology 53(1):* 53-60, 1987.

[Frijda 1996] Frijda, N. Passions: Emotion and Socially Consequential Behavior. In: Kavanaugh, R., Zimmerbag, B., and Fein, S. *Emotion: Interdisciplinary Perspectives.* New Jersey: Lawrence Erlbaum, 1996

[Frijda 1986] Frijda, N. *The Emotions*. Cambridge, UK: Cambridge University Press, 1986.

[Gelernter 1994] Gelernter, D. *The Muse in the Machine: Computers and Creative Thought.* New York, NY: The Free Press, 1994.

[Goleman 1995] Goleman, D. *Emotional Intelligence.* New York: Bantam, 1995.

[Hayes-Roth, Brownston, and Sincoff 1995] Hayes-Roth, B., Brownston, L., and Sincoff, E. Directed Improvisation by Computer Characters. Technical Report: KSL-95-04. Knowledge Systems Laboratory, Stanford University, 1995.

[Izard 1977] Izard, C. *Human Emotions.* New York: Plenum Press, 1977.

[Izard 1991] Izard, C. *The Psychology of Emotions.* New York: Plenum Press, 1991.

[Izard 1993] Izard, C. E. Four Systems for Emotion Activation: Cognitive and Noncognitive Processes. *Psychological Review* 100(1):68-90, 1993.

[Johnson-Laird and Oatley 1992] Johnson-Laird, P. and Oatley, K. Basic Emotions, Rationality, and Folk Theory. In: Stein, N. L., and Oatley, K. eds. *Basic Emotions,* 169-200. Hove, UK: Lawrence Erlbaum, 1992.

[Johnson-Laird and Oatley 1989] Johnson-Laird, P. and Oatley, K. The Language of Emotions: An Analysis of a Semantic field. *Cognition and Emotion, 3,* 81-123, 1989.

[LeDoux 1993] LeDoux, J. Emotional Memory Systems in the Brain. *Behavioral and Brain Research, 58,* 1993.

[Lewis and Haviland 1993] Lewis, M. and Haviland, J. eds. *Handbook of Emotions.* New York: Guilford Press, 1993.

[Loyall and Bates 1993] Loyall, A.B., and Bates, J. Real-time Control of Animated Broad Agents. In: *Proceedings of the Fifteen Annual Conference of the Cognitive Science Society.* Boulder, 1993.

[Ludlow 1980] Ludlow, A. The Evolution and Simulation of a Decision Maker. In: Toates, F., and Halliday, T. *Analysis of Motivational Processes.* London: Academic Press, 1980.

[Maes 1995] Maes P. Artificial Life meets Entertainment: Lifelike Autonomous Agents. *Communications of the ACM. Special Issue on Novel Applications of AI,* 1995.

[Maes et al. 1995] Maes, P., et al. The ALIVE System: Full-body Interaction with Autonomous Agents. In: *Proceedings of Computer Animation '95 Conference.* Geneva, Switzerland, 1995.

[Maes 1993] Maes P. Behavior-Based Artificial Intelligence. In Proceedings of *the second Conference on Adaptive Behavior.* MIT Press, 1993.

[Maes 1991] Maes P. ed. *Designing Autonomous Agents: Theory and Practice from Biology to Engineering and Back.* MIT Press, 1991.

[Maes 1989] Maes P. How To Do the Right Thing. AI Memo No. 1180, MIT Artificial Intelligence Laboratory, 1989.

[Minsky and Papert 1969] Minsky, M., and Papert, S. *Perceptrons.* Cambridge: MIT Press, 1969.

[Minsky 1986] Minsky. M. *The Society of Mind.* New York: Simon & Schuster, 1986.

[Nass, Steuer, and Tauber 1994] Nass, C.I., Steuer, J.S., and Tauber, E. Computers are Social Actors. In: *Proceedings of the CHI'94 Conference,*72-78. Boston, 1994.

[Niedenthal and Kitayama 1994] Niedenthal, P. and Kitayama, S. *The Heart's Eye: Emotional Influences in Perception and Attention.* San Diego: Academic Press, 1994.

[Oatley and Johnson-Laird 1987] Oatley, K., and Johnson-Laird. P.N. Towards a Cognitive Theory of Emotions. *Cognition and Emotion*, 1:29-50, 1987.

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

[Panksepp 1994] Panksepp, J. Basic Emotions Ramify Widely in the Brain, Yielding Many Concepts That Cannot Be Distinguished Unambiguously... Yet. Ekman, P., and Davidson, R. J. eds. *The Nature of Emotion: Fundamental Questions.* New York: Oxford University Press, 1994.

[Pfeifer 1988] Pfeifer, R. Artificial Intelligence Models of Emotion. In: Hamilton, V., Bower, G. H.. and Frijda, N. eds. *Cognitive Perspectives on Emotion and Motivation*, 287-320. Netherlands: Kluwer, 1988.

[Pfeifer 1995] Pfeifer, R. The "Fungus Eater Approach" to Emotion: A View from Artificial Intelligence. Japanese version in: Bulletin of the Japanese Cognitive Science Society 1(2):42-57, 1995.

[Picard 1995] Picard, R. W. Affective Computing, Technical Report No. 321, MIT Media Laboratory, 1995.

[Plutchik 1994] Plutchik, R. *The Psychology and Biology of Emotion*. New York: HarperCollins, 1994.

[Reilly 1996] Reilly S. Believable Social and Emotional Agents. Technical Report, CMU-CS-96-138, School of Computer Science, Carnegie Mellon University, 1996.

[Reilly and Bates 1992] Reilly, S., and Bates, J. Building Emotional Agents, Technical Report, CMU-CS-92-143, School of Computer Science, Carnegie Mellon University, 1992.

[Roseman, Spindel, and Jose 1990] Roseman, I. J., Jose, P. E., and Spindel, M. S. Appraisals of Emotion-Eliciting Events: Testing a Theory of Discrete Emotions. *Journal of Personality and Social Psychology* 59(5):899-915, 1990.

[Salovey and Mayer 1990] Salovey, P, and Mayer, J. Emotional Intelligence. *Imagination, Cognition, and Personality*, 9, 185-211.

[Shaver et al. 1987] Shaver et al. Emotion Knowledge: Further Exploration of a Prototype Approach. *Journal of Personality and Social Psychology, 52,* 1061-1086, 1987.

[Scherer 1988] Scherer, K. Criteria for Emotion-antecedent Appraisal: A Review. In: Hamilton, V., Bower, G., and Frijda, N. eds. Cognitive Perspectives on Emotion and Motivation. The Netherlands: Kluwer, 1988.

[Sloman 1991] Sloman, A. Prolegomena to a Theory of Communication and Affect. In: Ortony, A., Slack, J., and Stock, O. eds. *A.I. and Cognitive Science Perspectives on Communication*. Springer, Heidelberg, 1991.

[Stein, and Oatley 1992] Stein, N. L., and Oatley, K. eds. *Basic Emotions*, Hove, UK: Lawrence Erlbaum, 1992.

[Swagerman 1987] Swagerman, J. The Artificial Concern Realization System ACRES. A Computer Model of Emotion. Ph.D. Thesis, Department of Psychology, University of Amsterdam, 1987.

[Tinbergen 1969] Tinbergen, N. *The Study of Instinct.* London: Oxford, 1969.

[Toda 1993] Toda, M. Emotions and Urges. Technical Report, No. 93-1-01, School of Computer and Cognitive Sciences, Chukyo University, 1993.

[Toda 1994] Toda, M. Basic Structure of the Urge Operations. Technical Report, No. 94-1-01, School of Computer and Cognitive Sciences, Chukyo University, 1994.

[Toda 1995] Toda, M. A Decision Theoretical Model of Urge Operations. Technical Report, No. 95-1-01, School of Computer and Cognitive Sciences, Chukyo University, 1995.

[Toda 1996] Toda, M. The Problems of Altruism and Evolution Systems. Technical Report, No. 95-1-04, School of Computer and Cognitive Sciences, Chukyo University, 1996.

[Tomkins 1962] Tomkins, S. *Affect, Imagery, consciousness: Vol.1 The Positive Affects.* New York: Springer, 1962.

[Tomkins 1963] Tomkins, S. *Affect, Imagery, consciousness: Vol.2 The Negative Affects.* New York: Springer, 1963.

[Weiner 1986] Weiner, B. *An Attributional Theory of Motivation and Emotion.* New York: Springer, 1986.

[Wierzbicka 1992] Wierzbicka, A. Talking about Emotions: Semantics, Culture, and Cognition. In: Stein, N. L., and Oatley, K. eds. *Basic Emotions*, 285-319. Hove, UK: Lawrence Erlbaum, 1992.

[Wierzbicka 1992] Wierzbicka, A. Defining Emotion Concepts. *Cognitive Science 16*, 539-581, 1992.