THE UNIVERSITY OF SOUTHERN QUEENSLAND

THE SYNTHESIS OF EMOTIONS IN ARTIFICIAL INTELLIGENCES:

An Affective Agent Architecture for Intuitive Reasoning in Artificial Intelligences

A Dissertation submitted by
Penny Baillie, B.Info.Tech. (USQ)

For the award of
Doctor of Philosophy
2002

Abstract

This dissertation addresses several highly-critical issues in affective computing and agent architecture design including knowledge representation, motivation, emotion appraisal and affective decision making. The approach presented integrates motivational drives, goals and associated behaviours via a multi-dimensional *Affective Space.*

The research focuses on an emotionally motivated artificial intelligence (EMAI) architecture. This architecture dispenses with the ideas implemented in contemporary affective agent architectures where individual emotional states are modelled as individual variables, integrated and processed using complex algorithms. Contemporary approaches required significant programming effort to modify them for domains outside their realm, integration of new emotional states and high-level complex affective decision making. Unlike contemporary affective agent architectures, the EMAI architecture reasons using a multi-dimensional decision making process where emotional states are modelled as coexisting locations in a six-dimensional affective continuum called the Affective Space. Through use of the Affective Space, an EMAI agent can predict the effect that certain behaviours will have on its emotional state and in turn decide how to behave. Furthermore, the agent can use the emotions produced from its behaviour to update its beliefs about particular events and entities. The nature of the Affective Space also allows an EMAI agent to deal with processes related to emotion synthesis in a more effective manner than contemporary architectures. These processes include the natural diminishing of an emotional state's strength over time, the way in which emotions can influence an agent's perspective of a situation and the way in which an agent can migrate from one emotional state to another.

This dissertation contributes crucial and unique concepts and formalisations of emotion based intelligence for agent construction to the domain of Artificial Intelligence (in particular Affective Computing). It introduces a unique process for emotionally motivated decision making based on holistic and atomic appraisals made with respect to events. The thesis contained within has been supported through experimentation that has confirmed the effectiveness of the emotion synthesis technique in the EMAI architecture and how this is used to produce intelligent agents capable of emotional reasoning and decision making.

Certification of Dissertation

I certify the ideas, experimental work, results, analyses, software and conclusions reported in this dissertation are entirely my own effort, except where otherwise acknowledged. I also certify the work is original and has not been previously submitted for any other award, except where otherwise acknowledged.

____________________________ ____________________

Signature of Candidate Date

ENDORSEMENT

____________________________ ____________________

Signature of Supervisor/s Date

Acknowledgments

I would like to take this opportunity to thank those people who have influenced me the most during the past six years of my Ph.D. studies.

Foremost, I would like to thank Dr Dickson Lukose who has been there from the very beginning. At different times during the course of my studies, I had five supervisors. Many of them moved to other Universities making it difficult for them to continue their associations with my research. Although Dr. Lukose moved overseas and was no longer my principal supervisor he has continued to be an inspiration and continual mentor over the entire duration of my studies, in the capacity of Associate Supervisor. The best advice that he gave me was on day one when he said, “Start writing now. Don’t wait until the end!” Because of this, most of the work presented herein has already been published in numerous international conference proceedings, a journal and a book. I *hope* that I can pass this advice on someday. *Dickson, I do not know how I can show my gratitude.*

Secondly, I would like to thank my current principal supervisor, Dr. Mark Toleman. As I had already completed a significant amount of research when I met up with Mark, I believe that he was not initially convinced that artificial intelligences should exhibit emotions. Now, I believe he is a true advocate for the affective computing domain. He has provided me with invaluable advise and support over the past two years. I know that our numerous debates and discussions about this research have helped me to develop a deeper understanding of the topic. *Mark, my sincere thanks for your encouragement. I couldn’t have finished this without you.*

Next, I would like to thank my Mum, Dad and my brothers Tim and Craig. I know they have believed in me from the very beginning. *Thank-you so much for being there through the ups and downs of the past six years.*

Special thanks must go to Daniel and Deefa who have never known me to be without Ph.D. study. Without them, I do not believe I would have had the motivation to make it this far. I cannot express in words the emotional support and strength that I acquire from you. *LYVM*

And finally, to the late Gene Roddenberry, thank-you for the inspiration!

Table of Contents

List of Figures

List of Tables

1. Introduction

We are all in the gutter, but some of us are looking at the stars.
OSCAR WILDE.

1.1 The Ultimate Puzzle

> *How is it possible for a slow, tiny brain, whether biological or electronic, to perceive, understand, predict, and manipulate a world far larger and more complicated than itself?* (Russell and Norvig 1995)

Understanding the functions of the brain, according to Russell and Norvig (1995), is the ultimate puzzle the domain of Artificial Intelligence (AI) is attempting to solve. In an endeavour to develop intelligent systems that mimic the thinking and behaviour of *rational* humans, AI researchers have called on a number of disciplines including psychology, physiology, neurology and philosophy. However, since 1956 when the AI discipline's name was first coined, researchers have not made a considerable advancement toward a theory of general intelligence (Stork 1997). In Stork's interview with Marvin Minsky, Minsky conveyed his belief that too much time and effort is put into small projects that do not contribute to the progression of a model for general intelligence. Minsky concludes that AI researchers have but scratched at the surface of the fundamentals of true AI and he believes one of these fundamentals is emotion (Stork 1997).

When computers are viewed as logically intelligent machines, what would be the point in clouding their rational abilities with unpredictable and confusing qualities such as emotions? This may be a common view amongst AI purists and much of this opinion could be due to the way in which the term is perceived and defined (Baillie et al. 2000a). Traditionally, philosophers have defined emotions to be interruptions to otherwise logical states of being (Smith and Kirby 2000). The recent resurgence of research in the emotional realm in both psychology and cognitive science agrees with the view developed by the late Charles Darwin who, in the late 1800s, conceived that emotions play an important part in our cognition and serve to provide us with the mechanisms for adaptive behaviour in a dynamically complex world (Smith and Ellsworth 1985).

From a neurological standpoint, emotions act as a kind of filtering system for information processing within the brain (Pert 1997). The neuropeptide receptors, responsible for emotional states and original thought, were initially believed only to exist in the amygdala, hippocampus and hypothalymus (the traditional seats of emotion). Recent research has revealed that these receptors exist in high concentrations throughout the body including the spinal cord, where all bodily sensations and feelings are initially processed. Therefore, all sensory information passing between synapses undergoes an *emotional* filtering process. This operation assists the brain's ability to deal with the deluge of sensory input that it receives. Neurology indicates that emotion plays a role in the way humans process information and therefore suggests that emotion affects human intelligence. Furthermore, it can be argued that intelligence goes beyond logic and synthesising emotional responses in machines could greatly enhance the field of AI.

Although emotion is an integral part of the human psyche, traditional AI researchers have all but ignored the role that emotions play in intelligence. Understanding emotion generation and developing a reasonable computational model may hold the key to unlocking the potential of *affective* (emotion based) reasoning and decision making in machines.

One such line of research is the realm of Affective Computing. This relatively new domain examines the effect that emotions have on human intelligence and endeavours to use this to further enhance the field of AI. How the concept of emotions might heighten the intelligent functioning of artificial beings is still unclear, but through the variety of research programs that currently exist, areas that might benefit are being identified. In this domain much work is being done to develop artificial intelligences capable of identifying, processing and synthesizing emotions. Picard (1997) suggests that emotions are an integral part and a natural progression for artificial intelligence. She further states that: '... the inability of today's computers to recognize, express, and have emotions, severely limits their ability to act intelligently and interact naturally with us'.

A domain where this is becoming evident is the realm of interactive fiction where autonomous characters, generated by a computer are expected to act and interact in a believable and emotional way (Stern 1999). The fictional characters that inspire these computerised counterparts began with the intelligent machines of Isaac Asimov. These initially became the primary element in futuristic fictional environments. Before many others, writers realised that to create convincing storylines and allow human and machine interaction, emotions had to be an integral part of these electronic and mechanical intelligences. Examples of such automatons are Arthur C Clark's *HAL* (Clarke and Kubrick, 1993), *R2-D2* and *C3-PO* (Sansweet and Zahn 1998), Terry Nation's *O.R.A.C.* (Wells and Nazarro, 1997), Star Trek's *Commander Data* (Okuda and Okuda 1997), and

Asimov's robot, *R. Daneel Ovilaw* (Asimov 1950). Developing this type of believability in fictional characters impacts on the domain of Affective Computing in areas such as interactive fiction and computer games. Researchers are aware of the need to produce computerised characters that are believable (Reilly 1996) in order to suspend the user's disbelief in the same way attempted in fictional books and motion pictures. However, developers of computerised characters, in particular autonomous ones, need to understand how these characters can be programmed with the ability to construct and process synthetic emotions as well as convey them.

For years, scientists have debated over a general definition of emotion. However, one thing is agreed, emotions affect human behaviour. The common stigma associated with emotions is that they essentially cause irrational behaviour. To act in an 'emotional' way implies a fallacious, senseless and illogical demeanor. In fact, throughout this dissertation, there are terms such as *emotional decision-making* and *emotional behaviour*. Out of context these premises might lead the reader to assume them to mean *irrational decision-making* and *irrational behaviour*. However, this is not the case.

In a recent issue of *The Courier Mail*[1] dated the 11th of August 2001, an eye-catching advertisement was placed in the business section. This advertisement had the large banner heading of *EMOTIONAL INTELLIGENCE.* On further investigation it was discovered that it was an advertisement for a seminar on how to be emotionally intelligent. In the context of a newspaper section read by people *interest*ed in big business it could hardly be concluded that the term *emotional* in this advertisement could mean irrational. It was however, referring to the latest buzzwords in the business world. In 1995, Daniel Goleman coined the term *emotional intelligence* (Goleman 1995). He suggests that to get ahead in business it not only takes intelligence but also a considerable amount of knowing how to control and use the natural biological trait of emotion to one's advantage. Goleman (1995) defines emotional intelligence as: '…the capacity to form an accurate, veridical model of oneself and to be able to use that model to operate effectively in life'.

---

[1] A leading Australian current affairs newspaper produced daily in Brisbane, Queensland, Australia.

1.2 A Need for Emotional Computers

There has been much literature and research in recent years on the theories and generation of emotions in artificial intelligences. Many psychological theorists believe that emotions represent adaptive responses in different situations (Plutchik 1980, Smith and Ellsworth 1985). Many AI researchers (Picard 1997, Sloman 2001 and Minsky 2000) *hope* these mechanisms can be integrated into intelligent artificial devices for this same reason.

Hanley (1997) explains the need for emotions to act as motivational mechanisms to aid in decision making. In order to perform even the simplest task, a cognitive being (human or artificial) must decide between a number of options. Logical selection of a choice, at a minimum, requires weighing up the probable outcomes of the range of choices. The number of choices could be anywhere from one to infinity. Interpolating the choices that arise as outcomes of particular events and the further consequences of selecting from among these choices creates an explosion of options. In addition to tracking all possibilities, the being must also draw on its knowledge on how to complete a task that will satisfy these options while eliminating the irrelevancies. Hanley concludes that, a purely logical being (devoid of emotion) when faced with such a plethora of decisions would become overwhelmed with the immense amount of possibilities from which to select.

Emotional decision making provides a good solution for computer systems that face the problem of enumerating and evaluating multitudinous choices within an acceptable time frame. One application of AI that is benefiting by integrating emotional decision making mechanisms is that of intelligent agents. The word *agent* is used within the AI domain to refer to a number of different applications. The most popular use of the term pertains to an autonomous artificial being that has the ability to interact intelligently within a temporally dynamic environment.

Much of the current research for achieving these types of results with artificial agents is based in appraisal theories. Models such as the Affective Reasoner (Elliot 1992), the Oz project (Bates et al. 1992), PETEII (El-Nasr 1998), Silas (Blumberg et al. 1996) and Yuppy (Velasquez 1999) are just a few examples of how appraisals (cognitive and non-cognitive) are used to generate emotional states and behaviours within an artificial intelligence.

The variety of emotional concepts implemented in these models using these appraisal theories is quite expansive. No truly correct model has been identified (Velasquez 1999) and much debate continues on which models or theories are *best* for integration into AI architectures.

1.3 Perceived Shortfalls of Contemporary Research in Affective Computing

Synthesising and modelling emotions is a complex task. Most affective computing research efforts concentrate on particular aspects of generating emotions. There is no holistic view and with the limited understanding of how emotions are produced and the affect they have on humans, it would be arrogant to assume that one could be actualised so early in the study of this field. Much work has been done on small research projects that study particular aspects of the relationship between emotions and intelligence.

1. *Current affective computing architectures model emotions as individual variables.* The problem inherent with this method is that complex algorithms need to be programmed into the models to manage the relationships that exist between differing emotional states.

2. *In humans, the experiencing of emotional states provides feedback that influences future behaviour.* This mechanism is not used in current models to the extent that it allows the artificial being to adapt to its environment.

3. *Contemporary Affective Computing models are limited due to their lack of deep knowledge representation and knowledge manipulation capabilities.* Contemporary models can associate emotion with specific beliefs in the knowledge base, however, this restricts the way in which the agents can understand and interact with its environment. Emotional responses need to be associated with beliefs throughout the knowledge base.

4. *In contemporary research, the emotion associated with an event is evaluated holistically.* Contemporary affective agent architectures provide no means of attributing partial blame or credit of resulting emotions to individual elements that were part of the event.

5. *The natural decrease in emotion strength or emotional decay in previous models has been dealt with by using constant decay rate functions that act over time and require considerable programming and coordination effort.* This is a shortfall associated with the representation of emotional states as individual variables.

6. *Current models do not address the issue of dynamic goal interpretation (via dynamic planning) allowing the agents to change plans based on emotional experience.* Emotional responses should allow an affective agent to reevaluate its goals and plans through belief revision.

7. *Current models lack the ability to maintain a temporally dynamic emotional experience with an environmental element be it an object or another agent.* This means that contemporary agents are unable to associate positive emotions with an element in one instance and in another, through experience, change its emotional perspective about the element to associate a negative emotion with it.

8. *Contemporary agent models do not apply emotional bias in decision making.* While they do provide the means to generate agents with personalities and preprogrammed behaviours associated with differing dispositions, they do not allow for temporal changes in decision making using an emotional bias.

9. *As yet, no computational models of emotional intelligence have been described to formalise the emergence of emotion in an agent model.*

10. *While emotional states combine in current affective computing models to produce a single emotional state they require cumbersome programming and coordination of the individual emotional state to be achieved.*

The research presented in this dissertation examines how these limitations influence the affective decision making process and how the solutions presented within contribute significantly to the domains of affective computing and AI. These contributions will be explained and demonstrated by examining the new concepts and mechanisms introduced. Furthermore, this dissertation presents the current work taking place in the domain of affective computing and an improved affective agent architecture that integrates a new approach to emotion synthesis and affective decision making.

1.4 Structure of the Dissertation

By considering all the contemporary ideas and models of emotion, this dissertation presents an examination of the elements that make up an emotional event. It explores how the assessment of these elements, when combined, can be used to generate emotional states and responses in temporal situations. To this end, considerable time has been dedicated to the exploration of the psychological theories of emotion to develop a computational model for emotion generation of which emotional experiences are to be a key element.

The research in this dissertation proposes the use of physiological and psychological theories in the development of a new model for emotion synthesis. This model not only offers an artificial intelligence capable of generating behaviours from emotional states, but an artificial intelligence with the ability to use its emotional states to revise its beliefs and adapt to its environment. This model, known as the *Emotionally Motivated Artificial Intelligence* (EMAI) architecture, model presents the archetype of an artificial agent capable of reasoning about a selection of behaviours by emotionally assessing its alternative options.

As the research presented in this dissertation is unique to AI, Chapter 2 explores the concept of emotion, how emotions are defined, generated and the effect they have on human decision making and perception. This chapter also examines the neurological foundation for the generation of emotions and their effect on human intelligence. Varying theories of emotion and an empirical method for measuring and predicting emotional states will be presented. This is followed by an overview of the most recent research in the domain of affective computing, which examines the current state of affective computing with respect to intelligent decision making and reasoning. Due to the fledgling status of the domain, this is essential background information.

Chapter 3 outlines the EMAI architecture. It describes the Knowledge Area, the Constructive Area and the Deliberate Area and each of their sub-components. The mechanisms introduced in the EMAI architecture makes it possible for an EMAI agent to produce internal emotional states. These states are used to activate goals, develop plans to satisfy goals, make affective decisions about behaviour, develop attitudes and adapt to the environment.

Chapter 4 discusses and illustrates the EMAI architecture's Ontology. The Ontology defines the knowledge structure within the architecture. The EMAI Ontology comprises a goal structure called the Goal Hierarchy, structures of concepts called the Type and Relation Hierarchies and a set of fundamental conceptual structures that provide the agent with a deep knowledge representation essential for its functioning. Most contemporary affective agent systems use a set of production rules that are programmed into the agent. They associate fixed emotional states and reactions with elements contained in the agent's environment. This does not allow the agent to change the emotions it associates with items in the environment or its actions. This chapter shows the way in which knowledge is represented in the EMAI architecture. It illustrates how the EMAI architecture's deep knowledge representation provides an EMAI agent with the ability to construct complex representations of events, elements within its environment and associated emotions.

Chapter 5 discusses and illustrates the goal setting and motivational mechanisms integrated into the EMAI architecture. It examines in depth the Goal Hierarchy (introduced in Chapter 4) and explains how the agent can generate low level emotions that can lead to the activation of goals. This chapter explains how an EMAI agent can dynamically plan and re-plan through the interpretation of its goals and behaviours derived from motivational emotional states.

Chapter 6 investigates the way in which an EMAI agent can generate a set of competing plans of behaviour from a single activated goal. These plans are used to produce outward behaviour in an EMAI agent that will lead to the satisfaction of a goal. It will also be shown why it is important for an EMAI agent to plan and re-plan through its process of dynamic goal interpretation.

Chapter 7 examines EMAI's ability to form valenced reactions[2] toward elements that make up the agent's environment and use this information to make decisions. It discusses the psychological concept of attitude and presents the mechanisms integrated into the EMAI architecture that allow for development and use of these valenced reactions to perform single-dimension affective decision making. It will also be shown how an agent built with the EMAI architecture can use the valenced reactions that it forms to revise the beliefs about its behaviour and interactions with its environment. It can then use this information in future decision making tasks.

[2] [2] A positive or negative response.

Chapter 8 discusses the agent's mechanism for generating emotional states and how it uses these to prioritise its behaviours. It explains in detail the *Affective Space* and formalises how it can be used to determine an EMAI agent's current emotional state and predict future emotional states. The Affective Space is a key component of the EMAI architecture and allows the agent to apply emotional biasing in decision making, theoretically represent an infinite number of emotional states without restructuring or reprogramming and combine multiple emotional states to produce a single emotional state. It allows emotions to be represented in such a way that enables opposing emotions to naturally counter-react with each other and produce an automatic decay in emotion strengths as an EMAI agent's emotional state changes from one emotion to another. This chapter also provides a formalisation of the emergence of emotions in an emotionally intelligent decision-making agent.

Chapter 9 builds on the single-dimension affective decision making model in Chapter 7 by integrating it with the Affective Space from Chapter 8. The result is a model for multi-dimensional affective decision making. The chapter also explains how this can be used to give an EMAI agent the ability to assign emotions toward individual elements present during an event and re-evaluate its emotional reaction toward these elements. This assignment and reassignment of emotional states toward individual event entities gives the agent the ability to extrapolate future emotional states given the knowledge of the entities that will be present. Furthermore, the chapter explains how emotions are updated, blended and decayed within an EMAI agent.

To conclude, Chapter 10 presents an experiment conducted to evaluate an EMAI agent's emotional state generation and associated behaviours. In this chapter, an EMAI agent is used to produce a computer-generated character. An evaluation of the EMAI architecture will be presented. Assessment of human interaction with the computerised character will be used to evaluate the emotional behaviour of the agent and confirm the architecture's contributions of new concepts and theories to the realm of affective computing.

Finally, Chapter 11 summarises the content of the dissertation, reviews the contributions of the research and discusses further directions for this work.

1.5 Case Studies and Implementation of Prototypes

Throughout this dissertation are case studies, beginning at Chapter 5, demonstrating how the concepts within the EMAI architecture can be implemented and analysed. The prototypes developed and used for the experiments, presented in the case studies, were programmed by the author.

The case studies, presented in Sections 5.3, 7.6 and 7.8, were implemented using Java on a personal computer running Windows 95. Each experiment used the same base source code. The knowledge base was stored in a text file and uploaded each time the program was run. This allowed the knowledge base to be changed in order to run different experiments without recompiling the source code. In the case study of Section 7.7, Borland C++ Builder was used to create a simple graphical user interface to display a cell automata grid.

The Fido simulation system prototype, presented in Chapter 9 and evaluated in Chapter 10, was programmed in Borland C++ Builder on a personal computer running Window 98. The conceptual graph processor, integrated to manage the agent's knowledge base, is based on a version of CGKEE (Munday et al, 1996) ported by the author from UNIX to Windows 98.

All software is available from the author upon request.

2. Artificial Intelligence and Emotions

I have no inclination to keep the domain of the psychological floating as it were in the air, without any organic foundation... Let the biologists go as far as they can and let us go as far as we can. Some day the two will meet.
FREUD

2.1 A Missing Piece of the Puzzle

When emotions are such an integral part of the human psyche, how is it possible that traditionally they have been ignored in AI? It is beyond the scope of this work to philosophise about the historical division of emotion and otherwise indiscriminate faculties of the mind. However, it seems pertinent to mention how emotions and reason (logic) may have ever become separated. As a point of interest, Koestler (1967) traces the separation back to the ubiquitous rituals of human sacrifice where reason and emotions were split into the continuous delusional streak that runs throughout history. It seems logical that the killing of another person would conjure up a number of seriously debilitating emotions that would impede any reasonable excuses to sacrifice a life. Therefore, emotions became the illogical states that interrupt rational thought. At least this has been the viewpoint of philosophers and the majority of *purist* AI researchers (Smith 1997). Although emotions have long been acknowledged as an important part of the human mind, it was unclear how these conditions affected rational thought. The conclusion was that they did affect rational thought; they made it irrational.

The first question that should be asked is, *why do machine minds, that are disembodied from a natural biological system, need to exhibit states that are primarily produced in biological beings through a series of physiological reactions?* The answer resides in the present day research of affective computing. Whether an emotional system is being applied to an artificial being for use in interactional fiction (Jenkins 1998), computer games (Maes 1995), facial expression recognition (Rosenfeld 1997) or affective decision making (Baillie et al. 2000b) the consensus that allures researchers to the field of affective computing is that emotional mechanisms are essential in the intelligence of a human being and therefore, are a relevant issue in artificial intelligence. This chapter examines the psychological and physiological theories of emotion that have been used in describing the human emotional experience, and how theories on emotion have inspired AI researchers to implement models of emotion derived from these theories into artificial beings to enhance their intelligence.

It is difficult to deny the intertwining link between emotions and logic given the recent emergence of research into emotion in the AI domain. The inclusion of models of emotional synthesis into artificial intelligences may be necessary. Although it is the physical attributes of humans that generate emotions, undeniably and unavoidably these cognitive responses influence behaviour, perception, motivation and personality. Given that it is recognised that emotions are an important aspect of rational thought, reasoning and decision making, it needs to be determined how emotions can be synthetically generated. Obviously, contemporary machines lack the necessary neurotransmitters to produce emotions naturally and other discrete and computational methods need to be examined for this purpose.

This chapter begins by defining the concept of emotions and continues by examining the physiological basis of emotion in the human body, how these emotions are interpreted, and how they affect behaviour. The diversified theories and models of emotions are also discussed.

2.2 A Definition of Emotions

Freud's epigraph is a well-put observation of the fields of psychology and biology, not only at the time that he said it but also now. No position could be truer when it comes to defining emotion.

Emotion presents itself as a non-concise term in many of the domains that boast an understanding of the topic. These range from neurology (Fellous 1999, Pert 1997) and psychology (Smith and Kirby 2000) to artificial intelligence (Picard 1997). The reason may be that the term is used to describe a large range of cognitive and physiological states in sentient beings.

Emotions are often referred to in the broad sense, to describe not only familiar feelings such as *happiness* and *sadness*, but also biological motivational urges such as *hunger* and *thirst* (Pert 1997). Koestler (1967) summed up this general view defining emotions as: '...mental states accompanied by intense feelings and involving bodily changes of a widespread character'. The degree of difficulty plaguing this subject matter is that a person rarely experiences a *pure* emotion (Koestler 1967). For example, feelings of *hunger* may be accompanied by feelings of *frustration*. However, there is a logical intuitive difference in defining *hunger* as an emotion and *frustration* as an emotion.

Another general definition of emotion given by Lefton (1994) defines emotion as: '... a subjective response (feeling), usually accompanied by a physiological change, that is interpreted by the individual, then readies the individual for some action that is associated with a change in behaviour'. This proposition certainly recognises the meeting of the minds that Freud predicts and although this does not attempt to explicitly define an emotion, it meshes the logical theory that emotions are indeed physiologically produced in an individual and affect that individual psychologically.

Of course, this later stance is clear in the field of neuropsychology, where emotional behaviour is explained as physiological and treated at this level with chemicals (Kolb and Whishaw 1990). For example, patients with severe depression can be successfully treated with medications containing serotonin (an essential neurotransmitter found in the brain and spinal cord) that replaces essential neuropeptides lacking in the brain. However, much research in generating psychological theories and models ignores the biological counterpart. For example, many reasoning, decision-making, belief, personality and social behaviour models have been developed in complete isolation of the physiological domain.

There exists a plethora of theories and models on emotions. Many of these have been conceived to explore different emotional concepts from different perspectives. Some theories were designed to explain the physiological generation and effects of emotion, some place an emphasis on the lower primal emotions that influence and generate motivational states, while others examine the discrete categorisation of emotional reactions and their antecedents. The following sections provide a synopsis of the work accomplished in this domain and the theories and models developed.

2.2.1 The Physiology of Emotion

While it is ludicrous to contemplate emotions having the same physiological affect on artificial intelligences as they do on the human nervous system, it is crucial to examine the biological generation of emotions and physiological effects before a theory of the influence of emotions on cognition can be formulated.

The search for a physiological theory of emotion has lead many researchers to experiment with crucial areas of the brain thought to produce emotional responses in animals. These include the hypothalamus, the amygdala, and other cortical and subcortical areas (Lefton 1994). From the 1920s onwards, researchers have removed and electrically simulated these portions of the brain in order to examine the behavioural and emotional responses in their subjects. These particular areas of the brain are collectively known as the limbic system and traditionally emotions where thought to emanate from a series of neurons located within.

MacLean (1958), one of the pioneers of physiological-emotion theory, explains human emotions as the result of the interaction of three very differently structured brain types; the reptilian brain, the paleo-mammalian brain and the neo-mammalian brain. His theory is known as the *Three Brain Theory*. In brief, the human brain has inherited the structures of these three brains as the result of the evolutionary process. The reptilian brain corresponds to the basic structure in the phylo-genetically oldest parts of the human brain such as the medulla, hypothalamus, the reticular system and the basal ganglia (see Figure 2.1). This area is said to be responsible for visceral and glandular regulations, primitive activities, reflexes, arousal and sleep.

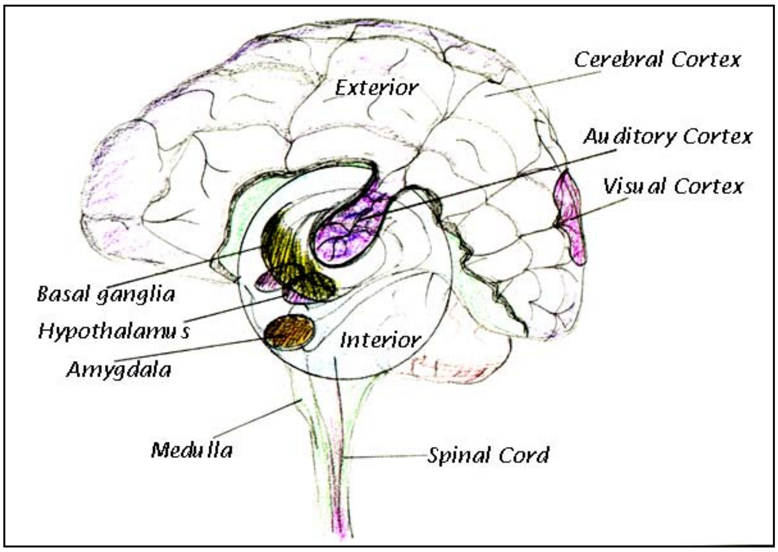


Figure 2.1 Rough Sketch of the Human Brain Displaying the External Cortex and Internal Older Parts of the Brain

Higher level behaviours such as intelligent problem solving, learning and emotion are made possible by the cortex: the 2.5 millimetre thick surface layer of the cerebral hemispheres that grows out of the brain stem and consists of more than a hundred different functional areas made up of different types of nerve cells. The cortex has three main cortical divisions. MacLean calls them the archicortex, mesocortex and neocortex, each respectively coordinated by the reptilian, paleo-mammalian and neo-mammalian sections. The archicortex and mesocortex combine to form the limbic system. Psychologists believe the limbic system to be the seat of all emotions, attention and memory. However, there are no clear dividing lines between the limbic system and the neocortex, which contains the visual and auditory cortex and is responsible for a majority of perceptual processing.

MacLean's opinion is that each brain has its own special subjective, cognitive memory and the evolution of the first and simplest form of the cerebral cortex coincides with the development of emotional behaviour.

Reactions and behaviours generated by the archicortex and mesocortex are otherwise known as *fast primary emotions* and can be likened to the stimulus response experiments of the behavioural psychologists (Koestler 1967). These are the primal responses made before the stimulus gets to the cortex in order to be interpreted at a higher level. These emotions refer to the quick acting emotions that serve as the human's initial response to a stimulus and act as a survival mechanism. They include effects such as *fear* and *startledness*. They are the types of emotions that cause a human to jump out of the way of a large incoming object only to find later that it was just a beach ball. Of course, it could have been a large rock, so these instinctive emotions are justified.

MacLean's theory of emotions is an attempt to explain the human ability to produce and interpret high-level emotional responses. How emotional and physiological states interact has been theorised by several other psychologists. Two opposing points of view are presented in the James-Lange and Cannon-Bard theories (Lefton 1994).

Around the turn of the 20th century, two independent researchers, James and Lange proposed similar theories of emotion generation. So similar were these views that they are now jointly referred to as the James-Lange Theory. In short, the theory states that an emotive situation will elicit an appropriate set of physiological changes that will be sensed by the brain and interpreted it as an emotion. Essentially, when an event occurs in our environment it triggers behavioural, autonomic and endocrine responses. These responses are sensed by the brain and perceived as emotions. This implies that bodily changes produce emotional responses and not the reverse. In the opinions of James and Lange, people do not cry because they feel *sad*, but rather, people feel *sad* because they are crying.

A contemporary approach to this theory suggests the action of creating facial expressions produce emotions. For example, when a person smiles (first the action), the blood flow and temperature in the brain changes and induces pleasant feelings (Zajonc et al. 1989).

A theory in opposition to the James-Lange theory was developed by Cannon and Bard. They questioned that if several oppositional emotions, such as *joy* and *anger*, are produced from similar physiological changes, such as increased blood pressure and heart rate, then how does the individual deduce their emotional state? The Cannon-Bard theory states that when an individual becomes emotional, two areas of the brain are stimulated concurrently. These areas are the cerebral cortex and thalamus. Arousal of the cerebral cortex produces emotional responses and the stimulation of the thalamus creates physiological changes.

While the theories reviewed in this section provide much to contemplate, they are just that; theories. There are many arguments for and against each theory, especially with the later two holding opposing viewpoints (Lefton 1994). A newer approach by Pert and Snyder (Pert 1997), reviewed next, is supported with stronger research evidence and examines the effects of emotions at the chemical level of physiological change and may provide an indisputable theory of emotion generation.

2.2.2 The Molecules of Emotion

In 1985, Pert and Snyder published a controversial paper entitled *Neuropeptides and Their Receptors: A Psychosomatic Network.* Due to the radical theory proposed within, mainstream idealists ignored the research (Pert 1997). It was not until several years later that the scientific community found merit in the theory.

Traditionally, the limbic system was thought to be the seat of emotions where the neurons and associated neuropeptides in this area were responsible for producing emotional states. Neurons are found throughout the body and form parts of the electrochemical communication network that relays all incoming sensory information to the brain. Neurons do not physically touch, but are separated by gaps called synapses. When a neuron receives a signal it becomes active and excites surrounding neurons by releasing neuropeptides that bridge the synapses and allow the signal to be transferred to the next neuron. There are many types of neuropeptides that affect the transfer of the signal in many and varied ways. They perform a filtering process that determines how the final signal is interpreted when it reaches the brain. Neuropeptides identified as being emotion producing were initially thought only to exist in a confined area of the brain. However, the work of Pert and Snyder found that emotion-producing neuropeptides existed throughout the body.

The theory was so revolutionary that it threatened conventional thinking about emotions and emotion generation. In short the theory states, *neuropeptides and their receptors thus join the brain, glands and immune system in a network of communication between brain* *and body, probably representing the biochemical substrate of emotion* (Pert and Synder 1985).

This research has revealed that the neuropeptide receptors observed as responsible for emotional states and originally thought only to exist in the amygdala, hippocampus (a gland residing behind the amygdala) and hypothalamus have now been detected in high concentrations throughout the body. This has included the backside of the spinal cord, the nervous system's first synapse where bodily sensations and feelings are processed. Therefore, all sensory information passing between synapses via the emotion producing neuropeptides undergoes an *emotional* filtering process (Pert 1997). This operation assists in the brain's ability to deal with the deluge of sensory input that it receives.

The nervous system also carries signals, not only from the body to the brain, but also from the brain to the body. Emotional states or moods occur when emotion-carrying peptides are produced in the body's neurons. The presence of different emotional neuropeptides can create dissimilar reaction in an individual when exposed to the same stimuli.

Leuba and Lucas (1945) conducted an experiment involving the description of six pictures by three people when in each of three different moods. Each mood was induced by hypnosis and then the pictures were shown. Interpretations of the scenes in the pictures related to the moods of the viewer. For example, a picture of several university students sitting on the grass listening to the radio was interpreted by the same person as relaxing when they were in a *happy* mood, irresponsible when they were in a critical mood and competitive when they were in an *anxious* mood. Furthermore, as moods are evoked, the production of emotion-carrying peptides activates a particular neuronal circuit. This circuit facilitates in learning and recall. When a person is in a *happy* mood and the associated neuronal circuit is activated, they are more likely to recall positive emotional experiences (Pert 1997).

Having gathered evidence that suggests emotions are a physiological fact present in the human body and they impact on the cognitive processes of the brain, the question remains, how do emotions in a physiological sense have an impact on the field of artificial intelligence? Furthermore, how can these theories of emotions be integrated into an artificial agent that has no biological physiology? Should they be integrated into an artificial agent? Would it be pertinent to give an artificial being the ability to feel and express, for example, *sadness* or *happiness*?

The very idea of an emotional robotic device or machine seems nonsensical and an avenue that many AI researchers wish to avoid. This is not to say that the concept of emotions has been entirely ignored by the AI domain but due to the lack of definition and understanding of these states it has been mostly put aside for more logical pursuits. If artificial agents are to be programmed with emotional capabilities, these agents need to be provided with the capacity to interpret and respond to emotional cues. Due to their lack of biological embodiment, it is not possible to rely on complex systems of neuropeptides and hormones to configure such an understanding.

In contrast to the physiological theories of emotion are the theories and models that work on the principles of discrete emotional states. These are the cognitive theories of emotion.

2.2.3 Cognitive Theories of Emotion

The way in which the brain interprets emotional states from physiological actions is the foundation of the natural evolutionary path taken in psychology towards the development of cognitive theories of emotion. The physiological theories examined in the last section support this view. This standpoint on emotion generation, otherwise known as *appraisal theory* and considered by a number of psychologists, explains the generation of emotions using distinct evaluations of situations to categorise arising affective states. These appraisal theories involves grouping emotional states as positive or negative and continues to decompose these into sub-states according to other discrete characteristics such as expected/unexpected, certain/uncertain and high control/low control.

Incorporating elements from the James-Lange and Cannon-Bard physiological theories, Schachter and Singer observed that an individual does not only interpret their emotional state from their bodily one but also through their perception of the environment (Lefton 1994). Not only does an individual perceive their physiological state to determine their emotional state but the context in which it is occurring is also significant. For example, a father crying at his daughter's wedding may be perceived as a *happy* emotional state, whereas the same physiological condition at a funeral may be perceived as *sad*. To prove their theory, Schachter and Singer injected volunteers with epinephrine (adrenaline, a powerful stimulant). Although this drug increases heart rate, excitement and energy, the volunteers were not aware of how it would affect them. Immediately after the injection the volunteers where asked to fill out a questionnaire. In the room with them were students who had been injected with saltwater unbeknown to the volunteers. About the time the epinephrine's effects surfaced in the volunteers, the *under-cover* students would react with either a *happy* or *angry* disposition towards the questionnaires. After the experiment, the volunteers in the room with the *happy* students reported feeling *happy* and the volunteers in the room with the *angry* students reported that the drug made them *angry*. The researchers concluded from this that an individual has no immediate explanation for their physical state. They categorise it in thoughts available to them at the time.

Shaver et al (1987) believes that this kind of emotional research has moved too far ahead of itself. He believes that although cognition plays an important role in emotional interpretation, before a researcher can understand and perceive complex emotional states they must first understand the structure and concepts of emotions in their entirety. Shaver and his colleagues identify six atomic emotional states: *love*, *joy*, *anger*, *sadness*, *fear* and *surprise*. Although they believe that emotional experience is very distinct in different cultural situations, these six basic emotions or a subset thereof are universal throughout the theories of emotion (Smith and Kirby 2000, Magai and Hunziker 1998, Plutchik 1980).

The overlapping of these emotions can produce all other emotional states, but the cultural emphasis placed on each of the six basic emotions will determine how a complex emotion is perceived. The example they give is that *love* is a powerful emotion in the United States, but in Sumatra, *nostalgia* is most important.

The idea of elementary emotion blending is not new. Plutchik (1980) suggested a *colour wheel* of emotions as shown in Figure 2.2 where the basic emotions could be blended to create complex emotions in a similar way to fusing primary colours to produce secondary colours.

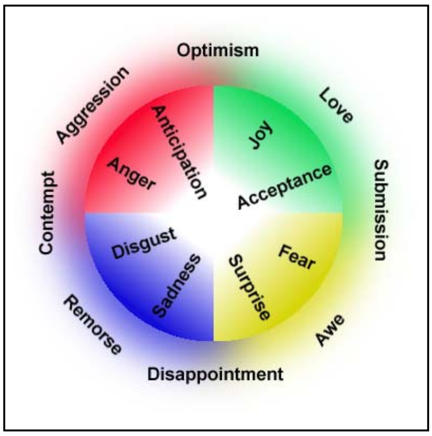


Figure 2.2 Plutchik's Emotional Range Circle

In his theory, Frijda takes the identification of emotions one-step further (Elliot 1992). Frijda combines emotional responses with appraisals and suggests that an individual not only uses cognitive appraisal to assess their emotional state in a situation, but as this cognition is taking place the individual is also preparing a set of behavioural actions in response to the situation. His theory combines the philosophies of the Behaviourist School of Psychology based on the stimulus-response theory (Koestler 1967) with an appraisal theory of emotion [3]. Often there is not enough time for a complex emotional evaluation of a situation, nor is it necessary.

Further definition for the process of appraisal, which formulates an emotional response, can be found in the cognitive appraisal theory developed by Ortony, Clore and Collins (OCC model) (Ortony et al. 1988). An illustration of this model can be found in Appendix A. The model incorporates and assesses the positive and negative *dispositions* that can occur in response to situations. It examines the valenced reactions toward one of three elements: the consequences of an event, the actions of an individual, and, the aspects of an object. The model assesses the human reaction to one or more of these elements as being: pleasing or displeasing, approving or disproving, liking or disliking. Given this assessment an emotional response category can be derived. Due to its discrete categorization of appraisal mechanisms, the OCC model lends itself admirably to implementation in AI systems and although not originally intended to act as a means to synthetically generate emotions in computers, it has provided a basis for a number of research projects in affective computing including Elliot's Affective Reasoner (Elliot 1992) and Reilly's Artificial Actors (Reilly 1996), examined further in this chapter.

Table 2.1 displays the values used in the OCC model (and subsequent AI applications) to appraise an event and determine an appropriate emotion. The table is a subset of the OCC appraisals displaying eight of the possible 22 emotional states mentioned in the theory.

Consequential Disposition	**Focus Object (self or other)**	**Consequences for Focus Object**	**Emotion**
pleased	other	desirable	Happy For
pleased	other	undesirable	Pity
displeased	other	desirable	Resentment
displeased	other	undesirable	Gloating
pleased	self	relevant	Hope
pleased	self	irrelevant	*Joy*
displeased	self	relevant	Fear
displeased	self	irrelevant	Distress

Table 2.1. A Subset of the OCC Cognitive Structure of Emotion

2.2.4 Understanding the Chaos

In all the chaos of emotional models, theories, experimentation and implementation, there is a consensus that emotional states can be divided into two distinct categories; lower

[3] This is what MacLean might have identified as the control of the reptilian brain (see Section 2.1.1) or with the complex emotional responses generated by the neo-mammalian brain.

(or survival) emotions and higher emotions. These two divisions are identified in MacLean's Three Brain Theory (Koestler 1967) as the difference in function between the reptilian brain and the neo-mammalian brain.

The first category is that of emotional drives. These drives or urges can be classified according to their source, pleasure rating and strength. This category of emotion is present in philosophies of the Freudian School of Behaviour (Furth 1987) as a sexual drive (Koester 1967) and the Behaviourist School of Psychology (Hull 1943) as a stimulus-deprivation drive (Koester 1967). MacLean (1958) refers to the two basic emotional drives in animal behaviour; *self-preservation* and *preservation of the species*. From these two ultimate urges all motivational goals of animal behaviour can be derived and are individualistic of a defined being. The four emergent classifications from these categories for emotional drive as identified by Toda (Padgham and Taylor, 1997) are *emergency*, *biological*, *cognitive*, and *social*.

The emergency drive refers to the mechanism that creates behaviours when an individual is unexpectedly and suddenly displaced from its homeostatic state. An example is how an individual reacts to an automobile accident. Usually unplanned, such an event might motivate an individual to perform acts that will rectify the situation or return the individual to (or as close as possible to) its original state. Biological urges are again drives to return an individual to homeostasis. Many of these drives are well known, cyclical and not surprising occurrences to the individual. For example, *hunger* occurs in a human when the body has a low blood sugar level. Simplistically, warning signals are sent to the brain, which in turn generate *hunger pain*s in the stomach (Lefton 1994). This stimulus creates a motivation to eat. The cognitive drive creates curiosity in an individual. This in turn motivates an individual to crave knowledge, to learn more about their environment and how they can manipulate and interact with it. Finally, the emergent social urge refers to an individual's need for companionship or interaction with others. Such drives cause individuals to comply with peers or seek acceptance among others. These have a significant effect on behaviour in social situations.

This set of emotions will be referred to, in this dissertation, as *survival* or *lower* emotions (*drives* or *urges)*. The second type of emotion that is to be addressed is that of the higher, neurological emotions.

If survival emotions are those that generate goals and motivation, then higher emotions are the resultant mental (and in turn physical) states generated by attempts to satisfy these goals. These emotions include feelings such as *happiness, anger, sorrow, guilt* and *boredom.*

Attempts to satisfy goals driven by the survival emotions give rise to an appropriate set of higher emotions. However, the emotions generated are difficult to categorise due to their complex relationships. Some emotions are viewed as being opposite such as *happiness* and *sadness.* These are easily distinguished because one is pleasant and the other unpleasant. On the other hand, *fear* and *anger* could also be considered opposite (when comparing the respective *flight* or *fight* responses), but clearly not in the same way (Smith and Ellsworth 1985). Popular theories for categorizing these emotions in psychology are the cognitive appraisal theories as mentioned in Section 2.2.3.

When considering all the research and published material on emotion generation, determining a way in which emotions could be of benefit in an artificial intelligence is a complex issue. There is no doubt in the minds of the researchers who developed the theories presented in this section that emotions are an important part of human intelligence. The next step in the AI domain is to determine which emotional theories and concepts could enhance the intelligent information processing in machines. While some researchers talk about machines being able to feel and have a conscience (Picard 1998, Sloman 2001), the foundation of this dissertation is not to contemplate such a development, but to examine how emotions may enhance human intelligence and to develop a model that captures these concepts and applies them to the way in which the artificial beings manipulate and process information.

The following section examines five components of emotion found in the field of affective computing. They are by no means reflective of any structure or model of affective computing, but rather lend themselves to categorise and present current research in a well-organised manner.

2.3 Affective Computing

As this area is a one of the latest research developments in the AI domain, it is also a fast growing one. Many designs and architectures for emotionally capable artificial intelligences have been proposed in the last five years. These designs focus their attention on specific aspects of emotion synthesis. Understandably, an emotionally complete architecture is yet to be developed. While the research in this dissertation does not boast to have achieved this, it too is another small part of the puzzle.

The motivation in this domain has been to allow computers to emulate areas of human behaviour in order to increase information-processing efficiencies. These abilities include flexible and rational decision making, reasoning with limited memory, limited information and relatively slow processing speed, social interaction and creativity.

According to Picard (1998), if computers are ever to exhibit or have emotions, they must be capable of synthesizing and generating them. Picard (1998) outlines five components present in a human emotional system, where all or subsets of these must be present, in an affective computer system. These are 1) emotional behaviour; 2) fast primary emotions; 3) cognitively generated emotions; 4) emotional experience: cognitive awareness, physiological awareness, and subjective feelings; and, 5) body-mind interactions. An examination of these categories is pertinent at this point and the following sections examine the current state of research with respect to each of these areas and collectively define the domain of affective computing.

2.3.1 Emotional Behaviour

The area of emotional behaviour refers to systems that display outward emotions. This does not necessarily mean the system has the means to internally process emotional responses as humans do, but nonetheless is capably of appearing emotional.

A robotic device that seemingly displays emotional behaviour is that of W. Grey Walter's tortoise mentioned in Penrose (1989) which was capable of autonomous motion and had the ability to plug itself into a power socket to recharge its batteries. Although a simple decision making algorithm was obviously deployed to cause this behaviour, it still exhibited *emotional behaviour* (Picard 1998).

Another example of the display of emotions is the well-known boot-up sequence smiley face on the Apple Macintosh computer. While it is very simple, the user could infer that the computer is in a *happy* state or not. Though logically it is not, some human users without in-depth thought about the computer's emotional state might gain a feeling of well being from the machine.

A more complex example of a computer-generated display of emotions is that developed by Haptek (Kaehms 1999). The company has developed a series of three-dimensional animated characters that are capable of accepting scripts of text and converting them into speech. Beside this ability, the characters also possess the ability to express emotion using voice inflections and facial expressions when explicitly programmed. These characters are an example of the computer being able to convey a more complex emotional state to the user while not having autonomously generated that state.

Another example of a display of emotions can be seen in Pixar Animation Studio's lamp animation (Luxo Jr. 1986). This short animation uses bodily movements to breathe life into a small desk lamp and metaphorically represent it as a small child playing with a ball. Although the lamp does not possess facial expressions, it conveys the supposed emotional state of the lamp and convincingly evokes sorrowful responses from the audience when the lamp's ball bursts and it divulges its *sadness* by hanging its head represented by the lamp shade.

The next stage in emotion-machine integration is to program the machine with the capacity to generate these emotional behaviours independently. In fact, in the case of W. Grey Walter's tortoise, this machine was programmed with the ability to recognise depletion in power and to act on it. This is synonymous with human *hunger* and fits into the next category of fast primary emotions.

2.3.2 Fast Primary Emotions

The psychological and physiological background of the fast primary emotions was previously examined in Section 2.2.1. The brain, in response to stimuli that evoke a survival mechanism, generates these emotions. One such application of fast primary emotions in computers would be to implement survival mechanisms in a robot where it could quickly react to protect itself. This could include such things as getting out of the way of moving objects, ducking, the heightening of perceptual sensors or renewing its power supply.

Another subset of these primary emotions is drives or motivational urges. Many affective systems implement motivational mechanisms in unison with goal-orientated belief systems (Reilly 1996, Padgham and Taylor 1997, El-Nasr 1998). These systems work by generating internal motivations that activate the goals of the artificial agent. The ideas behind these systems are essentially the same. A timing mechanism or change in state of the agent generates an urge that in turn triggers a goal. Balkenius (1995) identified several categories of motivational drives that he believes play an important role in cognitive processes and uses these in the simulation of animal behaviour.

The homeostatic drive or the drive that is generated by a violation of homeostasis includes, for example, *hunger* and *thirst,* but also the responses to heat and cold. The noxious drive considers the signals from noxious stimuli and includes the sensation of *pain*. Cyclical drives, which vary with the time of day or the year, are generated by an oscillator influenced by external stimuli such as the length of the day, odours, or the amount of light in the environment. These include *tiredness* and *wakefulness*, the *sexual* drives of most animals, and migratory drives. The default drive influences the animal when no external cue is present which commands the animal to do something else. The exploratory drive, which is similar to the default drive in that it tries to activate behaviour when the animal has nothing else to do. Finally, the anticipatory drive is internally generated but does not relate to any present need of the animal. This drive can influence any other drive and its purpose is to simulate a drive, typically a homeostatic one, which is not present.

One of more of these drives have been implemented in a number of affective agent systems such as El-Nasr's PETEEI (El-Nasr 1998), Balkenius's simulated artificial creatures (Balkenius 1995), Padham and Taylor's Dog World agents (Padgham and Taylor 1997) and Canamero's Abbotts (Canamero 1997). In all cases the drives are programmed into the agent as a number of gauges and cyclic mechanisms. The cyclic mechanisms update the drive gauges and when a threshold value on a gauge is reached a goal is triggered. For example, an agent may have a *hunger* gauge that is cyclically updated to simulate *hunger* increasing over time. At a predetermined threshold value on the *hunger* gauge, a goal to seek food will be activated.

These types of primary emotions are easily simulated and methods for implementing them in artificial beings are quite elementary. The next stage in implementation of such mechanisms in AI is that of generating appropriate internal state changes within a machine and have that machine understand, react and learn from them. This capability is called emotional experience.

2.3.3 Emotional Experience

An emotional experience is the ability for a machine to have an awareness of its emotional state. This awareness may be accompanied by changing physiological states (an example in humans being an increased heart beat) and may be generated by internal subjective feelings. These internal feelings, or *gut feelings* (Koestler 1967), are very difficult to implement in computers due to the fact that in humans they are produced biochemically (Pert and Synder 1985).

As previously mentioned, existing AI systems do not have the capacities for consciously recognizing or generating emotional states. Goal-orientated systems are constructed from motivation mechanisms that cause the machine to appear as though it *wants* and *cares* when in fact it has only been programmed to give the superficial appearance of these humanistic traits (Sloman 2001).

Sloman does not speculate on whether machines may one day have conscious feelings but does provide an informative architecture for a human-like agent that can emulate the process of emotional experience. He conjectures that the information processing capabilities of humans make use of three different architectural layers: a reactive layer, a deliberative layer and a self-monitoring layer. These layers mirror the three separate brain layers of MacLean's theory: the reptilian brain, the paleo-mammalian brain and the neo-mammalian brain respectively.

The reactive layer takes care of the fast primary emotions, the deliberative layer deals with higher emotional states and the creation and execution of plans and the self-monitoring layer oversees the improvement of the processes stemming from the other two layers. It is the final self-monitoring layer that gives the agent the ability to become aware of its emotional experiences, to identify them and to learn from them.

2.3.4 Body-Mind Interactions

This section refers to the influence that emotions have on the body and mental states of humans. The study of emotions from a neural basis is an active field of research (Fellous 1999). Emotions can influence the body's state and in turn the body's state can influence emotion. The physiological aspect of emotions is examined in-depth in Section 2.2.1. In particular, research by Pert (1997) examines the low-level physiological changes that take place within the human body with respect to emotions.

Canamero (1997) describes a AI system where emotional state changes cause changes to a system of synthetic hormones and in turn, changes to the states of these hormones also influences the emotional state of the artificial beings. The emotionally influenced beings, called the "Abbotts", inhabit a two-dimensional world with another race of unemotional beings called the "Enemies". The Abbot's behaviour is determined by motivational and emotional mechanisms that correspond to internal state changes within the agent. These internal states (constructed from synthetic hormones) represent the agent's physiological state. The agents can choose to attack, withdraw, drink, eat, play and rest. Each behaviour has an effect on the agent's internal state, which in turn may trigger a motivation that will cause a different behaviour.

2.3.5 Cognitively Generated Emotions

These emotions may be explained as the emotions that follow primary emotions. They are the emotions generated on appraisal of a stimulus. Over the past 15 years this categorical approach to the study of emotion has become prominent in psychology (Ortony et al. 1988). Groups of theorists known as cognitive appraisal theorists have attempted to explain different emotional reactions by categorising emotions from evaluating the situations from which the emotions arose. The long term interest in the classification of emotional experiences has traditionally only been defined in two dimensions*; pleasantness* and *arousal* (Smith and Ellsworth 1985). However, these appraisal dimensions were found to be inadequate in classifying the range of human emotional states and newer models arose using more complex appraisal dimensions.

One such model by Ortony, Clore and Collins (OCC) (Ortony et al. 1988) examines the valenced reaction to one of three elements: 1) the consequences of an event; 2) the actions of the agent; and, 3) the aspects of objects. The model assesses the human reaction to these elements as being either, pleasing, displeasing, approving, disproving, liking or disliking.

The Oz project (Bates et al. 1992) is the effort of researchers at Carnegie Mellon University. The primary focus of their work has been in the development of agents capable of generating emotions and moods for the purposes of autonomous animated characters. Their broad architecture called *Tok*, incorporates a number of sub-systems one of which, Em, is primarily for the generation of emotions. In the Em architecture, emotions are structured hierarchically and categorised as either positive or negative in nature. The emotions are generated in the goal-orientated agents when a goal is completed. The intensity of the emotion is calculated from the importance of the goal. Although based on the OCC model, Em differs slightly in that its likelihood of success or failure, rather than an appraisal of the event to which it relates, determines the importance of the goal. The inspiration for this research has been to build a set of tools that can be used by creative artists to build believable artificial characters and scenarios for dramatic purposes. Artists have the ability to create new characters and situations by manipulating, by hand, the hard-coded cognitive and behavioural mechanisms of *Em*.

Another application of the OCC model has been the development of the Affective Reasoner, a research project of Elliot (1992). The Affective Reason is a collection of AI programs personified into a multimedia-computing agent. Each agent has a two-part personality consisting of a disposition that influences how the agents perceive their environment and a temperament that controls their affective behaviour. As the agents react to the dynamic environment in which they have been placed, emotional responses to environmental stimuli are produced in one or more of the twenty-four different affective categories with which they have been programmed. The twenty-four categories and their associated emotional responses can be found in Appendix B. The agent's disposition and temperament combine to form the personality of the agent. This research demonstrates how modelling personalities and social relationships can generate emotions. The agents' are capable of modelling three types of social relationships and corresponding emotions; friendship; animosity and empathy.

Today, this type of reasoning is the most frequent way of generating emotions in artificial intelligences through cognitive reasoning.

2.4 Affective Agent Architectures

In recent times, there have been a number of architectures that have been designed to produce artificial agents capable of expressing and processing emotions. These models cover a wide range of affective phenomena and differ broadly between implementations. As a complete examination of these architectures would constitute a publication in its own right, a comprehensive review of these models will not appear in this dissertation.

Having said this, the current architectural designs, their application and evaluation provide essential research information that should be considered in the development of any new affective agent architecture. To this end, before the EMAI architecture (the product of this dissertation) is explained in Chapter 3, this section provides a brief yet expansive overview of several current models and ideas circulating in the Affective Computing research domain that have influenced the work contained in this dissertation.

2.4.1 Blumberg's Silas T. Dog

An artificial agent that uses primary emotions (see Section 2.2.1) to motivate behaviour is Blumberg's architecture for *Silas T Dog* (Blumberg et al. 1996). The architecture illustrated in Figure 2.3 is an autonomous, animated virtual dog, consists of three layers: the Behaviour System layer, the Motor Controller layer and the Geometry layer. While the motor controller and geometry layers perform the necessary tasks to represent the agent as a virtual computer character in a virtual three-dimensional world and control its movements, the behaviour system controls the agent's intelligent behaviours.

The Behaviour System is a network consisting of independent goal-orientated entities called *Behaviours*. *Behaviours* are grouped into mutually inhibiting groups called Behaviour Groups, which are in turn organised into a loose hierarchical structure. The goal of a *Behaviour* can range from very general (for example, "reduce hunger") to very specific (for example, "chew food").

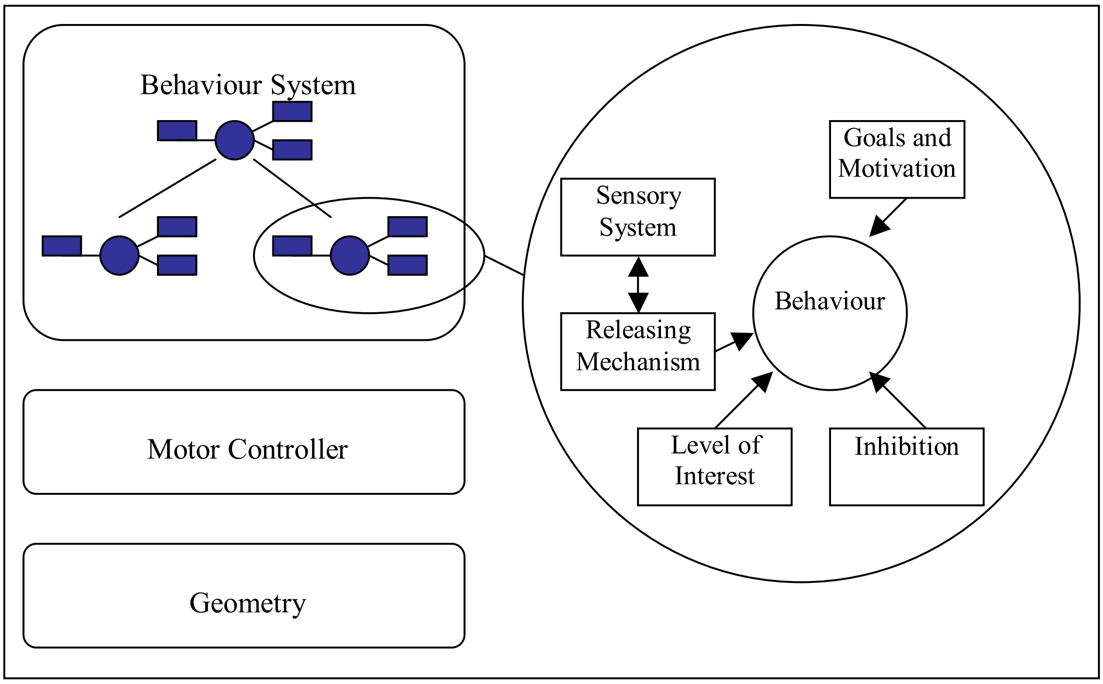


Figure 2.3 An Overview of Blumberg's *Silas T. Dog* Architecture

Before the agent can exhibit a given behaviour, the *Behaviour* entity must evaluate its relevance given external stimuli and internal motivations. External stimuli are passed from the Sensory System (that senses the external world) through the Behaviour's Releasing Mechanism. The Releasing Mechanism filters the sensory input for relevance toward the *Behaviour* entity. Motivations and goals are represented by Internal Variables. The *Behaviour* entity combines the values from the Releasing Mechanism and Internal Variables to determine the Behaviour's Level of Interest. This Level of Interest is used to model the strength of a particular Behaviour.

The process of Behaviour System begins by examining the Behaviour Group at the top of the system's Behaviour Group Hierarchy. *Behaviour* entities within this group compete for control of the agent. The winning *Behaviour* entity is decided by examining the Releasing Mechanism, Internal Variables and Level of Interest of each *Behaviour*. This *Behaviour* entity is then said to become active. If the active *Behaviour* is linked to a child Behaviour Group in the hierarchy, the process of competing begins again on a more specific level. This process continues until the most specific form of a *Behaviour* is identified (that is the bottom level of the hierarchy is reached). The chosen *Behaviour* entity then communicates with the Motor Controller layer to create the outward behaviour of the agent. This architecture provides a means of managing goal-orientation and motivational mechanisms controlled by lower level emotions such as *fear* and *hunger*.

2.4.2 PETEEI (A PET with Evolving Emotional Intelligence)

The PETEEI architecture (El-Nasr 1998) has three major components; an emotional process model, a cognitive process model and a learning process model. The architecture has been used to build an agent with the behaviours of a pet dog. A PETEEI agent operates by perceiving data in the form of an event occurrence. On receipt of this external stimulus the agent's emotional model evaluates the event's desirability and expectation and proceeds to activate an emotion and produces an appropriate emotional behaviour. The emotion is decayed linearly using a constant decay rate function. Each emotion is triggered according to predefined rules based on the OCC model (see Section 2.2.3). Behaviours are then selected using a set of predefined production rules. For example, *if* ***anger is high*** *and* ***event is dish-was-taken-way*** *then* ***behaviour is bark-at-user***. Besides reacting to external stimuli, a PETEEI agent also generates internal stimuli in the form of motivations. These motivations include states such as *hunger*, *fatigue*, *thirst* and *pain*.

These states are stored as individual variables and updated linearly using measures such as time.

2.4.3 Emotion Based Control (EBC) Framework for Autonomous Agent

Velasquez's Emotional Based Control (EBC) Framework (Velasquez 1999), as shown in Figure 2.4, integrates emotional processing with other agent systems that control perception, attention, motivation, behaviour and motor skills in a robotic agent.

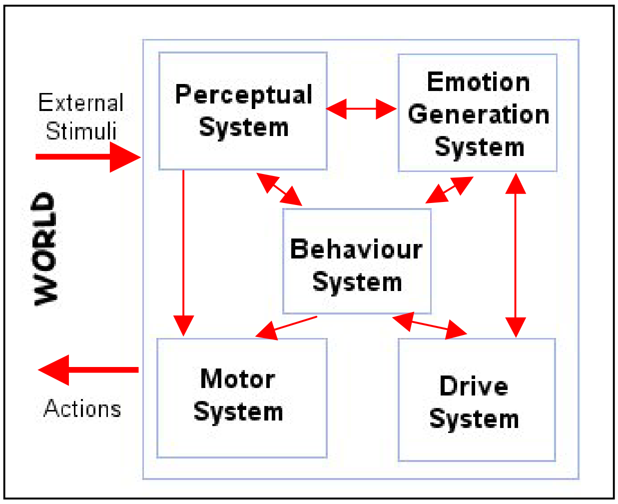


Figure 2.4 EBC Framework for Autonomous Agents

In an EBC agent, the Perceptual System accepts external stimuli from the outside world. This data is provided to the Emotion Generation System and the Behaviour System. Both these systems also receive data from the Drive System. The Emotion Generation System determines the emotional significance of incoming stimuli and ensures that future behaviour and perception are affected accordingly.

The EBC framework is based on the neurobiological premise of *affect programs*. Affect programs relate to preprogrammed brain systems that generate and coordinate stereotypical behaviour in biological agents (Velasquez 1999). The EBC framework has six different types of affect programs generated by the Emotion Generation System, that correspond to different primary emotions; *fear, anger, sorrow, happiness, disgust and surprise*. Each affect program activates its own series of circuitry in an EBC robot that causes it to behaviour in a preprogrammed manner. For example, the *fear* affect program can cause the robot to exhibit cowering and retreating behaviour. Each affect program is programmed with a temporal decay function that controls the time for which the affect program is in operation once activated.

Although not fully explored, the EBC framework could also be used to generate behaviours relating to higher order emotions such as *guilt*. For example, by implementing a system of emotion blending based on the theories of Plutchik (see Section 2.2.3) the affect programs for *fear* and *joy* could be activated simultaneously to produce the emotional behaviour for *guilt*.

2.4.4 Minsky's Commonsense Agent

Minsky (2000) suggests that the reason researchers have not yet developed a machine capable of performing many of the mundane tasks that people seem to have no problem executing is that AI researchers are not using adequate large-scale models that could be used to orchestrate systems capable of *commonsense* or *resourcefulness*.

The large-scale architecture conceptualised by Minsky examines the way in which the human brain employs its resources and he builds on a previous 3 layered architecture devised by Sloman (2001) to create a 5-layered model of an artificial mind.

Minsky's incomplete architecture consists of a reactive layer, a deliberate layer, a reflective layer, a meta-management layer and a self-conscious layer. The reactive layer is the basic instinctual area of the architecture and deals with external sensory input. It is responsible for the agent's reactive behaviours (reflex actions) that require very little processing, for example, moving out of the way of a large incoming object. The deliberate layer produces the agent's intentional behaviours or calculated actions that have been premeditated. The reflective layer of the agent contemplates and learns from past behaviours. It accesses behaviours produced by the other layers and learns and adapts based on the outcomes. The meta-management layer monitors the agent's temporal progress of the processing taking place in the other layers. It manages the processes and ensures that timely decisions are made and appropriate behaviours are performed. Finally, the self-conscious layer processes information pertaining to beliefs about the agent's own being or how the agent believes that it is perceived by others.

Minsky's large-scale architecture of the mind is currently incomplete. How the layers interact with one another and other parts of the agent remain undefined. Minsky (2000) explains that the architecture will be further developed and discussed in his forthcoming book, *The Emotion Machine*. While Minsky does not explicitly mention emotions with respect to intelligence he does hint that there would be emotional switches that would interact with the self-conscious and meta-management layers. The title of his forthcoming book and quotes made by Minsky (Stork 1997) suggest that he believes that concepts on emotion should be an integral and fundamental part of AI.

2.5 Summary

Emotions are an ill-defined concept. There have been many and varied theories developed over the years that attempt to explain emotions from a holistic view or concentrate on atomic components of emotional experience. Whatever the definition, there is current research to suggest that emotions play a crucial role in human intelligence.

Attempts to replicate human-like performance in the domain of AI have been successful with respect to mimicing elements of intelligent behaviour, but limited in a holistic sense. Visual and motor capabilities of current AI systems are far from being as precise and elegant as that of a real animal let alone emotional capabilities. Sloman (2001) suggests that researchers need to examine not only their understanding of the possibilities for information processing machines but they should also deepen their understanding of the concepts of emotions and how the two could be integrated. The question is not, *Which is the correct theory of emotion?* but rather, *Which theory bests suits the needs of a particular AI system?*

Theories of emotion have been integrated best into artificial beings that are goal-orientated. The domain of affective computing is laden with concepts such as *goal, motivation, desire, intention* and so on. Cognitive appraisals have been a popular concept for integration into current goal-orientated artificial agent architectures. However, like emotion definitions, cognitive appraisal theories are also many and varied. The research within the domain of affective computing covers all aspects of the concept of emotions. Some applications focus their efforts on one aspect, while others implement several components in order to create more believable and emotionally intelligent agents.

Using the emotional theories and concepts as presented in this chapter, this dissertation presents a unifying artificial agent architecture that is influenced by emotions. In keeping with the well documented goal-orientated approach, this new agent integrates a number of theories on emotion, builds on several of the agent designs already in use in the affective computing domain, improves on the approach of event assessment using appraisals and presents a new method for decision-making and reasoning about behavioural choices.

The next chapter introduces this new emotionally motivated artificial intelligence architecture and examines its components in illustrative detail. This architecture is the original concept of the author and is the major contribution of the work presented in this dissertation.

3. Emotionally Motivated Artificial Intelligence

Only the discipline of logic saved my planet from extinction.
MR SPOCK – Star Trek the Original Series, "Let That Be Your Last Battlefield".

3.1 Introduction

This chapter examines the Emotionally Motivated Artificial Intelligence (EMAI) architecture developed by the author as a result of the research herein. The design of this architecture was motivated by research into the affective agent domain and the identification of the shortage of agent architectures that have the capacity for decision-making influenced by a mechanism that simulates human emotional intelligence (Goleman 1995). Picard (1998) affirms the importance of this type of emotionally influenced decision-making in computers. She suggests that if affective decision-making were integrated into computers it would provide a competent solution to emulating the intelligence of humans where decisions are often made with insufficient knowledge, limited memory, and relatively slow processing speeds. Emotions are an integral part of human decision-making and by giving machines a similar mechanism it could help in problem solving where options cannot be fully explored, data is incomplete and processing time is short.

This type of emotional intelligence in computers is implemented in the EMAI architecture as three major conceptual areas; the Knowledge Area, the Constructive Area and the Deliberate Area. The Knowledge Area contains the agent's beliefs and states, the Constructive Area acts as the agent's cognitive domain and the Deliberate Area performs scheduling and execution of behaviours. The examination of these three areas is the focus of this chapter.

3.2 Overview of the EMAI Architecture

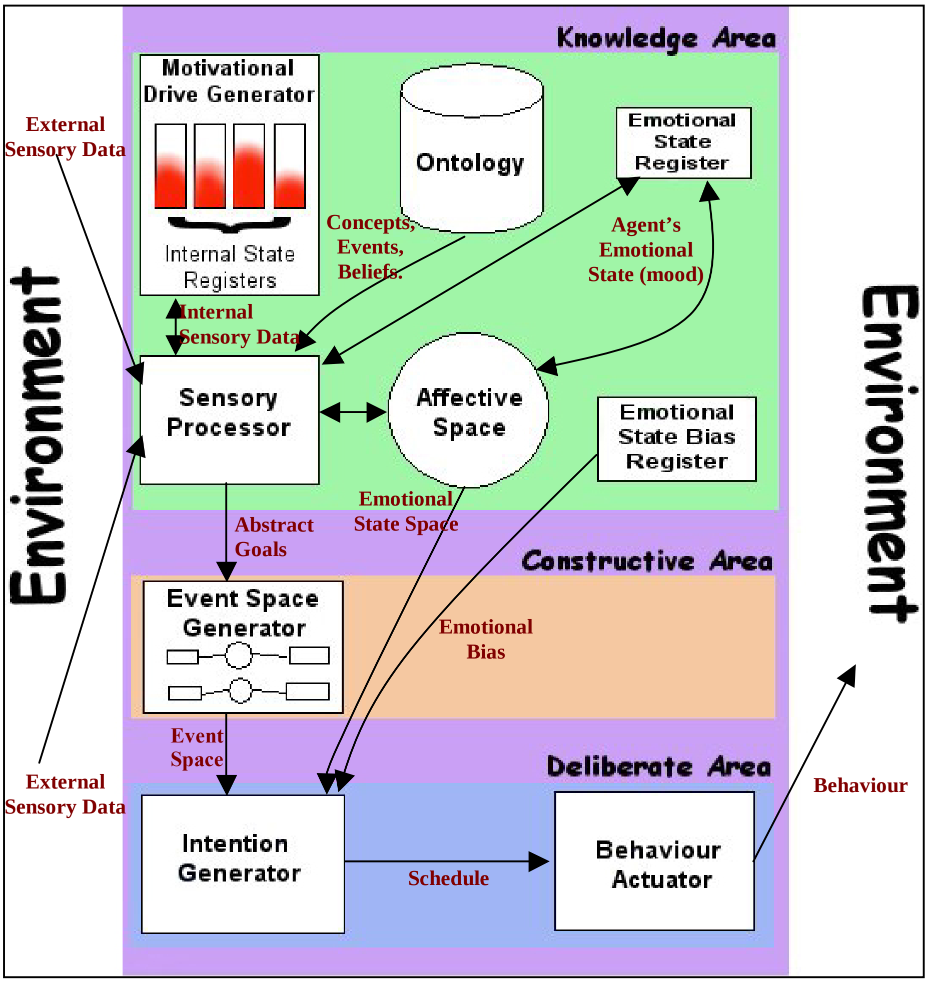
The objective of the development of the EMAI architecture has not been to examine all aspects of human behaviour, which would require the examination of elements such as physical conditions, cognitive capabilities and social status (Schmidt 2000), but to focus on emotional states and processes that affect human decision-making behaviour. As there are many aspects to human behaviour that should be modelled to build a complete and accurate representation of a human's cognitive abilities, the collection of these components makes for a difficult study of the individual components in isolation. For this reason, the

EMAI architecture is by no means complete. Because of the complex nature of emotions and the plethora of application areas, this architecture has been designed to focus only on the modelling and synthesis of emotional states in relation to decision-making about the performance of behaviours. Keeping this in mind, this architecture implements a psychological basis for emotion generation that concentrates on motivation and the selection of behaviour.

There are three conceptual areas of the EMAI architecture's structure that interact to motivate an agent by implementing emotional intelligence. They are the Knowledge Area, the Constructive Area and the Deliberate Area. Each of these components can be seen in Figure 3.1, which shows the conceptual working areas of the EMAI architecture. The EMAI architecture is a complex system consisting of an ontology, motivational drives, an Affective Space and decision-making mechanisms. A common enigma in building agents with emotional competence is organising the complex way in which these conceptual components interact. For example, emotions affect behaviour and in turn behaviour affects emotions.

For this reason, the Affective Space, shown in Figure 3.1, becomes the focus of the agent architecture. The Affective Space acts as an emotional filter that influences an EMAI agent's perception of its beliefs, the environment and as a consequence how it behaves. The basis for such a mechanism has been born from the physiological theories of emotion such as those of Pert (1997) explained in Chapter 2 where emotions are described as having a filtering effect on human sensory input. The implementation of the Affective Space in the EMAI architecture is unique in emotional agent architectures. While there have been a number of emotional agents that have preceded EMAI (see Chapter 2), none have explored emotions in the multi-dimensional sense as is presented in this dissertation. An in-depth discussion of the Affective Space will be delayed until Chapter 8.

The EMAI architecture consists of several major processing and knowledge representation areas. Each of these areas works together in a complex network of information gathering, manipulating and updating. As shown in Figure 3.1, any agent implemented using the EMAI architecture receives external sensory data from its environment. It also processes internal sensory data from the Motivational Drive Generator in the Knowledge Area. Internal State Registers in the Motivational Drive Generator simulate low level biological mechanisms. The Sensory Processor and the Affective Space integrate both types of sensory data into the agent's belief system via an emotional filtering process. Sensory input (both internal and external) received by the Sensory Processor may activate goals in the agent. The goals are processed by the agent's Constructive Area where plans are chosen that will satisfy these goals. These plans are generated by the Event Space Generator, which generates a series of events that could be performed by the agent to satisfy its goals. Before the agent schedules the events for execution in the Deliberate Area, each event is ordered by the Intention Generator in collaboration with the Affective Space and sorted from most *liked* to least *liked*. Once the agent has the list of events sorted by emotional effect, the Behaviour Actuator begins executing them in order. The events executed by the Behaviour Actuator at any moment in time become the EMAI agent's outward behaviour.



The in-depth examination of the analytical and information handling processes in each of the areas of the architecture is given in the following chapters. The remainder of this chapter is dedicated to an exploration of each of the three conceptual areas and their sub-components.

3.3 The Knowledge Area

The Knowledge Area of the EMAI architecture consists of the agent's belief system (represented by the Ontology), the agent's emotional state (represented by the Emotional State Register), the agent's preferred emotional state (represented by the Emotional State Bias Register), the agent's motivations (created by the Motivational Drive Generator), the agent's perceptions (processed by the Sensory Processor) and the agent's emotions (represented by the Affective Space). These components work in unison to represent the agent's knowledge and processes it needs in order to function.

The Affective Space is the focal point of the EMAI architecture as this construct provides any agent built using this architecture with the ability to make affective (emotion based) decisions. In the following sections, each of these components will be reviewed, while detailed discussion of the mechanics of these components will be outlined in Chapters 4 through 9.

3.3.1 Ontology

The Ontology of the EMAI architecture contains the Type and Relation Hierarchies, a set of Canonical Graphs and the Goal Hierarchy. These components work collectively to represent the EMAI's knowledge base. The knowledge base contains all the data that an EMAI agent needs to operate within a specific domain environment. Not only are the agent's beliefs about itself and the environment stored in the Ontology, but they are also structured into partially-ordered hierarchies that assist the agent in operations of deduction, association and belief revision. The Ontology will be examined in Chapter 4.

3.3.2 Motivational Drive Generator

The Motivational Drive Generator consists of three drive mechanisms (homeostatic, cyclic and default) and a set of Internal State Registers. The Internal State Registers record the agent's physical state. Each register is represented by a single gauge that stores the value for a particular physical state, for example *hunger*. The number of Internal State Registers implemented depends on the application for which the EMAI agent is being used. For the simple pet-dog agent that will be examined in Chapter 9, three Internal State Registers are implemented; these are *hunger*, *fatigue* and *arousal*. The method of storing the agent's internal state, as a series of individual variables is not new, other affective agent architectures such as Silas (Blumberg 1996) and PETEEI (El-Nasr 1998) (see Section 2.4.4) do the same.

The Internal State Registers are synonymous with the *Internal Variables* in Blumberg's architecture, Silas. The Internal Variables are used to model the agent's internal states such as *hunger*, *thirst, fear* and *aggression* (Blumberg 1996). The values of the Internal Variables are updated at each time-step of the agent's simulated life. A growth rate function, dampening rate function, the agent's behaviour and external stimulus affect the values of the variables. The Internal Variables of Silas represent motivations that drive the behaviour of the agent as they do in the EMAI architecture.

The same type of motivational mechanism is included in PETEEI. The PETEEI architecture has a number of *Motivational States* such as *pain*, *hunger*, *thirst* and *tiredness*. Each Motivational State is updated by its own function that assesses how a PETEEI agent is being affected by its environment or its own behaviours. Once the value of a Motivational State reaches a predetermined calibre, it sends a signal to the agent's cognitive process telling it to attend to the associated motivation.

Unlike the Motivational States of the PETEEI architecture, the Internal State Variables of the EMAI architecture do not perform any processes. They are updated and monitored by three independent drive mechanisms. The three drive mechanisms that monitor the values of the Internal State Registers are inspired by the work of Balkenius (1995) (see Section 2.3.2) who categorise the many motivational urges involved in animal behaviour into three types; *homeostatic*, *cyclical* and *default*. These drive mechanisms pass internal sensory input to the Sensory Processor (explained in the next section) when certain conditions are met.

The Homeostatic Drive helps to maintain an EMAI agent's equilibrium. It works to return an agent's internal state to the default state when the agent becomes excited. It does this by continually tracking the values of the Internal State Registers and sending internal sensory data to the Sensory Processor when the values on the registers fall outside of a normal range. For example, the agent may have an Internal State Register for storing its *body temperature*. The Homeostatic Drive would monitor this register. Should the value of the agent's *body temperature* register rise or fall outside of a threshold range, the Homeostatic Drive mechanism would send an internal sensory data value representing *cold* or *hot* to the Sensory Processor.

The Cyclical Drive maintains an EMAI agent's internal *body clock.* The Cyclical Drive synthesises a constant rate of change in the agent's Internal State Registers that are affected by time differentials in its environment. At predetermined temporal intervals the Cyclical Drive mechanism will increase or decrease the values on temporally sensitive Internal State Registers. For example, the agent may have an Internal State Register to represent the agent's *hunger*. As time passes, the Cyclical Drive mechanism would increase the value of this register. This would represent the agent becoming *hungry*. At a particular threshold level, the Homeostatic Drive would detect that the agent was *hungry* and send appropriate internal sensory data to the Sensory Processor for computation.

The Default Drive mechanism produces deliberate actions within the agent. This drive relays internal sensory data to the Sensory Processor that represents a need in the agent to explore its environment. Essentially, it is used to produce behaviours in the agent that are neither homeostatic nor cyclical. This mechanism prevents the agent from *doing nothing*, when the other drives are not sending internal sensory data to the Sensory Processor.

The internal sensory data passed to the Sensory Processor from the Motivational Drive Generator represents *instincts* or *urges* (characteristics of survival emotions discussed in Chapter 2). This information is used by the Sensory Processor to trigger goals within the agent and in turn formulate outward behaviours.

3.3.3 Sensory Processor

The Sensory Processor works by processing internal sensory data passed from the Motivational Drive Generator and external sensory data collected from the environment. When the Sensory Processor receives information it checks to determine which of the agent's goals (if any) should be triggered in order to deal with the data. The triggered goal's associated plan is passed to the Constructive Area for processing.

A processor that senses internal and external stimuli is fundamental in all affective agent architectures. In Blumberg's Silas architecture (Blumberg 1996), it is called the Sensory System and contains a number of smaller monitors called *Releasing Mechanisms* that determine how input will be handled. In El-Nasr's PETEEI architecture (El-Nasr 1998), external stimuli are handled by the *Cognitive Process* module and internal stimuli are handled by the *Learning* and *Emotional Process* modules. In Reilly's Emotional Agents (Reilly 1996), stimuli are first processed in the *Emotion Generation* module. In Velasquez's EBC Framework (see Section 2.4.3) sensory input is handled by the

Perceptual System. Whatever the processor or mechanism is called, a sensory input processor is used by all affective agent architectures to sense the surrounding environment and the agent's internal state. Having done this, the processor activates other processes within the agent to determine how the input will affect the agent.

In EMAI's Sensory Processor, a goal such as EAT[4] may be triggered by a threshold value of the agent's Internal State Register that represents *hunger*. In a non-active state the agent will become increasingly *hungry* at a constant rate in response to signals from the Cyclical Drive mechanism within the Motivation Drive Generator. At a particular threshold value of this register the Homeostatic Drive mechanism will send a sensory signal of *hunger* to the Sensory Processor. In turn, the Sensory Processor will determine which goals to activate in order to return the agent to a less *hungry* state.

In other cases, external sensory data from the environment, such as the presence of noxious stimuli including *a loud noise*, *pain*, *heat* and so on, can also affect goals of an agent. On receiving external sensory data, the Sensory Processor determines which goals should be activated and updates the Internal State Registers appropriately. For example, an external data reading of a drop in atmospheric temperature may also reduce the agent's *body temperature* register's value.

The goals identified and activated by the Sensory Processor are quite abstract and are passed to the Event Space Generator in the Constructive Area for further processing.

3.3.4 Emotional State and Bias Registers

As well as the Internal State Registers that may define internal states such as the *hunger*, *tiredness* or temperature of an agent, the Knowledge Area also stores another set of attributes that collectively define the agent's emotional state or *mood*. The inspiration for including a mood value in the EMAI architecture comes from the work of Leuba and Lucas (Leuba and Lucas 1945) who concluded through experimentation that an individual's view of the world varies from moment to moment and is influenced by the individual's emotional state (see Section 2.2.2). Therefore, a mood value is stored in the

[4] Throughout this dissertation, goals will be typeset in uppercase.

EMAI architecture and used when assessing the external and internal stimuli processed by the agent.

This dissertation introduces a unique way of representing an affective agent's emotional state. Until now architectures for affective agents have determined an agent's emotional state by assessing the state of individual emotion variables (PETEEI (El-Nasr 1998), Silas (Blumberg 1996), Emotional Agents (Reilly 1996)) where each variable represents a pure emotion (e.g. *happiness*, *sadness*, *anger*). In the EMAI architecture the mood of the agent is stored separately from the agent's emotional module.

The attributes that define an EMAI agent's mood and emotional state are *pleasantness, anticipated effort, certainty, attentional activity, responsibility* and *control* and are known as *appraisal dimensions.* Although this will be discussed at length in Chapter 9, the mood of an EMAI agent is the result of an analysis made when the agent executes an event.

All events are assessed by the same six attributes listed above. The agent's mood is the summing of all the attitudes the agent has thus far collected by executing events. Along with the agent's mood, an Emotional Bias Register is kept. This register defines the agent's preferred emotional state and is used later by the Intention Generator in the Deliberate Area to schedule and order the agent's intentions.

Not only are events as a whole appraised, but each action, object, time and context in the knowledge base is also assigned a set of the six appraisal dimensions that must be present and linked to all event elements. Appraisal is an assessment the agent places on each concept type in the Type and Relation Hierarchies of the Ontology. They are a set of summative values the agent keeps and updates each time it performs a behaviour described by an event. Each element involved in an event, be it action, object, time or context, is evaluated by the agent under the six appraisal dimensions listed above. How an element is involved with the agent during the execution of a behaviour determines how the agent rates it in terms of each category. The collection of these appraisals forms the basis of the Affective Space.

3.3.5 Affective Space

Chapter 8 provides the background and a complete elucidation of the Affective Space component. Suffice it to say that this component is the *heart* and *soul* of the agent. The Affective Space is a six-dimensional space defined by the six appraisal dimensions. The

Affective Space, based on the psychological model of Smith and Ellsworth (Smith and Ellsworth 1985), defines 15 emotions (*happiness, sadness, anger, boredom, challenge, hope, fear, interest, contempt, disgust, frustration, surprise, pride, shame* and *guilt*) with respect to the dimensions of *pleasantness*, *responsibility*, *effort*, *certainty*, *attention* and *control*. Each appraisal dimension is used to produce a six coordinate point that defines an individual emotional state. The location of these emotional states can be manipulated to produce different personalities for the agent. This will also be discussed in Chapter 9.

The Affective Space replaces the individual emotional state variables that are used in other affective agent architectures (PETEEI (El-Nasr 1998), Em (Reilly 1996) and Dog World agents (Padgham et al 1997)). The benefits of including a continuous emotional state space with respect to discrete emotional state values is that the Affective Space eliminates the need for the significant coordination between emotional states, complex emotional blending formulae and rules and the emotional decay rate functions implemented in the other affective agent architectures. These benefits are further explored in Chapter 8. As the agent's mood is stored as six attributes each representing a value in each of the appraisal dimensions of the Affective Space, the agent's mood can be compared (by a distance measure) to any pure emotion (for example, *happiness* or *sadness*) in the Affective Space. The pure emotion that is physically closest to the agent's emotional state is used to describe the agent's mood.

The Affective Space is not only used in determining the agent's emotional state. As will be explained in Chapter 9, the Affective Space is used primarily in the agent's decision-making mechanism to make decisions about event execution using emotions. This process occurs in the Deliberate Area. However, before the agent can make any decisions about which events to execute in order to satisfy its goals, a list of events must be created. This process occurs in the Constructive Area.

3.4 The Constructive Area

The Constructive Area receives abstract goals and generic events from the Knowledge Area. These events form plans that the agent knows if successfully executed, will satisfy the goals that have been triggered by the Motivational Drive Generator via the Sensory Processor. The Constructive Area also receives information about the concept types in the Ontology that are used to construct specific events the agent needs to perform in order to satisfy its goals.

Goals received by the Constructive Area may be general or specific depending on the level of goal that has been activated. Rather than a goal with a plan that states generally that *some action should be taken toward something in someplace at sometime*, the Constructive Area specialises this to more exact terms that specifically define an event plan such as *the cooking of the cake will take place in the kitchen at noon*. This task, performed by the Constructive Area, takes the generalised event plans and constructs specialised plans using concepts in the Type and Relation Hierarchies of the Ontology.

3.4.1 Event Space Generator

The Event Space Generator receives abstract goals from the Sensory Processor and detailed event and planning information from the Ontology. With this information, the Event Space Generator constructs a series of specific competing events and associated plans that can be executed in order to satisfy the agent's goals. This process is examined in Chapter 7. The competing event plans are collectively referred to as the *event space*.

In Blumberg's architecture for Silas (Blumberg 1996) plans also compete to become the agent's behaviour. Each behaviour in the Silas architecture is contained within an autonomous object called a *Behaviour*. The strength of a Behaviour is determined by any associated Releasing Mechanisms (see Section 3.3.3) that have activated it. Events in the PETEEI (El-Nasr 1998) seem to be much less complex. The actions of the agent are based on a simple *stimulus-reflex* process where events occur as the direct result of an external stimulus. Unlike these and other previous affective architectures (EBC Framework (Velasquez 1999), Emotional Agents (Reilly 1996)), the EMAI architecture can produce complex event plans and dynamically create new ones. Rather than generating just one behaviour to perform in order to satisfy its goals, the agent can create multiple plans to satisfy just one goal. This creates a series of competing event plans. The competing event plans within the event space create a decision-making situation for the agent. The agent must select an event plan from the event space to execute. This selection process is performed in the Deliberate Area, where the agent makes a decision about which event plan to execute based on how it *feels* about the plan. The agent will execute the most *liked* plans first. Thus the decision making process is based on emotions. However, before this type of decision-making can take place, the agent must assess each event plan and determine how each event will affect the agent's emotional state. This process occurs in the Deliberate Area.

3.5 The Deliberate Area

The Deliberate Area accepts input in the form of an event space from the Constructive Area. The purpose of the Intention Generator in the Deliberate Area is to order, prioritise and schedule events. The events in the agent's schedule become the outward behaviour of the agent via the Behaviour Actuator. The Deliberate Area creates a list of tasks for the agent to perform called the agent's *schedule*. This schedule is dynamic, continually being updated and reordered as new event spaces arrive from the Constructive Area. This process is performed by the *Intention Generator*.

3.5.1 Intention Generator

As event spaces arrive in the Deliberate Area, the Intention Generator takes each collection of plans and orders it in the agent's schedule based on a two-fold ordering process. The primary ordering of the plans by the Intention Generator is based on goal priority. When a goal is activated or triggered by the Sensory Processor an appropriate event space is created in the Constructive Area. If a goal that has already been triggered is triggered again, the associated event space is given a priority rating. The more often a goal is triggered, the higher the agent's priority to achieve that goal. Different event spaces will have different levels of priority based on the number of times the associated goal has been triggered. The Intention Generator uses this priority to order the event spaces within the agent's schedule.

This type of primary event scheduling also occurs in other affective agent architectures including Silas (Blumberg 1996), PETEEI (El-Nasr 1998) and EBC Framework (Velasquez 1999) where a priority is placed on a behaviour, based on the agent's motivational state. However, this is where the similarity between the EMAI's behaviour scheduling and previous affective architectures ends. As mentioned previously, the EMAI is not only capable of deducing one plan to satisfy a goal, but it can also infer multiple and mutually exclusive plans that can satisfy just one goal. This collection of plans constitutes a goal's event space.

Because each event space in the agent's schedule may consist of one or more event plans, these plans need to be ordered within the event space. The Intention Generator orders the events within an event space using an emotional sequence, from most desirable to least desirable. The Emotional State Bias Register determines what constitutes a desirable event. The most desirable of event plans are ones that, when executed, would produce emotions closest to the value of the Emotional State Bias Register. As is the case with the agent's mood, each event in the event space also has an associated set of values for each of the six appraisal dimensions in the Affective Space. The agent assesses each event and assigns values to each of these dimensions. The agent uses the six appraisal values for an event, the agent's mood and the Affective Space to interpolate the emotional state that an event would evoke in the agent if the agent were to execute the event. This process is examined in detail in Chapter 9. The Intention Generator orders the event plans, within the event spaces, using this interpolated emotional state. Events that will produce the most desirable emotional state within the agent are ordered first and least desirable plans are placed last within each priority ordering of the event spaces.

Once event spaces have been ordered, the agent can attempt to perform plans from them. This processing occurs in the Behaviour Actuator.

3.5.2 Behaviour Actuator

Before the agent translates an event into behaviour, checks are performed to ensure the agent's physical state and the environmental states are adequate for the performance of the task. The Behaviour Actuator performs this task which includes considerations such as prerequisite and co-requisite tasks being performed and sufficient simulated resource availability. For example, the agent cannot perform the goal EAT if there is no food present in the environment and the agent cannot perform the goal PREPARE FOOD if the task for the goal FIND FOOD has not been performed. The following is an example of the Deliberate Area at work.

Assume the agent has been programmed with the beliefs of a pet dog. The agent has three Internal State Registers; Hunger, Tiredness and Activity, which in turn will activate three appropriate goals when their threshold values are reached. These are EAT, SLEEP and PLAY. Assume, for this example, the EAT goal has been triggered three times, the SLEEP goal has been triggered twice and the PLAY goal has been triggered once. Each goal evokes an associated event space and subsequent events. The EAT goal evokes events for EAT BONE, EAT DOG FOOD and EAT BISCUIT. The SLEEP goal evokes SLEEP ON BED and SLEEP OUTSIDE and the PLAY goal rouses the events CHASE BALL and DIG HOLE. Given that the agent is in a *happy* mood and has an Emotional

State Bias Register that programs the agent to prefer *happy* events, the Intention Generator would create a schedule thus:

```
EAT BISCUIT: happiness
EAT DOG FOOD: interest
EAT BONE: boredom
SLEEP ON BED: happiness
SLEEP OUTSIDE: frustration
DIG HOLE: pride
CHASE BALL: interest
```

where the emotion assigned to each event is shown. The events are first ordered in their respective event space priority order, in this case EAT then SLEEP and finally PLAY followed by an ordering within each event space based on the events that will keep the agent closest to a *happy* mood being first. If the agent is successful in achieving the first task, in this case the event EAT BISCUIT, the value on the Internal State Register that caused the triggering of the EAT goal will be reduced. The reduction amount is related to the attributes of the concept involved in the event. For example, the EAT BISCUIT event would reduce the *hunger* register by less than the EAT DOG FOOD event as the objects involved (biscuit and dog food) have differing *hunger* reducing capacities. If the Internal State Register's value is reduced below the threshold (signifying the agent is no longer *hungry*) the EAT goal will have been satisfied and any remaining events that were evoked by the EAT goal will be removed from the list as they are no longer needed. The resulting schedule would be**:**

```
SLEEP ON BED: happiness
SLEEP OUTSIDE: frustration
DIG HOLE: pride
CHASE BALL: interest
```

If the successful event reduced the Internal State Register but not enough to reduce its value to below the threshold, the associated goal would still be active and the remaining events from the respective event space would remain in the schedule. Furthermore, if the agent could not successfully complete the task for EAT BISCUIT it would discard this task from the list. The EAT goal would still be activated and the agent would try the next task in the list, namely EAT DOG FOOD.

Other affective agent architectures also include a process or mechanism that turns the agent's scheduled behaviours into actual outward behaviours that can be seen by an onlooker. In Silas (Blumberg 1996) the *Motor System* performs this task. As Silas is a three-dimensional animated character that exists in a virtual world, the Motor System controls the agent's animation sequences that convey a particular behaviour. In the EBC Framework (Velasquez 1999) created for implementation in robotic devices, outward behaviours are also controlled by a *Motor System*, however, in this case the Motor System controls the robotic devices physical circuitry to make it move. In PETEEI (El-Nasr 1998) although no specific module has been defined for activating behaviour, an appropriate set of animation sequences are played as the result of a production rule assessment of the agent's physical and emotional state. A similar mechanism or process for generating outward behaviour in an affective agent will be found in all affective agent architectures. Without one, the agent would not be able to act out events and plans that could satisfy its motivated goals.

3.6 Summary

Each area in the EMAI architecture works to create an agent that can choose its behavioural options based on assessments using emotional state values. The agent can be programmed to simulate any set of characteristics or behaviours by altering the structure of the Goal, Type and Relation Hierarchies and the number of Internal State Registers in the Motivational Drive Generator. The personality of the agent is defined by the structure of the Affective Space. All other components work together in order to generate the outward behaviour of the agent and to update its belief system. While the EMAI architecture includes a number of fundamental mechanisms that are present in other affective agent architectures, it also introduces a number of new and unique concepts and processes that enhance the emotion processing capabilities of an EMAI agent.

The purpose of this chapter has been to provide a holistic examination of the EMAI architecture, how it compares to other affective agent architectures and to describe how the components separately process information, and also how they cooperate to create an agent with affective decision-making abilities. Many of the components of the agent architecture have intricate internal operations not discussed in this chapter. The purpose of the following chapters is to examine the analytical and psychological background for the major components of the architecture. The structure and functionality of each component is discussed in depth. The first of these chapters examines the processes and knowledge representation of EMAI's Ontology.

4. Structuring Knowledge

e know a subject ourselves, or we know where we can find information upon it.
- Samual Johnson (1709 – 1785)

4.1 Introduction

Ontology is the study of categorising elements within a specific domain. The result of such a study creates an *Ontology*. Sowa (2001) defines an Ontology as:

> "…a catalog of the types of things that are assumed to exist in a domain of *interest D* from the perspective of a person who uses a language *L* for the purpose of talking about *D*. The types in the Ontology represent the *predicates*, *word senses*, or *concept and relation types* of the language *L* when used to discuss topics in the domain *D*. An uninterpreted logic, such as predicate calculus, conceptual graphs, or KIF, is *ontologically neutral*. It imposes no constraints on the subject matter or the way the subject may be characterized. By itself, logic says nothing about anything, but the combination of logic with an Ontology provides a language that can express relationships about the entities in the domain of *interest*."

The EMAI Ontology organises elements from the agent's realm, using conceptual graph theory (Sowa 1984, 2001) and extends the formalism using the new concepts: the Goal Hierarchy (Baillie et al. 2000c), Event Graphs and Activity Digraphs. Event Graphs and Activity Digraphs are new abstractions defined in this dissertation for the purpose of representing events and plans. The Ontology presents the knowledge that an EMAI agent possesses about its environment.

In this chapter, the EMAI Ontology is described. An EMAI agent's Ontology consists of sub-components that work collectively to represent the agent's knowledge base. The Ontology consists of a Type Hierarchy, Relation Hierarchy, Goal Hierarchy, Event Graphs, Activity Digraphs, and a set of primitive Canonical Graphs from which all other graphs needed to represent the full extent of the agent's knowledge can be derived. In the following section, the knowledge representation used in the EMAI architecture is described.

4.2 Knowledge Representation

The knowledge representation scheme of conceptual graphs (CG) (Sowa 1984, 2001) has been chosen as it provides the necessary vehicle for the storage of the type of deep knowledge required by the agent using existing structure for conceptual relationships and knowledge processing through canonical graph formation rules. Conceptual graphs also form part of a semantic network thus giving flexibility for the agent to later be integrated with other perceptual input, procedures and motor mechanisms. An introduction to conceptual graph theory can be found in Appendix C of this dissertation. In addition to the knowledge structures of conceptual graphs, this dissertation also introduces a structure called the Goal Hierarchy. Furthermore, it introduces two forms of abstraction called the Event Graph and Activity Digraph to represent events and plans, respectively. Collectively, these structures are used to represent the EMAI Ontology.

Note the Type Hierarchy and Relation Hierarchy have been defined by Sowa (1984), but re-produced here for completion and ease of understanding for the rest of the abstractions.

4.2.1 Type Hierarchy

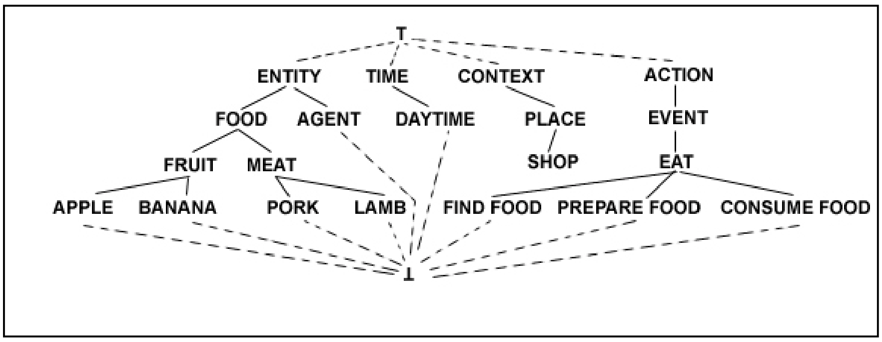
The Type Hierarchy in the EMAI Ontology contains all the concept types known to the agent in a partially-ordered set arranged as supertypes and subtypes. Sowa defines a Type Hierarchy as follows.

Definition: A *Type Hierarchy* is a partially-ordered set *T* whose elements are called *type labels*. Each type label in *T* is specified as *primitive* or *defined*.

- For any concept *c*, the type of *c* is either a type label in *T* or a monadic lambda expression.
- The Type Hierarchy *T* contains two primitive type labels Entity, called the *universal type*, and Absurdity, called the *absurd type*. The symbol $\top$ is synonymous with Entity, and the symbol $\perp$ is synonymous with Absurdity.
- For every defined type label, there is a monadic lambda expression, called its *definition*.

- A defined type label and its definition are interchangeable: in any position where one may occur, the other may replace it.
- The partial ordering over *T* is determined by the *subtype* relation, represented by the characters "≤" for *subtype*, "<" for *proper subtype*, "≥" for *supertype*, and ">" for *proper supertype*. If t is any type label, $\top \geq$ t and t $\geq \perp$; in particular, $\top > \perp$.
- The partial ordering of type labels must be consistent with the rules of inference. (Sowa 2001).

The most abstract concepts appear in the top of the hierarchy. As the hierarchy decomposes into lower levels the concepts become more specific. Figure 4.1 illustrates an example of a simple Type Hierarchy. A concept type such as ENTITY is abstract and can be decomposed into the mores specific concept types of APPLE, BANANA, PORK and LAMB. It can be stated that APPLE is a subtype of ENTITY and ENTITY is a supertype of APPLE.



4.2.2 Relation Hierarchy

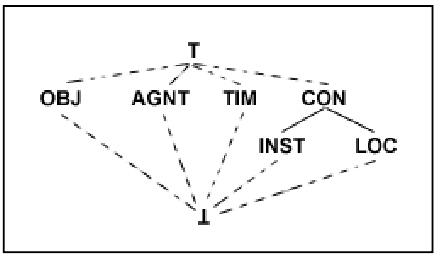
The Relation Hierarchy in the EMAI Ontology defines the supertype and subtype relationships that exist between the conceptual relations that are used by an EMAI agent to define the roles amid the concepts in the Type Hierarchy. Sowa defines a Relation Hierarchy as follows.

Definition: A *Relation Hierarchy* is a partially-ordered set *R* whose elements are called *relation labels*. Each relation label is specified as *primitive* or *defined*, but not both.

- For every relation label in *R*, there is a nonnegative integer *n* called its *valence*.

- For every *n*-adic conceptual relation *r*, the type of *r* is either a relation label in *R* of valence *n* or an *n*-adic lambda expression.
- For every defined relation label of valence *n*, there is exactly one *n*-adic lambda expression, called its *definition*.
- A defined relation label and its definition are interchangeable: in any position of a CG where one may occur, it may be replaced by the other.
- The partial ordering over *R* is determined by the *subtype* relation, with the symbols ≤ for *subtype*, < for *proper subtype*, ≥ for *supertype*, and > for *proper supertype*.
- The partial ordering of relation labels must be consistent with the rules of inference (see Sowa 2001).
- If *r* is an *n*-adic relation label, s is an m-adic relation label, and *n* is not equal to m, then none of the following is true: r<s, r>s, r=s. (Sowa 2001)

Figure 4.2 displays an example of a simple Relation Hierarchy. These relations are used to join the concepts of the Type Hierarchy into conceptual graphs. For example, the relation AGNT can be used to describe how an ENTITY concept can be the instigator of an ACTION concept. A relation such as CON is abstract and can be decomposed into more specific relations of INST and LOC. It can be said that INST is a sub-relation of CON and CON is a super-relation of INST.



4.2.3 Goal Hierarchy

Human behaviour is goal-orientated (Furth 1987, Rao and Georgeff 1995). This property allows people to recover from failed plans and to recognise and create new ones. Koestler (1967) cites MacLean (1958) and references the two basic motivational drives (or goals) in human behaviour; *self-preservation* and *preservation of the species*. These goals, according to Koestler, are, by instinct, built into all animals and are programmed into the reptilian and paleo-mammalian brain constructs described in MacLean's Theory in Section 2.2.1. From these two goals (U_1 and U_2), all other motivational goals (G_1 ... G_n) of human behaviour can be derived and are individual to the human for whom they are defined. Expression 4.1 attempts to formalise this relationship[5].

$$\{G_1 \cdots G_n\} \subseteq \{U_1 \cup U_2\} \qquad 4.1$$

Thus, an EMAI agent may posses a set of goals at any one point in time. Each of these goals has a particular level of priority. The level of priority placed on a goal indicates the level of desire the agent has for carrying out activities that fulfill that goal. The greater the desire to fulfill a goal the greater the chance the agent will be working towards fulfilling that goal. Besides the desire to fulfill the goal, the ability of an agent to carry out a plan that will satisfy the goal is also dependent on its beliefs resulting from the environment and its internal states.

The Goal Hierarchy contains the entire set of goals that are used to generate tasks within the agent. In order to emulate human-like behaviour the agent is driven by this set of goals. In a similar structure to a conceptual graph Type Hierarchy, the goals within the hierarchy have supergoals and subgoals. The subgoals inherit properties from the supergoals and the supergoals are more abstract than the subgoals. The goals that are programmed into the Goal Hierarchy reflect the type of agent the architecture is being used to simulate. For example, if the agent were to represent an animal, then it would be programmed with a Goal Hierarchy suitable for simulating the behaviour of a goal-orientated animal. This Goal Hierarchy may look similar to Figure 4.3.

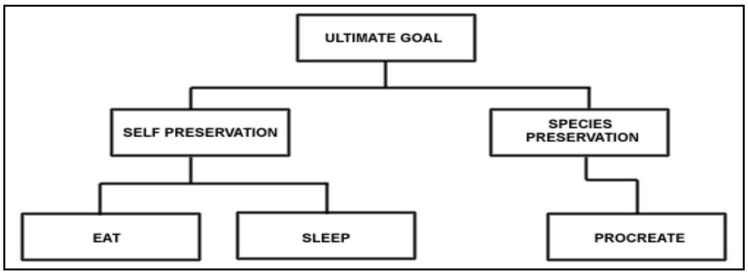


Figure 4.3 A Simple Goal Hierarchy for an Animal

The goals represented in the Goal Hierarchy of Figure 4.3 are quite abstract and could be representative of any living creature. The idea behind the Goal Hierarchy is that it can be representative of the agent's goals on any level. The more this Goal Hierarchy is decomposed into primitive goals the closer it comes to representing specific events. At the

---

[5] This occurs as the two ultimate goals have subgoals, which in turn have their own subgoals and so forth. The set of all these goals (including subgoals) constitutes all of an individual's motivational goals.

most primitive level of the Goal Hierarchy there exists a partially-ordered set of goals that are synonymous with the processes or events and plans that can be performed to satisfy a goal. By decomposing any goal in the hierarchy into primitives, the agent can identify a plan of events that can be performed to satisfy the initial goal. Figure 4.4 illustrates part of a Goal Hierarchy. This figure shows a partial Goal Hierarchy displaying the goal EAT decomposed into three partial subgoals; FIND FOOD, PREPARE FOOD and CONSUME FOOD.

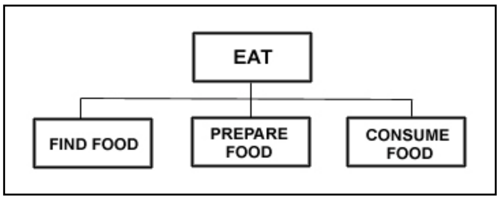


Figure 4.4 Goal Hierarchy for EAT

As stated earlier, the goals in the Goal Hierarchy of the EMAI Ontology are either abstract or primitive. An abstract goal can be broken down into subgoals (of which some will be abstract, while others may be primitive). Subgoals are either AND-ed or OR-ed. A set of AND-ed subgoals indicates that all of the subgoals need to be successfully achieved to satisfy the supergoal. A set of OR-ed subgoals indicates that at least one of the sub-goals needs to be successfully achieved to satisfy the supergoal. The AND and OR are used to indicate process inclusion. A new construct called an Event Graph is used to indicate process ordering (see Section 4.3.1). Each goal corresponds to an associated event that can be carried out in order to achieve (or partially achieve) the goal. The definition of Goal Hierarchy can be stated as follows:

Definition: A Goal Hierarchy is a partially-ordered set G whose elements are called *goals*. Each goal in G is specified as *abstract* or *primitive*. Considering the symbol '≤' indicates ordering and g and h are goals, G is such that:

- If $g \leq h$, then g is called the *subgoal* of h; and h is called the supergoal of g, written $h \geq g$.
- The Goal Hierarchy G contains one ultimate goal, called the *ultimate goal (U)*, and two *subultimate* goals called *self-preservation* (U_1) and *species-preservation* (U_2).
- The partial ordering over G is determined by the *subgoal* relation, represented by the characters "≤" for *subgoal*, "<" for *proper subgoal*, "≥" for *supergoal*, and ">" for *proper supergoal*. If g is any goal, $U \geq g$ and $g \leq U$.
- For any goal g where $g \neq U$, $g \neq U_1$ and $g \neq U_2$ then g will be a proper subgoal of U_1 and/or U_2.
- If a goal g has a set of subgoals S such that $S = \{\phi\}$, g is said to be *primitive*.

- The subgoals, $\{s_1..s_n\}$, of the set of subgoals, S, for a goal, g, can be arranged in sentences of propositional logic where the logical connective AND is represented by the character "Λ" and OR is presented by the character "V". A sentence such as $(s_1 \, Λ \, s_2 \, Λ \, s_3) \Rightarrow g$ indicates that of all the subgoals in S, only s_1, s_2 and s_3 need to be satisfied in order to satisfy the g. A sentence such as $(s_1 \; V \; s_2 \; V \; s_3) \Rightarrow g$ indicates that only one of the subgoals (s_1,s_2 or s_3) needs to be satisfied in order to satisfy the g.
- The subgoals $S=\{s_1...s_n\}$, of the goal, g, can be arranged into subsets of S, $A_1...A_m$, such that all goals in A_i are joined in a sentence of propositional logic using Λ as the logical connective and $A_1 V A_2 V ... V A_n$.

4.2.4 Relationship Between Type Hierarchy and Goal Hierarchy

A typical Type Hierarchy is shown in Figure 4.5. In this structure, there is type EVENT as a subtype of type ACTION, and FIND, PREPARE and CONSUME as subtypes of EVENT, and so on, as shown at the bottom of this figure as an expansion of the FIND FOOD, PREPARE FOOD and CONSUME FOOD subtypes[6].

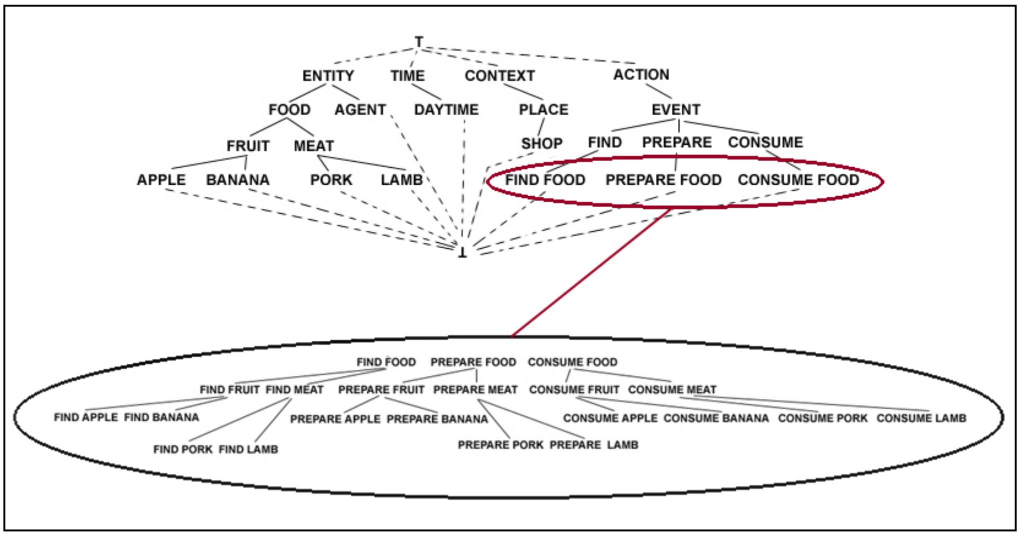
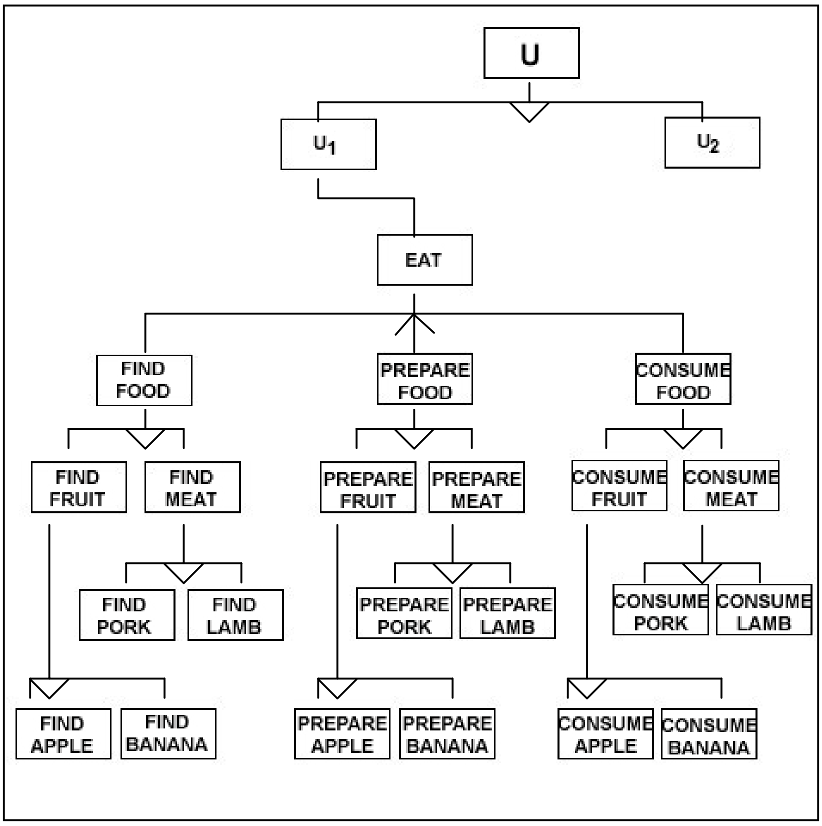


Figure 4.5 Type Hierarchy Showing Event Types with the FIND FOOD, PREPARE FOOD and CONSUME FOOD Subtypes Shown Expanded in the Oval Section

It is fundamental to the understanding of the definition of this EMAI Ontology to comprehend the relationship between the Type Hierarchy and the Goal Hierarchy. In the

[6] Combinatorial explosion is an issue with this representation scheme. However, in this dissertation, the emphasis of the study is in exploring emotion in decision making, and as such, combinatorial explosion is not the focus of this research.

Type Hierarchy, all ENTITYs are *objects types*[7], while all subtypes of EVENTs are *action types*. *Actions* are performed against the *objects*. In the Goal Hierarchy, all the *action types* from the Type Hierarchy can form the sub-goals under the *subultimate* goals called *self-preservation* *(U_1)* and *species-preservation* *(U_2)*, as shown in Figure 4.6.



4.6 An Extended Goal Hierarchy

Each goal corresponds to an associated event that can be carried out in order to achieve (or partially achieve) the goal. In the Goal Hierarchy in Figure 4.6, there are sub-goals for FIND-FOOD, PREPARE-FOOD and CONSUME-FOOD AND-ed, while FIND-FRUIT and FIND-MEAT is OR-ed. Similarly, FIND-PORK and FIND-LAMB are OR-ed, and so on. This Goal Hierarchy simply states that to satisfy the EAT goal, the subgoals: FIND-FOOD, PREPARE-FOOD, and EAT-FOOD have to be satisfied. To satisfy FIND-FOOD, either FIND-FRUIT or FIND-MEAT has to be satisfied, and so on.

[7] Note, in other knowledge bases not all entities are classified as object types. However, in this representation it is unnecessary to classify them as other types. Refer to Section 6.4 for further details on entity classifications.

An event can be represented by an abstraction in the form of a conceptual graph. These events can be ordered in a temporal manner to form plans. Plans are also represented by an abstraction in the form of conceptual graph. In the next section, these two forms of abstraction are described and defined.

4.3 Event and Plan Definitions

4.3.1 Event Graph

Each goal in the Goal Hierarchy has an associated event (with the same name) that can be performed by the agent to satisfy or partially satisfy the goal. An event is represented in the agent's knowledge base as a conceptual graph in the form dictated by the Theory of Reasoned Action (Petty and Cacioppo 1996) (see Chapter 5). This theory defines a behavioural event as having four elements; action, object, time and context. A generic event is represented by the conceptual graph shown in Figure 4.7 where an event is defined as *some entity performing an action on some entity at some time and in some context.* Associated with this graph are three sets of conceptual graphs that represent the precondition, postcondition and delete graphs that are used during event reasoning (otherwise referred to in this dissertation as *event execution*).

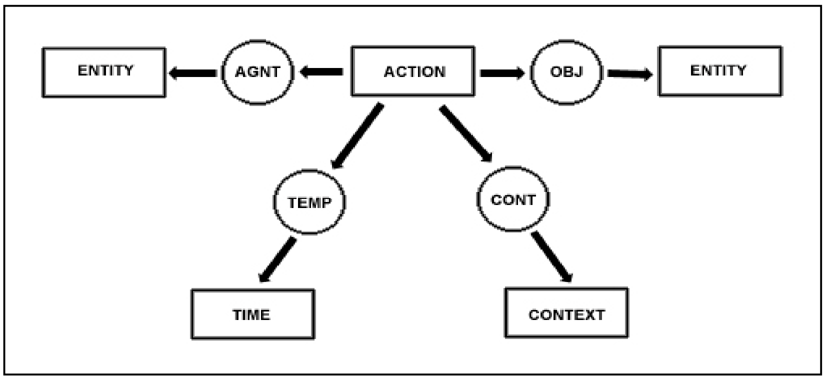


Figure 4.7 A Conceptual Graph Representation of an Abstract Event

This can also be represented in linear form as:

```
[ACTION] –
          (OBJ) -> [ENTITY]
          (AGNT) -> [ENTITY]
          (TIM) -> [TIME]
          (CON) -> [CONTEXT]
```

In other words, a goal a in the Goal Hierarchy G, at any level of abstract, n, written as $G_{a,n}$ is represented by a conceptual graphs abstraction called the Event Graph. Each Event Graph has associated with it a set of graphs that represent the precondition, postcondition and delete list. An Event Graph is defined as follows.

Definition: An *Event Graph* is an *n-adic* abstraction, $\lambda\ v_{1..n},\ w_{1..m},\ x_{1..o}\ u$, where u is the conceptual graphs representing the Event Graph, $v_{1..n}$ *represents* a set of n conceptual graphs that represents the pre-condition that must be satisfied before this event can be executed, $w_{1..m}$ represents a set of m conceptual graphs that represents the post-conditions that will result from executing this event, and $x_{1..o}$ represents a set of o conceptual graphs that will be removed from the memory, after execution of this event.

In other words, the preconditions denote the knowledge that must exist before the event is executed, the postconditions denote the knowledge that is created as a result of executing the event and the delete list represents knowledge that is no longer needed after the event has been executed. For example, the Event Graph for the FIND-FOOD could be represented as in Figure 4.8.

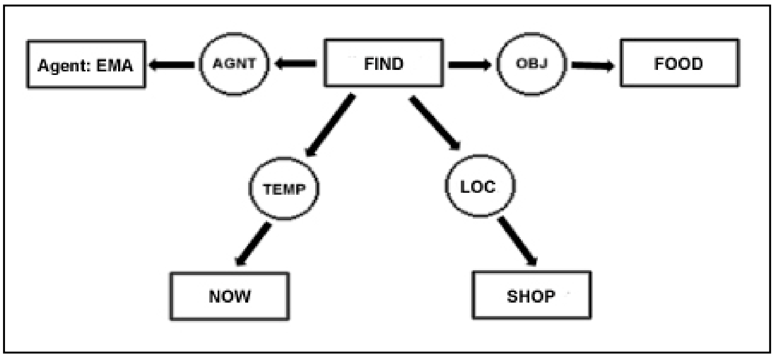


Figure 4.8 Event Graph Representation of FIND FOOD Event

In linear form, the Event Graph in Figure 4.8 is written as follows:

```
[FIND] –
      (OBJ) -> [FOOD]
      (AGNT) -> [AGENT: EMAI]
      (TIM) -> [TIME: #now]
      (LOC) -> [BACKYARD]
```

where the corresponding precondition, postcondition and delete sets of conceptual graphs are as shown below:

```
Precondition:   {[AGENT: EMAI]<-(AGNT)<-[ACCESS]->(OBJ)->[FOOD]}
Postcondition:  {[AGENT: EMAI]<-(POSS)->[FOOD]}
Delete list:    {[AGENT: EMAI]<-(AGNT)<-[ACCESS]->(OBJ)->[FOOD]}
```

To facilitate knowledge passing during the graph execution process, it is essential that the graph processor utilises *coreference links* within a graph, and *line of identity* between graphs. Definitions of these concepts are given below.

Definition: *Coreference links* show which concepts refer to the same entities. In linear form representation of conceptual graphs, co-reference links are represented using variables.

For example, consider the following conceptual graph:

```
[FIND] –
      (OBJ) -> [FOOD]
      (AGNT) -> [AGENT] -> (LOC) -> [SHOP:a*]
      (TIM) -> [TIME: #now]
      (LOC) -> [?a]
```

where the location of the AGENT is identical to the location where the event FIND took place.

Definition: A line of identity is a connected, undirected graph g whose nodes are concepts and whose arcs are pairs of concepts, called co-reference links. No concept may belong to more than one line of identity. A concept a in graph g is said to dominate another concept b if there is a path $<a_1, a_2, ..., a_n>$ in g where $a=a_1$, $b=a_n$, and for each i, either a_i and a_{i+1} both occur in the same context or the context of a_i dominates the context of a_{i+1}.

For example, consider the following Event Graph, that is made up of conceptual graphs g_1, g_2, g_3, and g_4:

```
g1:  [FIND] –
           (OBJ) -> [FOOD:*a]
           (AGNT) -> [AGENT: *b] -> (LOC) -> [SHOP:*c]
           (TIM) -> [TIME: #now]
           (LOC) -> [?c]
g2:  [AGENT: *b]<-(AGNT)<-[ACCESS]->(OBJ)->[FOOD:*a]
g3:  [AGENT: *b]<-(POSS)->[FOOD:*a]
g4:  [AGENT: *b]<-(AGNT)<-[ACCESS]->(OBJ)->[FOOD:*a]
```

In the above Event Graph, the concept SHOP in g_1 is a co-reference (indicated by the variable *c. There exists a line of identity between all four graphs via the concepts FOOD and AGENT. So, during graph processing, if the generic concept [AGENT: *b] in g_1 is specialised to [AGENT: EMAI], the graph processor will automatically follow the coreference links and line of identity to specialize g_2, g_3 and g_4 as follows:

g_2: [AGENT: EMAI]<-(AGNT)<-[ACCESS]->(OBJ)->[FOOD:*a]

g_3: [AGENT: EMAI]<-(POSS)->[FOOD:*a]

g_4: [AGENT: EMAI]<-(AGNT)<-[ACCESS]->(OBJ)->[FOOD:*a]

As described above, each goal in the Goal Hierarchy has a corresponding Event Graph. The Event Graph describes the process that must take place to satisfy the goal. This process is called an event. The goals define the why of an agent's behaviour and the events define the how. They are so closely related that it is difficult in the English language to divide them. For example, the goal may be to *eat an apple* and the event will be *eating an apple*. For each event, there is a corresponding goal, whatever it may be. We cannot have the how without the why. Therefore, each goal in the Goal Hierarchy has a corresponding event at the same level of abstraction as the goal and a set of events on sub-levels the agent may choose from to achieve the goal. The events on subsequent levels are arranged into Activity Digraphs that form plans to satisfy the associated goal.

4.3.2 Activity Digraphs

Alone, the Goal Hierarchy does not include any plans that show how the subgoals interact to satisfy the parent goal. To this end, the Goal Hierarchy is accompanied by a set of Activity Digraphs that explain how goals combine to form plans. Activity Digraphs can be formed at various levels of abstraction. For example, at level 3, the Activity Digraph accompanying the Goal Hierarchy of Figure 4.6 might look like that in Figure 4.9. In an Activity Digraph (not to be confused with a conceptual graph) the vertices correspond to the events (represented by Event Graphs) that must be performed and an edge is drawn from vertex i to vertex j if event i must precede event j.

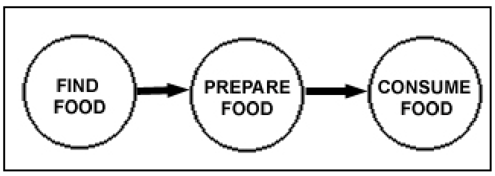


Figure 4.9 Activity Digraph for EAT

The plans in the Activity Digraphs need not be as simplistic as simple sequencing (for example, process repetition is allowed). Activity Digraphs are synonymous with Problem Maps described in (Lukose, 1992, 1996). They can represent any degree of complexity and can easily represent all the major control structures (such as FOR, WHILE, REPEAT and CASE). Each event in the plan will have a set of pre-conditions, co-conditions and post-conditions that need to be analysed by the agent before, during and after execution. Each event will alter the state of the agent or the environment during and after execution. The agent's processing of an Activity Digraph falls into the domain of plan reasoning and task analysis and will not be discussed in this dissertation. Suffice it to say that each event in the Activity Digraph is analysed by the agent using plan reasoning techniques (Allen, Kautz et al 1991) that as a result affect the state of the agent, the agent's environment and the agent's behaviour.

To construct a conceptual graph representation of an Activity Digraph, a partial temporal ordering of the Event Graphs is specified. The two temporal relation types that are used to specify the temporal ordering of the Event Graphs are *FINISH_BEFORE_START (FBS)* and *START_WHEN_START (SWS)* (Lukose 1996). An Activity Digraph can be defined as follows:

Definition: An *Activity Digraph* labeled t is defined as a n-adic abstraction $\lambda v_1,..,v_n$ $w_1,...,w_m$ u, where u is the conceptual graph representing the *activity digraph*, $v_1,..,v_n$ which is in turn the set of conceptual graphs representing the initial state, while $w_1,...,w_m$ is the set of graphs representing the final state.

This definition is analogous to the definition of *Problem Map* (Lukose 1993, 1996). The difference between Problem Map and Activity Digraph is that the Problem Maps are Executable Conceptual Structure (dynamic graphs), while the Activity Digraphs are not; they can only be reasoned upon (static graphs). The Activity Digraph in Figure 4.8 could be represented as:

```
[FIND-FOOD]->(FBS)->[PREPARE-FOOD]->(FBS)->[CONSUME-FOOD]
```

The initial state of the above Activity Digraph is the union of the pre-conditions of all the Event Graphs associated to the Activity Digraph. In the above example, they are the union of the preconditions of FIND-FOOD, PREPARE-FOOD, and CONSUME-FOOD. The final states are the union of all the post-condition graphs of the same Event Graphs.

4.4 Example of an EMAI Ontology

This section examines an example of a typical EMAI Ontology. Consider the Type Hierarchy, Relation Hierarchy and the Goal Hierarchy shown in Figures 4.10, 4.11, and 4.12, respectively. Note that for this example, only the relevant sections of the hierarchies are displayed.

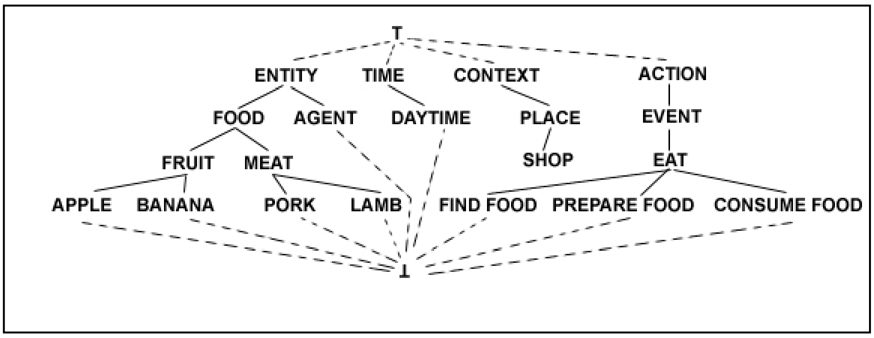


Figure 4.10 A Type Hierarchy

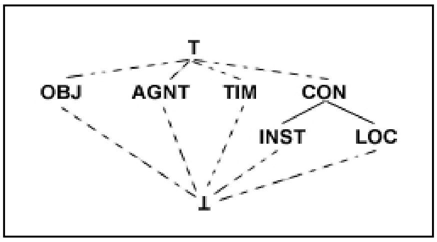


Figure 4.11 A Relation Hierarchy

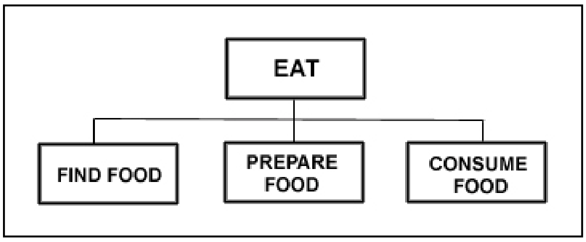


Figure 4.12 Goal Hierarchy for EAT

As described earlier, each goal in the Goal Hierarchy is associated with an Event Graph. In this example, the Event Graphs for FIND-FOOD, PREPARE-FOOD, and COMSUME-FOOD, are as depicted in Figures 4.13, 4.14, and 4.15, respectively[8]. Note the subgoals of EAT are AND-ed. This implies there is an Activity Digraph that links these

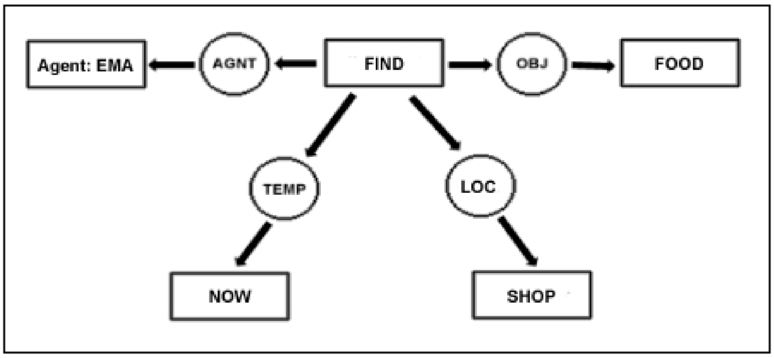


Figure 4.13 Event Graph Representation of FIND FOOD Event

Event Graphs to form the plan associated with the event EAT.

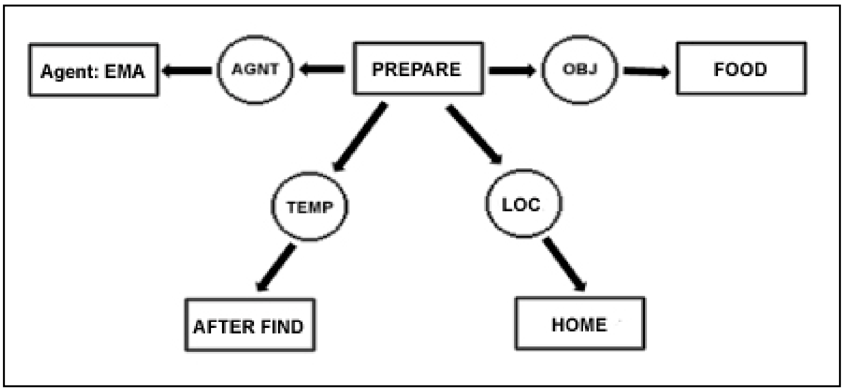


Figure 4.14 Event Graph Representation of PREPARE FOOD Event

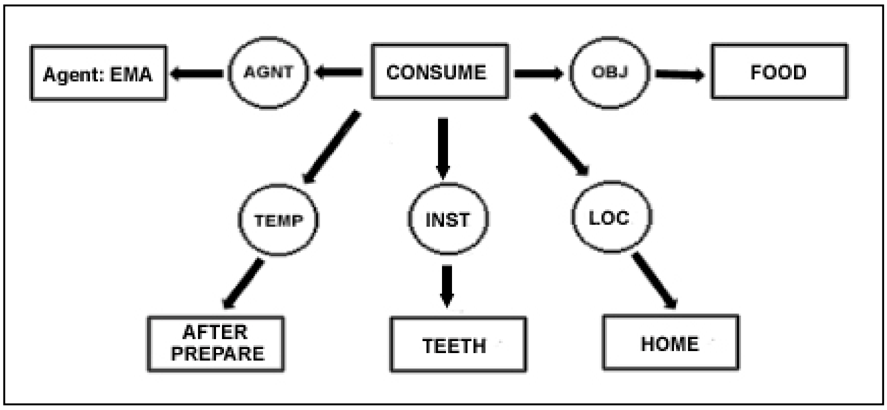


Figure 4.15 Event Graph Representation of CONSUME FOOD Event

The Activity Digraph that represents the plan for EAT is represented (as described earlier) using the temporal relation called FINISH-BEFORE-START (FBS), and can be written in linear form as shown below:

```
[FIND-FOOD] -> (FBS)->[PREPARE-FOOD] -> (FBS) -> [CONSUME-FOOD]
```

By expanding each of the above concepts with its corresponding Event Graphs, the Activity Digraph shown in Figure 4.16 is obtained.

[8] Notice the CON relation has been specialised from Figure 4.7 into the relations INST and LOC (see Figure 4.11) to illustrate how these relations might be implemented.

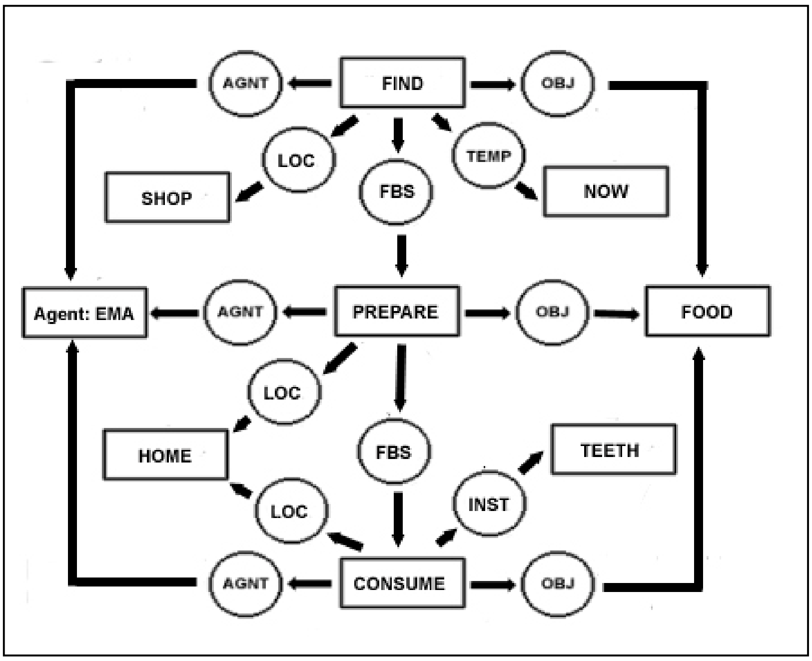


Figure 4.16 Activity Digraph Representation of EAT Plan

4.5 Summary

The EMAI Ontology achieves the depth of knowledge required by an affective agent by using conceptual graph theory (Sowa, 1984) and Event Graphs and Activity Digraphs based on the Executable Conceptual Structure constructs (Lukose, 1996). To these existing formalisms, a new structure, called the Goal Hierarchy, has been added to strengthen the architecture's representation of the knowledge in a goal-orientated agent. The Goal Hierarchy gives the agent a logical method for organising goals. Each goal in the hierarchy decomposes into a set of subgoals. Each subgoal has an associated event that forms a part of a plan that achieves a parent goal. An Event Graph in turn represents each event, and an Activity Digraph represents the plan.

The knowledge contained in the EMAI Ontology is necessary to develop the emotional behaviour of an agent. However, before an EMAI agent is compelled to perform a behaviour it must first *want* to do so. The internal processes that cause the agent to behave are contained within the Motivational Drive Generator. This, in collaboration with the

Sensory Processor, activates an EMAI agent's goals and begins the process of affective behaviour generation.

5. Motivation and Prioritisation

Obviously, for intelligence to function, it must be motivated by an affective power. A person will not ever solve a problem if the problem does not interest him.
- PIAGET

5.1 Introduction

Unlike machines, many factors drive human behaviour. Each human is an autonomous, socially interdependent entity. They are each motivated by many and varied needs and often do not act in a logical manner. This makes their behaviour unpredictable and beyond simple statistical extrapolation (Baillie and Lukose 2001a). In the struggle to define human behaviour, Koestler (1967) makes similar observations. He suggests the human organism is not merely a mechanical device, but it reacts to an ever-changing world and how it reacts to the world is based on its goals at the time. For example, at one instant, a person may have the intent to work towards achieving goal *A*, but at a later time the same person may have a completely different intent. This is one reason why it has recently become evident that in order to produce artificial agents with intelligent behaviour it is necessary to supply these agents with the ability to react to the dynamic world (environment) where they live; be it virtual or real (Rao and Georgeff 1992, Blumberg et al. 1996).

The EMAI architecture relies on the Knowledge Area's system of individual beliefs the agent holds about its environment and its physical and emotional states. These beliefs continually interact with an EMAI agent's desire to achieve a goal activated by the Sensory Processor and its ordered event plans and schedule of behaviour. Thus, the goal-orientated nature of the agent allows it to demonstrate planning and re-planning. The agent exhibits *reactive* and *deliberative* behaviours. It has a repository of *reactive* plans that it can use in various situations that occur within its environment, while at the same time, it also has a repository of *actions* that it is capable of performing. To achieve this, the agent must plan the sequence of actions to execute, in response to changes in the environment or itself. Following this, the agent can attempt to achieve its goal. The agent's beliefs are dynamic and may change with time, the agent's physical position within its environment, and its justification of beliefs. To simulate this kind of behaviour using the EMAI architecture, the mechanics of goal setting, goal prioritising and motivation are utilised to construct the agent's Knowledge Area.

This chapter examines the internal mechanisms of the EMAI's Motivational Drive Generator and takes a further look at the structure of the Goal Hierarchy in the EMAI architecture's Ontology. An illustration of how the motivational urges (represented by the corresponding Internal State Registers) influence the activation and prioritisation of an EMAI agent's goals is given. These processes within the Motivational Drive Generator, which have been inspired by human behaviour, are where the final behaviours of an EMAI agent originate.

This chapter begins by examining the structure of the EMAI Ontology's Goal Hierarchy (as presented in (Baillie and Lukose 1999)). This is followed by an illustration of the Motivational Drive Generator and Sensory Processor.

5.2 Motivating an EMAI Agent

Inspiration for the motivational and prioritising mechanisms of the EMAI agent have been taken from psychological literature. Many renowned psychologists (Freud (Rothgeb 1973), Piaget (Furth 1987), Toda (Padgham and Taylor 1997), and Koestler 1967) speak of this motivation in terms of urges. Toda (Padgham and Taylor 1997) identifies four motivational urges (emergency, biological, cognitive, and social). The combination of these urges and their intensity levels determine the behaviour of the individual and which goals are made high priority. Section 2.3.2 identifies other drives and the cyclic and concepts used in contemporary agent architectures (El-Nasr 1998, Balkenius 1995, Padgham and Taylor 1997 and Canamero 1997). The EMAI architecture uses these concepts to emulate urges within the Motivational Drive Generator.

5.2.1 Revisiting the Goal Hierarchy

Recall from Chapter 4 two basic motivational goals in human behaviour; self-preservation (U_1) and preservation of the species (U_2). From these two goals (U_1 and U_2), all other motivational goals (G_1 … G_n) of human behaviour can be derived and are individualistic of the human for whom they are defining. Expression 5.1 formalises this relationship.

$$\{G_1 \cdots G_n\} \subseteq \{U_1 \cup U_2\} \tag{5.1}$$

When a goal becomes the focus of an agent's belief and the agent wants to satisfy that goal, each and every subgoal of that goal becomes active. A very simple example is shown in Figure 5.1. An Activity Digraph represents the inter-relationships between these subgoals, where the diagrammatic representation of an Activity Digraph is made up of vertices and edges as described in Expression 5.2:

$$D = (V, E) \qquad 5.2$$

where V is the set of goals that are the vertices of D, and E is the set of edges that connect these vertices. As can be seen in Figure 5.1, this agent has seven subgoals on *Level n+1*. When the agent's goal $G_{a,n}$ becomes active, so too will all the subgoals on *Level n+1*. In this case the agent can choose from Activity Digraph $D_{1,n+1}$, $D_{2,n+1}$ or $D_{3,n+1}$. The common subgoals of $G_{a,n}$ will appear in all Activity Digraphs $D_{x,n+1}$, where $1 \leq x \leq 3$. These subgoals are easily identified in the set produced by expression 5.3[9]:

$$\bigcap_{i=1}^{j} D_{j,n+1} \qquad 5.3$$

where j is the number of goals on *Level n+1*.

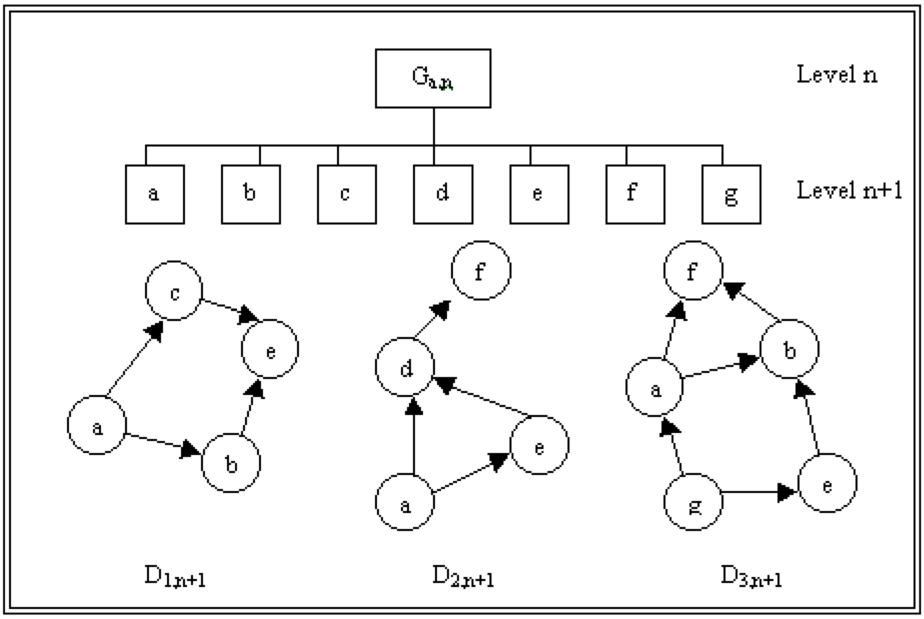


Figure 5.1 Hierarchy and Associated Activity Digraph for *Level n+1*

[9] These common subgoals are found through the intersection of each set of subgoals that constitute an Activity Digraph.

In the example shown in Figure 5.1, the agent has two common subgoals in the Activity Digraphs that satisfy $G_{a,n}$: *a* and *e*. These common subgoals will define the success factor for $G_{a,n}$ being achieved. If these subgoals cannot possibly be achieved then the parent goal cannot be achieved. Once it has been proven that these dependencies can be satisfied, total satisfaction of $G_{a,n}$ depends on the successful completion of other subgoals in the chosen plan (represented by one of the Activity Digraphs)

5.2.2 Goals

Goals are classified hierarchically, branching down into subgoals from one ultimate goal. The two sub-ultimate goals in human behaviour are as *self-preservation* and *preservation of the species*. The term *ultimate goal* is used for the goal at the very top of the hierarchy for completeness. This Goal Hierarchy is depicted in Figure 5.2. Incremental numbers identify each level of goals with Level 0 being the ultimate goal.

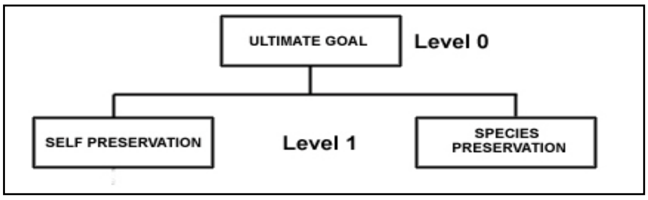


Figure 5.2 Goal Hierarchy

All other agent goals fall into this hierarchy as subgoals of Level 1. Each subsequent level of subgoals is appropriately named Level 2, Level 3 … Level *n*. These sublevels are partially-ordered sets of goals that are combined in the form of Activity Digraphs to satisfy the goal on the level above and partially satisfy goals on subsequent higher levels. Each level of the hierarchy can be displayed as an Activity Digraph that represents the paths through this network of goals that lead to the achievement of the parent goal. Figure 5.3 shows a partial view of an example Goal Hierarchy for a pet dog with Figure 5.4 displaying the associated Activity Digraphs.

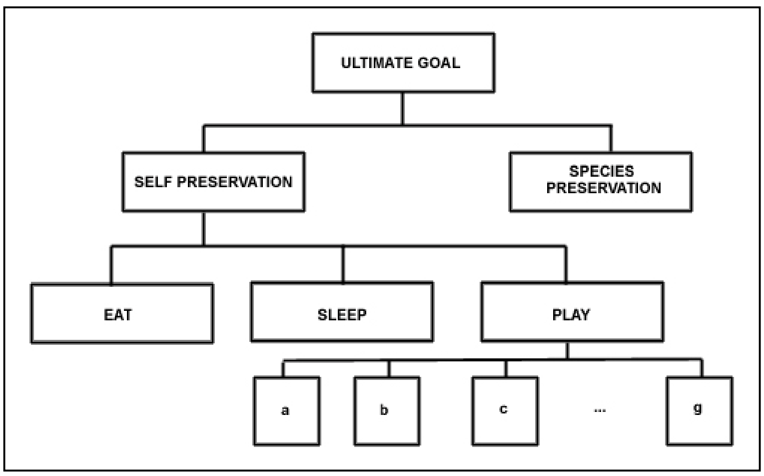


Figure 5.3 Partial View of the Goal Hierarchy for a Pet Dog

For the sake of this example, assume there are three ways in which the PLAY goal can be achieved. As more subgoals are added to the example, it becomes complex and difficult to examine the application of the Goal Hierarchy in the EMAI architecture. Therefore, for this example the number of subgoals is three.

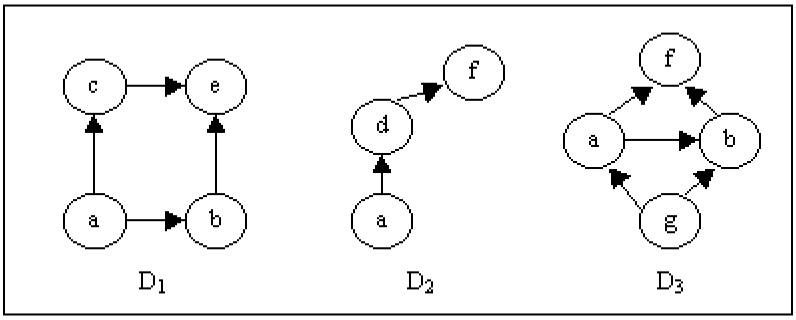


Figure 5.4 Activity Digraphs for PLAY

For a goal *a* at *Level n* ($G_{a,n}$), there may exist a number of Activity Digraphs ($D_{a1,n+1}$, $D_{a2,n+1}$, …, $D_{ak,n+1}$) consisting of subgoals of $G_{a,n}$ from *Level n+1*. When one of these Activity Digraphs is executed successfully, the goal $G_{a,n}$ will be satisfied as explained by Expression 5.4.

$$Satisfied(G_{a,n}) = Completed(D_{a1,n+1}) \ \ or \ \ Completed(D_{a2,n+1}) \ \ or....or \ \ (D_{ak,n+1}) \quad 5.4$$

where *k* is the number of digraphs in *Level n+1*. A goal can be defined in terms of its subgoals as shown in Expression 5.5:

$$S(G_{a,n}) = \bigcup_{i=1}^{k} V(D_{i,n+1}) \quad 5.5$$

where *S* is a function that when applied to a goal, will return a set of all its subgoals, and *V* is a function that when applied to an Activity Digraph will return a set of goals represented in the Activity Digraph (represented by the vertices of the Activity Digraph). In other words, a goal is defined in terms of its immediate subgoals. However, note that a goal $G_{a,n}$

can be achieved by satisfying any one of the Activity Digraphs, which may only contain a sub-set of all the subgoals of $G_{a,n}$. The union of all vertices of the Activity Digraphs gives the set of all individual subgoals of $G_{a,n}$.

The set of vertices V_t of the Activity Digraph $D_{t,n+1}$ will form a sub-set of all the subgoals of $G_{a,n}$ as shown in Expression 5.6:

$$V_t = V(D_{t,n+1}) \subseteq S(G_{a,n}), \qquad 1 \leq t \leq k \tag{5.6}$$

and the intersections between the Activity Digraphs are not necessarily empty, as shown in Expression 5.7:

$$V(D_{x,n+1}) \cap V(D_{y,n+1}) \neq \{\}, \qquad 1 \leq x, y \leq k, and \ x \neq y \tag{5.7}$$

Not every subgoal need be completed to satisfy the supergoal. If a goal consists of more than one Activity Digraph, only one of the digraphs need be completed to satisfy the goal. This can be seen in Figure 5.4. To satisfy the PLAY goal there are three choices of Activity Digraphs. No matter which digraph is used to satisfy the goal, subgoal *a* will always be satisfied as it appears in all three digraphs. However, the subgoals, *c*, *d, e, f* and *g*, may or may not be satisfied depending on which digraph is chosen.

At the lowest extreme level of the hierarchy exist the atomic goals. These goals have only one process. At this point the hierarchy could be reduced to where goals become purpose (for example, the purpose of the *heart* is to keep beating and circulate blood throughout the body). The *heart* has a fixed process. How far the Goal Hierarchy decomposes is dependent on its use. If the agent was simulating a pet dog, the purpose goals become holistic behaviours and it is acceptable to leave them out. If the agent was needed to simulate heart disease and the effects of eating fatty foods, then the Goal Hierarchy may be decomposed to this level.

Using the partial Goal Hierarchy in Figure 5.5, the atomic goal CHEW BONE could be represented as an event by the conceptual graph:

```
[CHEW] –
      (OBJ) -> [BONE]
      (AGNT) -> [AGENT: EMAI]
```

```
(TIM) -> [TIME: #now]
(LOC) -> [BACKYARD]
```

where the concept types NOW and BACKYARD are arbitrary values given in this example. The time and context (in this case *location*) could easily have been given as TOMORROW and ON THE BEACH, respectively. However, for the goal CHEW BONE the satisfaction of the goal is the primary focus of the event and the time and context, in this case are irrelevant.

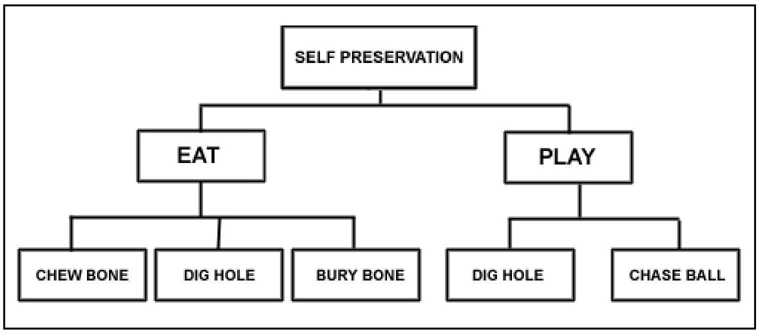


Figure 5.5 Hierarchy Showing Repetition of Goals on Different Levels

It is also apparent that as the hierarchy is decomposed further, goals and their subgoals are repeated as subgoals of other goals. Logically this creates a complex network of goals, but keeping with the hierarchical structure it is important to remember the goal to subgoal relationship cannot be shown in a network and that a goal cannot decompose into itself on any level. A partial hierarchy such as the one shown in Figure 5.5 is quite acceptable. In this figure, the Goal Hierarchy is shown including the subgoals EAT and PLAY. When the agent's EAT goal is triggered, it identifies the options, CHEW BONE, DIG HOLE and BURY BONE. The subgoals of the EAT goal are all part of the Activity Digraph that must be completed to satisfy the EAT goal. In contrast the PLAY goal has two mutually exclusive goals, either of which may be completed to satisfy the PLAY goal. Notice the DIG HOLE subgoal is repeated. In the hierarchy, this goal represents the same DIG HOLE goal, which in turn decomposes into the same DIG HOLE event. This illustrates how a subgoal of one goal can be a subgoal of another goal. Often the same behaviours can be performed with respect to satisfying different goals.

Also, illustrated here is a goal that has two different subgoals and associated processes for satisfying it. The PLAY goal can either be satisfied by successfully executing the DIG HOLE or CHASE BALL subgoals and the associated actions. This would give the agent two courses of action. How the agent chooses between subgoals is based on the priority of the goal calculated by assessing its urgency, if the agent's physical state and environment states are adequate for completing the goal, how the agent feels emotionally about the goal and the agent's mood.

5.2.3 Goal Activation

Each goal in the Goal Hierarchy can be derived from one or a combination of motivational urges. In the case where the agent has a complex Goal Hierarchy the mechanisms that cause the agent's goals to become activated and subsequently satisfied become elaborate. The agent desires to achieve all the goals in its Goal Hierarchy but not all at the same time. The goals that have been activated or *triggered* and are currently driving the behaviour of the agent have been extracted from the Goal Hierarchy, processed into events and placed in the agent's schedule. The amount of commitment the agent has to complete a goal is the intensity of desire it has toward completing the goal. For example, the agent may have an EAT goal in its Goal Hierarchy, but unless the goal is activated, the agent will not be attempting to satisfy that goal.

The agent will behave in response to activated goals. Therefore, in order to generate behaviour in the agent, a goal has to be triggered. The Sensory Processor (see Section 3.3.3) performs this process. The Sensory Processor receives internal sensory data from the Motivational Drive Generator that it then matches to appropriate goals from the Goal Hierarchy. The Motivational Drive Generator produces the internal sensory data by monitoring a series of Internal State Registers that represent motivational urges. Each register has a level value that records the strength of the associated urge. This level is influenced by changes in time and environmental conditions. The levels on the registers are controlled by drive mechanisms that simulate the motivational drives identified by Balkenius (1995) (see Section 2.3.2). These drive mechanisms are the Homeostatic Drive, the Cyclical Drive and the Default Drive.

The Homeostatic Drive is used to return the agent to an unexcited state. This drive causes the agent to seek a *drive-reduction* remedy. Initially the agent assumes an unexcited, normal physical and emotional state. Should the physical or emotional states change to extend beyond threshold values, the Homeostatic Drive sends the appropriate internal sensory data to the Sensory Processor, which triggers appropriate goals within the agent's Goal Hierarchy. On satisfaction of these goals the agent will have returned to a less excited state, if not having returned it to its norm. For example, the agent may be programmed with a simulated *body temperature* of 20°C. This temperature would be represented internally as an Internal State Register with a level value of 20. If the value on this *body temperature* register changes, the Homeostatic Drive will cause the triggering of goals that make the agent seek a treatment to return its *body temperature* register to its normal position. Other examples of simulated changes that would initiate the Homeostatic Drive would be the detection of *pain*, *boredom* or *social situations*.

The EMAI architecture models the Cyclical Drive in a similar manner to the representation of emotion in Padgham and Taylor (1997). The urges are modelled using a Temporally Cyclic Internal Engine and Internal State Registers. The Cyclic Engine synthesises a constant rate of change in the physical states of the agent that are affected by time differentials in its environment. The Internal State Registers keep a count of the updates made by the Cyclic Engine and at a threshold value the Motivational Drive Generator will send appropriate internal sensory data to the Sensory Processor that will trigger an appropriate goal. For example, an agent with an EAT goal will have a corresponding register called *hunger*. At preset temporal intervals the Cyclic Engine will increase the value on the *hunger* register. This simulates the agent becoming *hungry* over time. When the *hunger* register reaches a critical threshold level, the Homeostatic Drive will detect it and cause the agent's EAT goal to become activated via the Sensory Processor.

The Default Drive acts as a *deliberate* mechanism to produce behaviour in contrast to the other two drives, which are essentially *reactive*. Without the Default Drive the agent would appear inactive when it was not attending to its physical or emotional needs. This drive is used to generate behaviours that appear exploratory or anticipatory. Behaviours produced are not due to changes in the agent's physical, emotional or environmental state as is the cause with the Homeostatic and Cyclical Drives.

As already briefly discussed, the motivational urges are implemented in the Motivational Drive Generator as a series of Internal State Registers. The Motivational Drive Generator monitors the levels of the registers via the drive mechanisms and passes appropriate data to the Sensory Processor when the preset threshold values for the registers are reached. One or more registers, depending on the depth of the agent's Ontology and what the agent is simulating, may represent each urge. A *hunger* register, a *body temperature* register and a *fatigue* register could represent the urges produced by the Homeostatic Drive. If the implementation of the EMAI architecture was not concerned with simulating *body temperature* and *fatigue* these registers could be omitted. Whatever the number of registers implemented, the levels on these registers are maintained by the Homeostatic, Cyclical and Default Drives and used via the Sensory Processor to activate goals in the Goal Hierarchy. The goal EAT could be triggered by a certain level on a *hunger* register that is updated by the Cyclical Drive and read by the Homeostatic Drive. The goal SEEK WARMTH could be triggered by a certain level on the *body temperature* register that is updated by the Homeostatic Drive. To model the Internal State Registers, the set of registers are represented as *R*, where a particular register *r* has a neutral intensity $I_0(r)$, a positive value r^+ and a negative value r^-. A register with a positive intensity level will trigger a different set of goals to the negative intensities. For example, a positive value on the *body temperature* register (that represents an increased temperature) will cause the agent's SEEK COOLING goal to become active. In contrast a negative value on the *body temperature* gauge would trigger the SEEK WARMTH goal. Figure 5.6 illustrates an example Internal State Register.

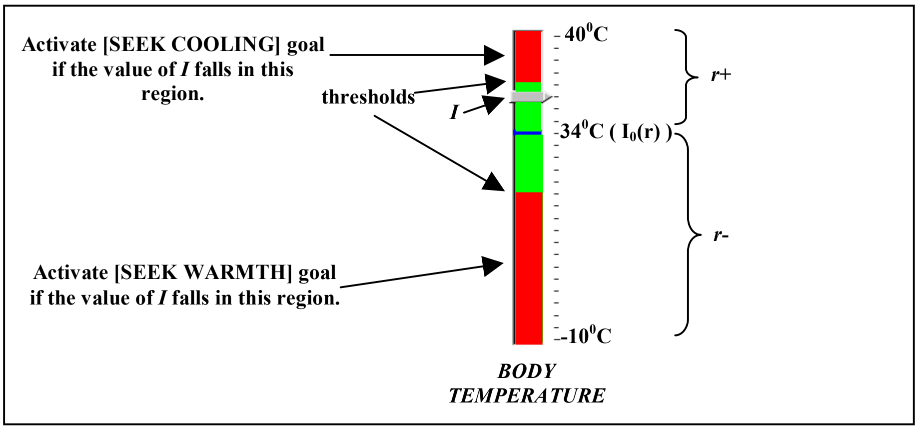


Figure 5.6 An Example of an Internal State Register Representing *Body Temperature*

The register that causes the goal to become active is also affected by the outcomes of the processes used to satisfy the goal. The agent may begin to feel hot (a value of r^+), at a threshold value of r^+, the agent's SEEK COOLING goal will have a priority proportional to the intensity of the value on the register. The agent may be programmed to withstand temperatures that are a threshold value 10°C above its norm. If the register exceeds this threshold value, an appropriate goal will be triggered to attempt to reduce the value on this register. The agent will seek to satisfy its goal with an appropriate process. The atomic goal or process may be REMOVE CLOTHING. On successful completion of the process the agent will have satisfied the SEEK COOLING goal and the intensity of the agent's *body temperature* register will have been reduced. Of course the REMOVE CLOTHING process may not reduce the register completely (that is setting it back to $I_0(r)$) and the agent's SEEK COOLING goal will not have been satisfied. In this case the agent's SEEK

COOLING goal will still be active and the agent will continue to seek other ways of satisfying this goal.

In summary, there is an Internal State Register *r* with intensity, *I(r)*. At any point in time ($t + \delta$) the intensity of *r* can be calculated using Expression 5.8:

$$I_{t+\delta}(r) = I_t(r) + GR\ (r,\delta) - SA\ (r, goal\_outcome) \qquad 5.8$$

where *GR*[10] is a function of the growth in the intensity of the register since time *t* and *SA*[11] is the level of satisfaction for a motivational urge (represented by the register *r*) incurred by the *goal_outcome*. The intensity of a register is equal to the previous value of the register plus any growth in the intensity minus any satisfaction of that intensity due to a goal outcome. An agent's urge has been satisfied when:

$$SA\ (r, goal\_outcome) \geq I_t(r) + GR\ (r,\delta) \qquad 5.9$$

For example, assume that an EMAI agent has a neutral intensity level ($I_0(r)$) of 34^0C and a r^- threshold value of 20^0C for the Internal State Register representing *body temperature*. If the agent is placed in an environment such that the rate of growth of its intensity is -1^0C per second, after 5 seconds, the value of *G(r, 5)* will be -5. The intensity value represented by the corresponding Internal State Register, *I(r)*, would be 29^0C (34 - 5). After a further 10 seconds the value of *I(r)* would be 19^0C. The value of *I(r)* would now be below the r^- threshold value. The Homeostatic Drive mechanism would detect this drop in *body temperature* below the threshold and send appropriate internal sensory data to the Sensory Processor to indicate the agent was cold. The Sensory Processor would in turn trigger the SEEK WARMTH goal. The agent may decide to execute a process for ADD CLOTHING. This may result in a *body temperature* increase of 5^0C, which would increase *I(r)* to a value of 24^0C. As the new value for *I(r)* is above the r^- threshold, the SEEK WARMTH goal will be deactivated. If however, the value of *I(r)* were to fall to a

[10] The growth function is performed in an EMAI agent by the Cyclical Drive Mechanism that applies growth values to the intensities of the Internal State Registers with each cycle.

[11] The satisfaction function occurs after the agent has performed a behaviour. The Sensory Processor receives external sensory data relating to the outcomes of behaviours and the appropriate Internal Sensory Registers are updated. The satisfaction function is used to test if the values on the Internal Sensory Registers have returned to normal.

value of 10^0C the execution of the ADD CLOTHING event would only raise the value of *I(r)* to 15^0C. As this value is still below the r^- threshold, the SEEK WARMTH goal would remain active. In this case the agent would have to continue performing the ADD CLOTHING event until the value of *I(r)* returned to a level above the r^- threshold.

5.2.4 Prioritisation of Goals

Each time the drive mechanisms within the Motivation Drive Generator detect a register's value is outside the threshold values, internal sensory data is sent to the Sensory Processor, which activates the associated goal. If the goal is already active, the priority of the goal increases. At time *t*, the priority placed on a goal, $P(G_{a,n})_t$, is proportional to the total intensity of the register/s that triggered the goal as shown in Expression 5.10.

$$P(G_{a,n})_t = \sum_{f=1}^{p}(I_t(r_f)) \qquad 5.10$$

As shown by Expression 5.10 the priority of a goal can be determined by more than one Internal State Register. Two or more registers may influence the priority of a single goal. For example, if there existed a register for *hunger* and a register for *boredom*, the internal sensory data sent to the Sensory Processor from both these registers via the Homeostatic Drive mechanism could be used to trigger the EAT goal. This also implies the satisfaction of a single goal can affect the values on more than one Internal State Register. For example, satisfaction of the EAT goal may affect the values on not only the *hunger* and *boredom* registers but it may also affect the *body temperature* register.

It is important to note that all goals in the hierarchy are not always active. It is not until the register reaches a critical threshold point that the goal is activated and the agent believes its objective is to fulfill the goal and begin a set of processes that will satisfy the goal. In other words, the goal lays dormant in the Goal Hierarchy until it is triggered. At this point the goal becomes active and the agent makes further investigations into the processes that will achieve the goal, thus suppressing the register and in turn the drive that began the process.

The Intention Generator uses the prioritising values placed on the goals to order them for execution. Further details of this process are delayed until Chapter 9 where this motivational intensity prioritising is coupled with emotional ordering to produce the agent's schedule of behaviours.

5.3 Case Study: Urge Thresholds, Goal Setting and Prioritisation

5.3.1 Introduction

The purpose of this case study is to illustrate the urge thresholds, goal setting and goal prioritisations of the Motivational Drive Generator and the Sensory Processor described earlier in this chapter. To do this, the EMAI agent has been programmed with the Goal Hierarchy shown in Figure 5.7[12]. The inspiration for this hierarchy stems from the basic instinctual behaviours present in a pet dog-like animal. Other agent architectures have been examined by simulating pet animals and in particular dogs such as PETEEI (El-Nasr 1998) and DOGZ (Stern et al. 1998). The reason for this is that a pet dog has a simple set of behaviours (compared to humans) and most humans have an idea on how these animals should behave under varying conditions (El-Nasr 1998), making analysis simpler.

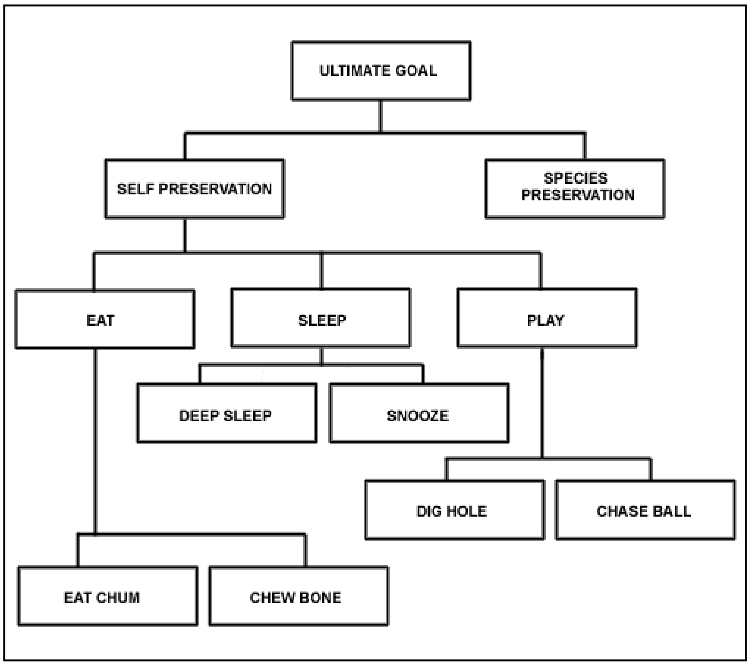


Figure 5.7 Simple Goal Hierarchy for a Dog-like Agent

The behavioural traits of the pet dog have their basis in the evolutionary process. Through natural selection, a canine model has formed giving the species physical

[12] Note: CHUM is a dog-food brand in Australia.

characteristics and mental abilities to continue their existence. Their behavioural patterns guide them through their day to day and year to year existence helping them to make the right choices to protect themselves and their offspring (Foster and Smith 2000). Beside the essential survival goals of eating and sleeping, a dog's playing behaviour stems from their puppy play. This taught them hunting skills and how to interact with other dogs in the wild. Even digging is evolutionary and dogs instinctually dig holes for protection. For these reasons, each of these goals have been integrated into the EMAI architecture for the demonstration of this case study.

The EAT, SLEEP and PLAY goals each have two subgoals. Each subgoal is mutually exclusive. This means that any of the two subgoals can be satisfied in order to complete the associated supergoal. If the PLAY goal is activated, either DIG HOLE or CHASE BALL needs to be completed to satisfy PLAY. To keep this example simple, the agent is programmed with three Internal State Registers that each trigger one parent goal when a threshold value is reached. At each time interval in the simulation, each register's value is updated by a predetermined amount. In the simulation the registers for *eat, sleep* and *play* are set to high values indicating a low depletion in these areas. For example, a high value on the *eat* register indicates the agent's nutritional reserves are high. As the simulation runs the values on the registers decrease. On reaching a threshold level, an appropriate goal is triggered. The agent schedules the subgoals for execution. (This process is examined in depth in Chapter 6 and will not be covered here.) When a subgoal is successfully completed, the goal is satisfied and the respective register's value is increased. Figure 5.8 graphically displays the simulation run. Each register value is displayed with respect to time and a narrative produced by the agent during the simulation is shown at the appropriate time intervals. A time interval without narrative illustrates the agent's inactivity. Beside the Internal State Register levels, the threshold value at which a goal is triggered is also shown.

5.3.2 Results

As the simulation begins, all agent goals are inactive. Therefore, the agent remains behaviourally idle. As time passes, the graph shows the depletion of the agent's Internal State Registers. Between time interval 4 and 5, the levels on the play and sleep registers reach their threshold values and the PLAY and SLEEP goals are activated. This event places the processes for DIG HOLE, CHASE BALL, DEEP SLEEP and SNOOZE into the agent's schedule. As the agent cannot perform PLAY and SLEEP activities at the same time it has to choose just one. At this point in the simulation the PLAY and SLEEP goals have been activated the same number of times (in this case, once) and therefore, have the same level of priority placed on them. For this case study the behaviour selected is the one entered into the schedule first (further prioritisation mechanisms will be examined in Chapter 9).

At time interval 5 the agent selects the DIG HOLE activity. This activity is completed successfully and the level on the *play* register can be seen to increase above the threshold value, thus satisfying the PLAY goal. The activities corresponding to DIG HOLE and CHASE BALL are removed from the agent's schedule. At the same time the *sleep* and *eat* registers continue to decrease at their temporally cyclical rate plus a small variation caused by the PLAY activity. In the next time interval, the agent no longer has an active PLAY goal and it therefore moves to its next scheduled activity that will satisfy the still active SLEEP goal. By time interval 7, both the PLAY and SLEEP goals have been satisfied, the appropriate registers updated, all scheduled activities removed and the goals become inactive.

By time index 9 the agent's *eat* register has reached its threshold. The EAT goal is activated, the activities for EAT DOG FOOD and CHEW BONE are scheduled, but the agent does not immediately begin executing these as the SLEEP goal has also been activated again. At time interval 10 when the agent has finished the event for the SLEEP goal, the agent is free to attempt to satisfy the EAT goal by executing one of the two associated activities. The agent begins by attempting the EAT DOG FOOD activity. However, this activity requires the presence of food in the agent's environment of which there is none. Therefore, this activity fails and the agent tries the next activity of CHEW BONE. Again, there are no simulated *bones* present in the environment and this activity fails too. Furthermore, the failure of these activities has not caused the *eat* register to increase and the EAT goal is still active.

The EAT goal remains active and the agent tries to satisfy it when it is not performing an activity associated with a goal that can be satisfied. At time index 14, the user adds some simulated food to the agent's environment. At time index 15, the agent's attempt to satisfy the EAT goal is successful and the *eat* register's value is increased.

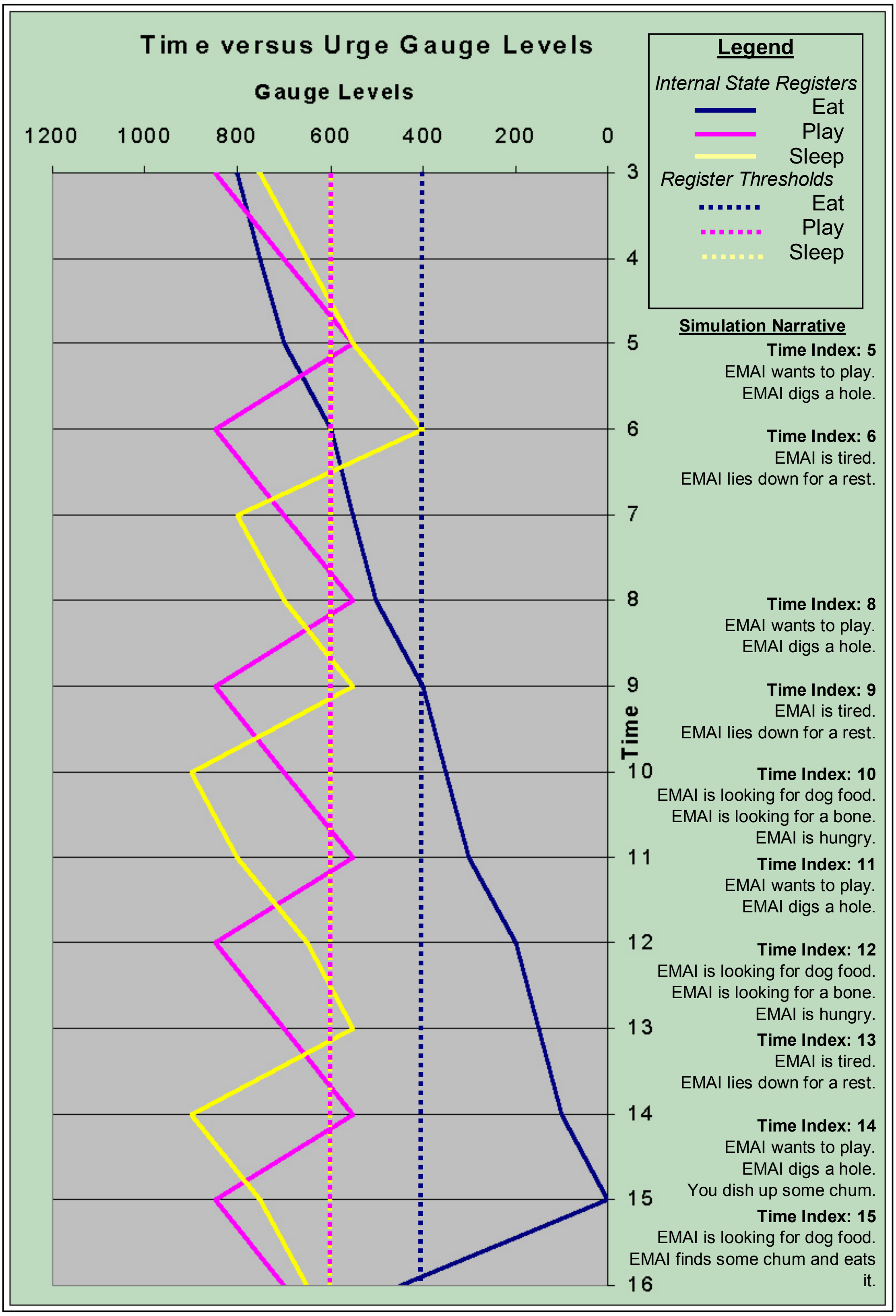


Figure 5.8 Simulation Run with EMAI Agent Displaying Internal State Register Levels, Thresholds and Narrative

5.3.3 Discussion

This case study illustrates how an EMAI agent prioritises goals and how the Internal State Registers and cyclical drive mechanisms are used to simulate motivation. These drives are used internally by the agent to produce behaviours by activating the goals within the Goal Hierarchy of the Ontology. Once a register has reached its threshold value and the corresponding goal is activated, that goal remains active until the register has returned and surpassed the threshold point. In the simulation depicted in Figure 5.8, each time the agent satisfactorily completes an activity, the appropriate register is returned to a value on the other side of the threshold. This may not always be the case as demonstrated in Figure 5.9.

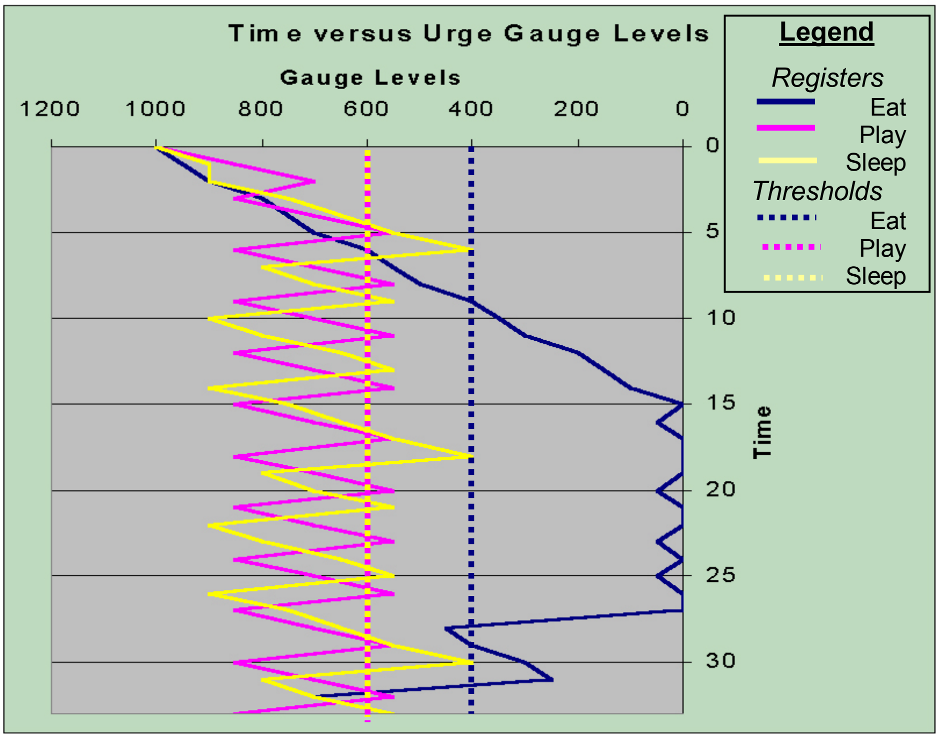


Figure 5.9 Simulation Run with Fluctuating Gauge Values Beyond the Threshold

At time interval 15, the eat register begins to fluctuate in value. This is the result of the agent's successful attempts to complete an activity that satisfies the EAT goal but does not raise the level of the eat register beyond the threshold. In this simulation the small increases in the eat register are caused by the agent completing the CHEW BONE activity. While this activity does suppress the agent's eat register a little, it does not return the eat register to a value above the threshold and therefore, is simulating that the agent is still hungry. Therefore, the agent's EAT goal remains active and the agent will continue to execute EAT activities. During the time intervals 15 to 22, the agent's environment was populated with a number of simulated bones but no other food. Therefore, the agent could only satisfy the CHEW BONE activity and not the EAT DOG FOOD activity that would have raised the eat register more substantially. At time interval 23, the user adds other simulated food item to the environment. This item allows the agent to successfully complete the EAT DOG FOOD action and this results in a raise in the eat register value to beyond the threshold; the EAT goal being deactivated and the EAT events being removed from the schedule.

Note the priority associated with a goal (as per Expression 5.10) equals the sum of the intensity of all the registers (motivational urges) that affect the goal. In this case study, there is only one register (motivational urge) associated with a particular goal. Therefore, the priority associated with a goal, like PLAY is calculated as:

$$P\ ([PLAY])_t = \sum_{f=1}^{1}(I_t(r_f)) \quad \text{where } r_1 \text{ represents the } \textit{play} \text{ register.}$$

One final point that should be made about this simulation is that the agent had subgoals and corresponding activities that were never executed. These were the SNOOZE and CHASE BALL subgoals. As this simulation worked purely on prioritising activities based on the urge register levels, all subgoals of the same parent goal were assigned the same priority. This meant that when the agent came to selecting an activity in the schedule based on priority, activities for the same goal were ranked evenly. Whenever this occurs, the agent selects the item that entered the schedule first by a FIFO method to resolve any conflict or indecision. Therefore, whenever the PLAY goal was activated the DIG HOLE process would be executed before the CHASE BALL process. As the DIG HOLE process was always successful in satisfying the PLAY goal, the agent never had to execute the CHASE BALL process. Although this is an unlikely simulation of real behaviour in solving conflict and selecting behaviour, it is adequate for the purposes of this simulation.

5.4 Summary

The EMAI architecture implements a Knowledge Area that includes a Motivational Drive Generator, Sensory Processor and an Ontology with a Goal Hierarchy. Before an EMAI agent will act, an appropriate goal in the Goal Hierarchy has to become active. The Motivational Drive Generator contains a series of Internal State Registers that maintain the motivational urges of the agent. Three drive mechanisms (homeostatic, cyclical and default) monitor the values of the Internal State Registers. When an Internal State Register reaches a threshold value, the appropriate drive mechanism sends data to the Sensory Processor. The Sensory Processor processes the data and determines which goal in the Goal Hierarchy to activate. When a goal in the Goal Hierarchy becomes active, the agent is able to decompose the goal to locate one or more Activity Digraphs that could be performed to satisfy the goal. This mechanism simulates a human-like goal-orientated behaviour in the agent. This goal-orientated behaviour gives the agent the ability to recover from failed plans and to recognise and create new plans. This capability is illustrated in the case study.

The processes, demonstrated in the case study presented in this chapter, illustrate the underlying goal-orientated processes and motivation mechanisms working in the EMAI architecture. It shows how lower level emotions represented by the eat, sleep and play registers are used to generate behaviours within the agent. This produces an agent capable of goal-orientated behaviour, but with little variation in action choices. These mechanisms make the agent quite predictable and less complex than real animal or human behaviour. The next chapter examines the concept of higher emotional evaluation using attitude theory and discusses the ways in which this has been integrated into the agent architecture in order to add depth to the personality of the agent.

6. Event Space

It is easy to be wise after the event.
- 17th century proverb.

6.1 Introduction

Humans (and other intelligent biological animals) are generally unaware of their cognitive processes (Lefton 1994). Psychologists refer to these processes as *thinking*. The term thinking in psychology refers to the sorting through of choices and deciding what behaviours to perform. In brief, thinking refers to *reasoning*, *decision making* and *problem solving* (Galotti, 1989). In psychology, reasoning refers to the process of generating and evaluating situations and decision making refers to assessing and choosing among alternatives. In the EMAI architecture, these same processes occur within the Event Space Generator.

This chapter introduces the EMAI architecture's Event Space Generator. It begins by examining the structure of a single event and how a partially-ordered set of events can be used to define a plan. An EMAI agent uses these plans to achieve its goals. Each event represents a specific outward behaviour that can be exhibited by the agent. As there may be more than one plan that could be executed to achieve a single goal, an EMAI agent must determine the nature of all events involved by specialising each plan into elementary concept types from its Ontology. This collection of specialised, competing events (that form alternative plans) is called the *event space*.

This chapter continues by formalising an event space and examining a case study that demonstrates the processes undertaken by the Event Space Generator in turning an abstract goal into a series of specialised plans. The discussion begins by re-examining the earlier definition of an *event*, and creating an alternative definition in terms of its constituents. This new definition is necessary for measuring the emotional attachment that an agent has towards each and every entity associated with a particular goal.

6.2 Events

As described earlier, events are represented as activities in the EMAI's Activity Digraphs (see Section 4.3.2). Each event represents a self-contained independent activity the agent can perform. It is represented in the EMAI Ontology as an Activity Digraph with a set of preconditions, postconditions and a delete list. Figure 6.1 illustrates a simple Activity Digraph with three events.

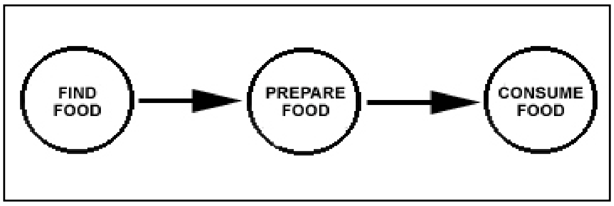


Figure 6.1 A Simple Activity Digraph with Three Events

This Activity Digraph can be represented as a conceptual graph by using the temporal relation FINISH-BEFORE-START (FBS), as shown below:

```
[FIND FOOD]->(FBS)->[PREPARE FOOD]->(FBS)->[CONSUME FOOD]
```

where FIND FOOD, PREPARE FOOD and CONSUME FOOD are events.

An event is also represented in an EMAI's Ontology as a conceptual graph. It defines a behavioural event as having the four elements of action, object, time and context. The type definition of an event is thus:

```
type EVENT(x) is:
                    [ACTION:*x]-
                          (OBJ) -> [ENTITY:*m]
                          (AGNT) -> [ENTITY:*n]
                          (TIM) -> [TIME:*t]
                          (CON) -> [CONTEXT:*c]
```

The type EVENT appears in the Type Hierarchy as a subtype of ACTION as shown in Figure 6.2. Specific events, such as the ones shown in Figure 6.1 are subtypes of the type EVENT.

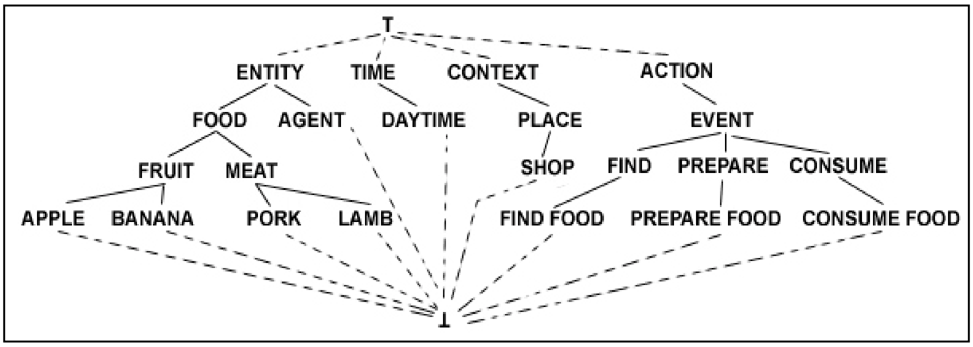


Figure 6.2 Type Hierarchy Showing Event Types

By specialising the conceptual graph for the generic event the graph for the specific event can be obtained. The conceptual graph representation of FIND FOOD could be:

```
[FIND] –
      (OBJ) -> [FOOD]
      (AGNT) -> [AGENT: EMAI]
      (TIM) -> [TIME: #now]
      (CONT) -> [BACKYARD]
```

This graph provides the agent with precise knowledge about how, when, where and with what an action should occur to execute the associated event. Thus an event can be re-defined as Expression 6.1:

$$E = \{a, o, c, t\} \tag{6.1}$$

where a is the action that relates to the event, o is the object involved or affected by a, c is the context or condition where a is taking place or being performed and t is the temporal component of a. Each of these elements may vary along a dimension of explicitness. At the most exact level of identifying the event, an EMAI agent will intend to perform a specific action, with or towards a certain object, in a particular context or situation at an exact point in time. For example, an EMAI agent may intend to *buy a bunch of flowers* (action) for *Georgina* (object) from *Barry's House of Posies* (context) at *12:30 p.m. on Tuesday* (time). At the other end of the spectrum, an EMAI agent may intend to be affectionate without referral to any exact elements of the event. For example, an EMAI agent may intend to *buy a gift* for a *friend*. How each component of the intended behaviour is categorised is inconsequential. The *bunch of flowers* could also be defined as an object of the EMAI agent's behaviour. It could also just as easily intend to buy Georgina a *box of chocolates*. As there may be multiple elements of the same type involved, the event and its elements are evaluated using the same method, it is unimportant how they are precisely defined.

To continue with the EAT goal example (from previous chapters), the Event Space Generator constructs events from the abstract goals that are passed to it. From this information the Event Space Generator has the Activity Digraph of:

```
[FIND FOOD]->(FBS)->[PREPARE FOOD]->(FBS)->[CONSUME FOOD]
```

Because of the level of abstraction of the EAT goal, no specific item of food or other contexts have been designated, the agent needs to construct specialised conceptual graphs that give the agent an exact perception of how it is meant to behave. The agent's Event Space Generator component takes the conceptual graph and specialises its concept types into subtypes in the Ontology that exist on sublevels of the Type and Relation Hierarchies. This creates a set of specialised events from which the agent can select in order to satisfy a goal. For example, given the FIND FOOD conceptual graph:

```
[FIND] –
          (OBJ) -> [FOOD]
          (AGNT) -> [AGENT: EMAI]
          (TIM) -> [TIME: #now]
          (LOC) -> [SHOP]
```

By using the Type Hierarchy shown in Figure 6.2, the agent can specialise to better define the FIND FOOD event by specifying the exact type of food for which the agent should look. Subtypes of FOOD include MEAT and FRUIT and therefore, the graphs:

```
[FIND] –
          (OBJ) -> [MEAT]
          (AGNT) -> [AGENT: EMAI]
          (TIM) -> [TIME: #now]
          (LOC) -> [SHOP]
```

and

```
[FIND] –
          (OBJ) -> [FRUIT]
          (AGNT) -> [AGENT: EMAI]
          (TIM) -> [TIME: #now]
          (LOC) -> [SHOP]
```

will be constructed. In this agent's knowledge base, the types of MEAT and FRUIT are still abstract and could be further specialised into graphs that substitute MEAT for PORK or LAMB and FRUIT for APPLE or BANANA. If this were all the specialisation the agent were to perform on the graph for EAT, instead of one event plan, the agent would now have four specific event plans, one for each type of food. The specialisation of concept types and the resultant set of new specific events form the basis of the event space.

6.3 Event Space

A goal can have one or more associated Activity Digraphs that could, when executed successfully, achieve the goal (see Section 4.3.1). Each of the Activity Digraphs associated with a goal can embody one or move events. A goal can therefore, have any number of individual events associated with it. This collection of events will be called the *event space*. As with the definition of *event* in terms of its constituent, the definition of Activity Digraph, *AD*, in terms of fundamental elements can be defined as in Expression 6.2:

$$AD = \bigcup_{j=1}^{n} E_j \qquad 6.2$$

Given the definition of an event shown in Expression 6.1, the definition of an Activity Digraph can be rewritten as Expression 6.3:

$$AD = \bigcup_{j=1}^{n} \{ a_j, o_j, c_j, t_j \} \qquad 6.3$$

An event space for a goal *g*, written as ES_g, can be defined as shown in Expression 6.4

$$ES_g = \bigcup_{j=1}^{m} AD_j \qquad 6.4$$

where *m* is the number of Activity Digraphs, *AD*, associated with the goal, *g*. By substituting AD in Expression 6.4 for Expression 6.3, an event space can be defined in terms of all elementary entities involved in achieving the goal as Expression 6.5:

$$ES_g = \bigcup_{i=1}^{m} \left(\bigcup_{j=1}^{n} \{ a_j, o_j, c_j, t_j \} \right)_i \qquad 6.5$$

Through the use of conceptual graphs, that provide structure for events, and temporal sequences for Activity Digraphs, the Event Space Generator accepts an abstract goal from the Sensory Processor and builds the event space through a series of canonical graph formation rules. This process is examined in the following case study.

6.4 Case Study

The Event Space Generator accepts a primitive set of goals from the Sensory Processor and constructs actual behavioural processes in the form of specialised Activity Digraphs. This representation contains atomic and individual concepts from the Type and Relation Hierarchy. This case study examines how the Event Space Generator creates an event space from the abstract goals used to generate behaviours in an EMAI agent.

Assume that an EMAI agent has a goal of EAT and two Activity Digraphs that can be performed to satisfy the EAT goal. These Activity Digraphs, ad_1 and ad_2, are defined as:

```
ad1: [FIND FOOD]->(FBS)->[PREPARE FOOD]->(FBS)->[CONSUME FOOD]
ad2: [FIND MONEY]->(FBS)->[BUY FOOD]->(FBS)->[CONSUME FOOD]
```

and illustrated in Figure 6.3.

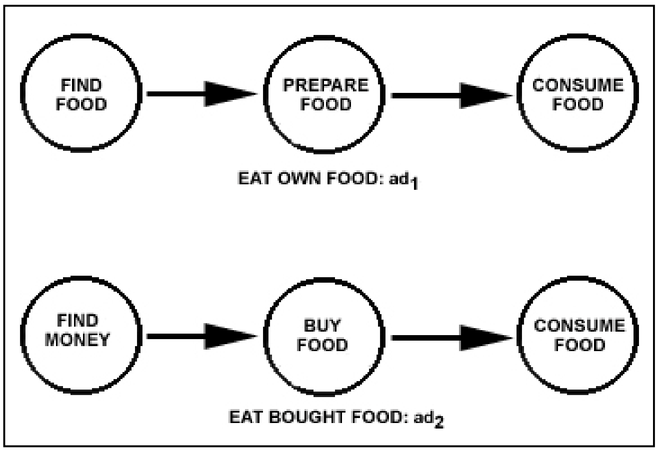


Figure 6.3 Two Activity Digraphs to Satisfy the Goal EAT

As these Activity Digraphs describe generic events, the Event Space Generator must specialise each using atomic concepts from the Type Hierarchy to identify actual elements that exist in the agent's environment. However, before this can occur the Event Space Generator must expand the Activity Digraphs so the events and their elements are visible. It does this by projecting the Event Graphs into the Activity Digraph from the agent's Ontology that represent each of the events in the Activity Digraphs with the conceptual graphs for the digraphs. Assuming the conceptual graph for the FIND FOOD event is:

```
[FIND] –
          (OBJ) -> [FOOD]
          (AGNT) -> [AGENT]
          (TIM) -> [TIME: #now]
          (CONT) -> [AT HOME]
```

the conceptual graph for the PREPARE FOOD event is:

```
[PREPARE] –
          (OBJ) -> [FOOD]
          (AGNT) -> [AGENT]
          (TIM) -> [AFTER FIND]
          (CONT) -> [AT HOME]
```

and the conceptual graph for CONSUME FOOD event is:

```
[CONSUME] –
          (OBJ) -> [FOOD]
          (AGNT) -> [AGENT]
          (TIM) -> [AFTER PREPARE]
          (CONT) -> [AT HOME]
```

the Event Space Generator will process the graph representing ad_1 (EAT OWN FOOD):

```
[FIND FOOD]->(FBS)->[PREPARE FOOD]->(FBS)->[CONSUME FOOD]
```

into the expanded graph as shown below:

```
[FIND] –
          (OBJ) -> [FOOD: *y]
          (AGNT) -> [AGENT: *x]
          (TIM) -> [TIME: #now]
          (CONT) -> [AT HOME]
          (FBS) -> [PREPARE] –
                      (OBJ) -> [?y]
                      (AGNT) -> [?x]
                      (TIM) -> [AFTER FIND]
                      (CONT) -> [AT HOME]
                      (FBS) -> [CONSUME] –
                                  (OBJ) -> [?y]
                                  (AGNT) -> [?x]
                                  (TIM) -> [AFTER PREPARE]
                                  (CONT) -> [AT HOME]
```

The different coloured regions of the above conceptual graph represent the individual events that constitute the plan represented by the Activity Digraph. Rather than a graph that generally states that *some action should be taken toward something in someplace at sometime*, the Event Space Generator reduces this to more exact terms that specifically define an event such as *the finding of the food will take place in the shop now*. The same process that specialised ad_1 will also be performed on ad_2. Assuming the conceptual graph for the FIND MONEY event is:

```
[FIND] –
          (OBJ) -> [MONEY]
          (AGNT) -> [AGENT]
          (TIM) -> [TIME: #now]
          (CONT) -> [AUTOMATIC TELLER MACHINE]
```

the conceptual graph for the BUY FOOD event is:

```
[BUY] –
          (OBJ) -> [FOOD]
          (AGNT) -> [AGENT]
          (TIM) -> [AFTER FIND MONEY]
          (CONT) -> [SHOP]
```

and the graph representing ad_2 (EAT BOUGHT FOOD):

```
[FIND MONEY]->(FBS)->[BUY FOOD]->(FBS)->[CONSUME FOOD]
```

will be processed into the expanded graph as shown below:

```
[FIND] –
          (OBJ) -> [MONEY]
          (AGNT) -> [AGENT: *x]
          (TIM) -> [TIME: #now]
          (CONT) -> [AUTOMATIC TELLER MACHINE]
          (FBS) -> [BUY] –
                      (OBJ) -> [FOOD: *y]
                      (AGNT) -> [?x]
                      (TIM) -> [AFTER FIND]
                      (CONT) -> [AT SHOP]
                      (FBS) -> [CONSUME] –
                                  (OBJ) -> [?y]
                                  (AGNT) -> [?x]
                                  (TIM) -> [AFTER PREPARE]
                                  (CONT) -> [AT HOME]
```

Given the Type Hierarchy of Figure 6.2 the above graph can be specialised further to include the elementary types of FOOD: APPLE, BANANA, PORK and LAMB. Before continuing the relations of (TIM) and (CONT) and their linked types will be dropped from the example as this example is going to concentrate on the specialisation of the FOOD type. The reader should assume they are still part of the graphs although they will not be shown or specialised any further.

Once the Event Space Generator has produced an expanded conceptual graph from the Activity Digraph graph and the graphs that represent the individual events, the Event Space Generator attempts to further specialise each type in the graph. The process continues until the Event Space Generator has a set of graphs containing the most elementary types in the Type Hierarchy. Given the four FOOD types the Event Space

Generator will specialise the above conceptual graphs of *ad*$_1$ into a further four graphs thus:

```
1.  EAT OWN APPLE is
    [FIND] –
          (OBJ) -> [APPLE: *a]
          (AGNT) -> [AGENT: *x]
          (FBS) -> [PREPARE] –
                      (OBJ) -> [?a]
                      (AGNT) -> [?x]
                      (FBS) -> [CONSUME] –
                                  (OBJ) -> [?a]
                                  (AGNT) -> [?x]
2.  EAT OWN BANANA is
    [FIND] –
          (OBJ) -> [BANANA: *b]
          (AGNT) -> [AGENT: *x]
          (FBS) -> [PREPARE] –
                      (OBJ) -> [?b]
                      (AGNT) -> [?x]
                      (FBS) -> [CONSUME] –
                                  (OBJ) -> [?b]
                                  (AGNT) -> [?x]
```

```
3.  EAT OWN PORK is
    [FIND] –
          (OBJ) -> [PORK: *p]
          (AGNT) -> [AGENT: *x]
          (FBS) -> [PREPARE] –
                      (OBJ) -> [?p]
                      (AGNT) -> [?x]
                      (FBS) -> [CONSUME] –
                                  (OBJ) -> [?p]
                                  (AGNT) -> [?x]
4.  EAT OWN LAMB is
    [FIND] –
          (OBJ) -> [LAMB: *l]
```

```
        (AGNT) -> [AGENT: *x]
        (FBS) -> [PREPARE] –
                    (OBJ) -> [?l]
                    (AGNT) -> [?x]
                    (FBS) -> [CONSUME] –
                                (OBJ) -> [?l]
                                (AGNT) -> [?x]
```

As the initial goal of EAT could have been satisfied by one of two Activity Digraphs, the Event Space Generator will repeat the above process for the Activity Digraph ad_2 (EAT BOUGHT FOOD). As there are four types of FOOD this graph will also be specialised into four new graphs: EAT BOUGHT APPLE, EAT BOUGHT BANANA, EAT BOUGHT PORK and EAT BOUGHT LAMB as shown below:

```
1. EAT BOUGHT APPLE is
   [FIND] –
         (OBJ) -> [MONEY]
         (AGNT) -> [AGENT: *x]
         (FBS) -> [BUY] –
                     (OBJ) -> [APPLE: *a]
                     (AGNT) -> [?x]
                     (FBS) -> [CONSUME] –
                                 (OBJ) -> [?a]
                                 (AGNT) -> [?x]
```

```
2. EAT BOUGHT BANANA is
   [FIND] –
         (OBJ) -> [MONEY]
         (AGNT) -> [AGENT: *x]
         (FBS) -> [BUY] –
                     (OBJ) -> [BANANA: *b]
                     (AGNT) -> [?x]
                     (FBS) -> [CONSUME] –
                                 (OBJ) -> [?b]
                                 (AGNT) -> [?x]
3. EAT BOUGHT PORK is
   [FIND] –
```

```
            (OBJ) -> [MONEY]
            (AGNT) -> [AGENT: *x]
            (FBS) -> [PREPARE] –
                        (OBJ) -> [PORK: *p]
                        (AGNT) -> [?x]
                        (FBS) -> [CONSUME] –
                                    (OBJ) -> [?p]
                                    (AGNT) -> [*x]
  4.  EAT BOUGHT LAMB is
      [FIND] –
            (OBJ) -> [MONEY]
            (AGNT) -> [AGENT: *x]
            (FBS) -> [PREPARE] –
                        (OBJ) -> [LAMB: *l]
                        (AGNT) -> [?x]
                        (FBS) -> [CONSUME] –
                                    (OBJ) -> [?l]
                                    (AGNT) -> [?x]
```

For an EMAI agent to choose among the alternative plans (or to perform meaningful affective decision making), it has to measure its emotional attachment to each and every element in the plan. For this reason, the event space is defined in terms of its constituent elements as in Expression 6.5. In this case study, the event space for the goal EAT, written as ES_{EAT}, is as shown below:

$$ES_{EAT} = \bigcup \begin{pmatrix} E_{FIND-FOOD}, E_{PREPARE-FOOD}, \\ E_{FIND-MONEY}, E_{BUY-FOOD}, E_{CONSUME-FOOD} \end{pmatrix}$$

where each of these generalised events (except FIND MONEY because it has no subtypes) are specialised into:

$$E_{FIND-FOOD} = \bigcup \left(E_{FIND-APPLE}, E_{FIND-BANANA}, E_{FIND-PORK}, E_{FIND-LAMB} \right)$$

$$E_{PREPARE-FOOD} = \bigcup \left(E_{PREPARE-APPLE}, E_{PREPARE-BANANA}, E_{PREPARE-PORK}, E_{PREPARE-LAMB} \right)$$

$$E_{BUY-FOOD} = \bigcup \left(E_{BUY-APPLE}, E_{BUY-BANANA}, E_{BUY-PORK}, E_{BUY-LAMB} \right)$$

$$E_{CONSUME-FOOD} = \bigcup\left(E_{CONSUME-APPLE}, E_{CONSUME-BANANA}, E_{CONSUME-PORK}, E_{CONSUME-LAMB}\right)$$

Furthermore, each of the specialised events in the above expressions can be written in the format of Expression 6.1 thus:

$E_{FIND\ APPLE}$ = {FIND, APPLE, AT HOME, TIME:#now}
$E_{FIND\ BANANA}$ = {FIND, BANANA, AT HOME, TIME:#now }
$E_{FIND\ PORK}$ = {FIND, PORK, AT HOME, TIME:#now }
$E_{FIND\ LAMB}$ = {FIND, LAMB, AT HOME, TIME:#now }
$E_{FIND\ MONEY}$ = {FIND, MONEY, AUTOMATIC TELLER MACHINE, TIME:#now }
$E_{BUY\ APPLE}$ = {BUY, APPLE, AT SHOP, AFTER $E_{FIND\ MONEY}$}
$E_{BUY\ BANANA}$ = {BUY, BANANA, AT SHOP, AFTER $E_{FIND\ MONEY}$}
$E_{BUY\ PORK}$ = {BUY, PORK, AT SHOP, AFTER $E_{FIND\ MONEY}$}
$E_{BUY\ LAMB}$ = {BUY, LAMB, AT SHOP, AFTER $E_{FIND\ MONEY}$}
$E_{PREPARE\ APPLE}$ = {PREPARE, APPLE, AT HOME, AFTER $E_{FIND\ APPLE}$}
$E_{PREPARE\ BANANA}$ = {PREPARE, BANANA, AT HOME, AFTER $E_{FIND\ BANANA}$}
$E_{PREPARE\ PORK}$ = {PREPARE, PORK, AT HOME, AFTER $E_{FIND\ PORK}$}
$E_{PREPARE\ LAMB}$ = {PREPARE, LAMB, AT HOME, AFTER $E_{FIND\ LAMB}$}
$E_{CONSUME\ APPLE}$ = {CONSUME, APPLE, AT HOME, AFTER $E_{PREPARE\ APPLE}$}
$E_{CONSUME\ BANANA}$ = {CONSUME, BANANA, AT HOME, AFTER $E_{PREPARE\ BANANA}$}
$E_{CONSUME\ PORK}$ = {CONSUME, PORK, AT HOME, AFTER $E_{PREPARE\ PORK}$}
$E_{CONSUME\ LAMB}$ = {CONSUME, LAMB, AT HOME, AFTER $E_{PREPARE\ LAMB}$}
$E_{CONSUME\ APPLE}$ = {CONSUME, APPLE, AT HOME, AFTER $E_{BUY\ APPLE}$}
$E_{CONSUME\ BANANA}$ = {CONSUME, BANANA, AT HOME, AFTER $E_{BUY\ BANANA}$}
$E_{CONSUME\ PORK}$ = {CONSUME, PORK, AT HOME, AFTER $E_{BUY\ PORK}$}
$E_{CONSUME\ LAMB}$ = {CONSUME, LAMB, AT HOME, AFTER $E_{BUY\ LAMB}$}

The event space, ES_{EAT}, could then be defined as containing the union of all of the above event elements as shown below:

ES_{EAT} =

{FIND, BUY, PREPARE, CONSUME, MONEY, APPLE, BANANA, PORK LAMB, AT HOME, AT SHOP, AUTOMATIC TELLER MACHINE, AFTER $E_{FIND\ MONEY}$, AFTER $E_{FIND\ APPLE}$, AFTER $E_{FIND\ BANANA}$, AFTER $E_{FIND\ PORK}$, AFTER $E_{FIND\ LAMB}$, AFTER $E_{BUY\ APPLE}$, AFTER $E_{BUY\ BANANA}$, AFTER $E_{BUY\ PORK}$, AFTER $E_{BUY\ LAMB}$, AFTER $E_{PREPARE\ APPLE}$, AFTER $E_{PREPARE\ BANANA}$, AFTER $E_{PREPARE\ PORK}$, AFTER $E_{PREPARE\ LAMB}$}

For this example, the Event Space Generator's processing results in the creation of eight specialised graphs, four for the EAT OWN FOOD Activity Digraph and four for the EAT BOUGHT FOOD Activity Digraph. The EMAI agent now has eight plans from which to choose from to satisfy the initial goal of EAT. Any one of the eight plans, when successfully executed will satisfy the goal. So the agent must choose one plan out of eight. How the EMAI makes a choice among these plans is explained in the next chapter.

6.5 Summary

An event describes a potential behaviour of an EMAI agent. It constitutes an action taking place with respect to a target object at a particular time in a particular context. An event is partially-ordered with other events and arranged into Activity Digraphs. These digraphs represent plans that an EMAI agent can follow to achieve its goals. For any one goal there may exist one or more Activity Digraphs that could be used to achieve the goal. Each Activity Digraph may contain one or more events. Activity Digraphs are stored in the EMAI Ontology as a series of conceptual graphs representing generic events and plans. The objective of the Event Space Generator is to take these generic graphs and specialise the concept types into elementary types from the Type Hierarchy. Having achieved this, the Event Space Generator can turn one goal into one or more specialised plans the EMAI agent can execute to achieve its goal. As only one plan needs to be executed to satisfy the initial goal, an EMAI agent is placed in a decision making situation. For this purpose, the Activity Digraph is viewed in terms of all its constituent elements.

The EMAI agent uses an affective decision making method to select among its plans. This concept and process is examined in the next chapter.

7. Single-Dimensional Affective Decision Making

There are two tragedies in life. One is not to get your heart's desire. The other is to get it.
- George Bernard Shaw

7.1 Introduction

Today's lifestyle means that people are constantly bombarded with decisions. Very rarely do these decisions involve right and wrong answers. Often they have to choose from a number of solutions that would all suit a purpose equally as well. For example, a person might decide to purchase a car. Having chosen the make and model, the final decision may come down to, *"Do I take the red one or the green one?"* In this case, the colour will not affect the ability of the car to fulfill the requirements for which it is being purchased. So a decision is made based on personal preference. Whatever the collection of past experiences that makes someone prefer green cars to red cars, it defines who they are as individuals, their likes and dislikes and determines their behaviour. If all people held the same attitudes they would all be driving the same car, eating at the same restaurants and using the same washing power. But this is not the case.

The domain-independent EMAI architecture allows its agents to assess their beliefs and satisfy their goals based on their *attitudes* towards other agents and objects in their environment. Before the implementation of the Theory of Reasoned Action (to be examined in this chapter), the EMAI architecture used an opportunistic approach to carrying out processes in order to satisfy an active goal. By establishing an algorithm by which the agent can apply a theory of attitude in decision-making, the agent becomes emotionally intelligent with the ability to make affective decisions and order intentions with human reasoning-like ability. For example, in the case study presented in the previous chapter, the agent had two Activity Digraphs that could be executed to achieve the EAT goal. From these, the Event Space Generator created an event space embodying eight individual specialised plans. The agent did not perform any assessment of the plans or explore how it might order them. Given an event space with a number of plans, each equally capable of satisfying the same goal, the agent must choose just one.

In this chapter, the Theory of Reasoned Action is used to generate a one-dimensional appraisal model for emotional decision making. The theory applies an *intention* to each plan of events based on the agent's beliefs. The intention toward a plan indicates an EMAI agent's preferences toward executing it.

This mechanism allows an EMAI agent to deal with situations where problem solutions are innumerable or time dependent. These mechanisms will be examined through four case studies that explore individual preferences and affective decision making.

7.2 Intentions, Attitudes and Subjective Norms

While there is no clear definition of attitude, there is popular consensus among social psychologists that the term refers to the general enduring disposition to feel positively or negatively towards an object, person or issue (Petty and Cacioppo 1996). By using a relevant collection of a person's attitudes, it is possible to closely predict the behaviour of that person where the same set of attitudes is applied (Fishbein and Ajzen 1975). Ajzen and Fishbein as noted by Petty and Cacioppo (1996) concluded this behaviour can be viewed as consisting of four key elements. They are:

1. the action being performed;
2. the target or targets that are the object of the action;
3. the context of the action, for example, where it is being performed; and
4. the temporal alignment of the action, for example, the time of day or month.

These four elements can be used to define a behavioural event (as per Chapter 6). All the elements of each of the events within all the competing Activity Digraphs form the *event space*.

In an empirical sense there is very little mention of attitude in the cognitive appraisal theories that have been integrated into intelligent agent architectures. The models (as discussed in Chapter 2) use valenced reactions to situations and events to determine an emotional state. By determining the effect that an event will have on an agent's goal, an emotional response can be generated. For example, the occurrence of a desirable event will synthesise *joy*. But how is an event evaluated as being desirable? What if more than one desirable event exists for the agent to select? How does it decide? To solve this issue the EMAI architecture implements a model of attitude. In general, attitudes serve as convenient summaries of a person's beliefs (Petty and Cacioppo 1996). A statement such as, "*I liked the movie*" sums up the speaker's feelings about an event conveniently for the listener rather than having the speaker explicitly state all the reasons why. They also help others to predict our behaviour. For example, if one was to say, "*I do not like horror movies*", their friend may predict they will not be going to see the Wes Craven movie *Scream II*.

It is this method of predicting behaviour that is the focus of the ideas on attitude that have been integrated into the EMAI architecture. Fishbein and Ajzen's *Theory of Reasoned Action* (Petty and Cacioppo 1996) specifies that the prediction of behaviour is based on a person's attitudes toward a behaviour in addition to the person's disposition toward the effects of social pressures. The person's attitudes are based on learned experiences and beliefs about a particular behaviour, the consequences of performing that behaviour and an evaluation of those consequences.

Petty and Cacioppo (1996) give an example of the Theory of Reasoned Action. The example is of selecting a favourite car where the colour and make of a car are considered as elements for assessment in determining the overall attitude towards a car. This example will be explored in relation to decision-making later in the chapter.

For the purposes of the research presented in this dissertation, attitude is calculated in order to prioritise the plans of events in an EMAI agent's event space. The Intention Generator (see Chapter 3) performs this function after the event space has been placed in order into the agent's schedule by assessing its motivational priority as explained in Chapter 4. Briefly, in the EMAI, the event space contains the plans the agent has formalised in response to an activated goal. This event space is scheduled for performance. The behaviour of the agent is determined by a plan in the event space being acted upon. Not all plans in an event space become the agent's behaviours, and when one plan satisfies the agent's goal, any other plans in the same event space, that would also satisfy the same goal, are no longer necessary and may be discarded. The Intention Generator of the EMAI architecture's Deliberate Area prioritises the event spaces firstly by assessing the urgency associated with the goal needing to be satisfied and secondly it orders the plans within the event space by preference using the Theory of Reasoned Action (Fishbein and Ajzen 1975). Although Fishbein and Ajzen outlined Expressions 7.3, 7.6, and 7.7, this dissertation attempts to clarify these expressions. It extends the use of the Theory of Reasoned Action to determine the *intention* that an EMAI agent has towards a plan of events by introducing and formalising the key concept of event space appraisal.

7.2.1 Attitude and its Definition in EMAI

The Theory of Reasoned Action is twofold. Firstly, the attitude towards performing the behaviour is calculated. This is determined by the summation of the agent's assessment of the consequences of performing the behaviour. Expression 7.1 states that each consequence C_i is equal to the product of the agent's belief b_i about the consequence being on a scale between *likely* and *unlikely* and the agent's evaluation e_i of that belief as being between *bad* and *good*.

$$C_i = b_i e_i \tag{7.1}$$

The attitude A, towards performing a behaviour B (A_B), can thus be determined by the number, n, of beliefs b_i, that performing the behaviour will lead to consequences C_i, with respect to the agent's evaluation e_i of those consequences. Thus:

$$A_B = \sum_{i=1}^{n} C_i \tag{7.2}$$

In other words, the above two formulae can be combined to form Expression 7.3:

$$A_B = \sum_{i=1}^{n} b_i e_i \tag{7.3}$$

In brief, the agent's attitude toward performing certain behaviours is determined by the set of beliefs the agent holds about the behaviour and how the agent assesses each of these beliefs. Before a behaviour is expressed or performed by the agent, it is referred to as an intention. To illustrate the attitude assessment of an intention, the example in Table 7.1, is given.

Determining Attitude from Beliefs and Evaluations i Consequences of Eating at a Five Star Restaurant	Belief (b) 3 = likely -3 = unlikely		Evaluation (e) 3 = good -3 = bad		$(b_i)(e_i)$
1. Excellent Service.	3	×	3	=	9
2. Have to Tip the Waiter.	1	×	-3	=	-3
3. Food would be of highest quality	2	×	3	=	6
4. Food would be expensive	3	×	-2	=	-6
$A_B = \sum_{i=1}^{n} b_i e_i$				=	6

Table 7.1 Determining an Attitude for a Restaurant

In this example, the attitude is being determined about the plan for *eating at a five star restaurant*. In order to calculate the attitude, all resultant consequences of eating at a five star restaurant are considered. Firstly, each one of the consequences is rated on the likelihood of them occurring. Secondly, the consequences, if they were to occur, are assessed on whether they would be good or bad. In Table 7.1 the seven point scale used in (Petty and Cacioppo 1996) is applied where –3 is *unlikely* or *bad* and +3 is *likely* or *good.*

7.2.2 Subjective Norm and its Definition in EMAI

Not only are attitudes determined by assessing beliefs and evaluations about behaviours, but a second factor should be considered. The predictability that an intention will become a behaviour must also be evaluated with respect to the agent's assessment of what others (artificial or human agents) will think about it performing the behaviour. This evaluation has been termed the subjective norm (Fishbein and Ajzen 1975). The subjective norm can be seen as the agent's sense of moral obligation. The agent assesses the value placed on its behaviour by an individual j (written as V_j) by calculating the product of the individual's endorsement of the belief b_j as being on a scale between *likely* and *unlikely* and the agent's motivation to comply m_j with the individual as being between *always* and *never*. This can be written as Expression 7.4:

$$V_j = b_j m_j \tag{7.4}$$

Thus, the agent's subjective norm to performing behaviour B, (SN_B), is a summation of all the V_j's as shown in Expression 7.5:

$$SN_B = \sum_{j=1}^{p} V_j \tag{7.5}$$

The above two formulae can be re-written as Expression 7.6:

$$SN_B = \sum_{j=1}^{p} b_j m_j \tag{7.6}$$

where b_j is the belief of individual j towards behaviour B, m_j is the motivation of individual j to comply to behaviour B, and p is the number of individuals concerned. The use of this calculation is demonstrated in Table 7.2. Here, the group of individual's that will have an influence on the agent's decision about eating as a five star restaurant are listed and how they will affect the agent's attitude is calculated. Firstly, each individual is assessed on how much the agent believes the individual will approve of the agent's decision, and secondly an assessment is made on how often the agent takes the advice of that individual. In the given example, the individual Tom is rated most likely to agree with the agent's choice to eat at a five star restaurant and is also rated as someone with whom the agent often complies. On the other hand the individual Katherine is considered to be someone with whom the agent believes would disagree with its choice to eat at a five star restaurant and is also someone from whom the agent rarely takes advice.

Important Referents Subjective norm about eating at a five star restaurant.	Belief (b) 3 = likely -3 = unlikely		Motivation to Comply (m) 3 = always -3 = never		$(b_j)(m_j)$
1. Tom	3	×	2	=	6
2. Harry	-1	×	-2	=	2
3. Katherine	-3	×	-2	=	-6
$SN_B = \sum_{j=1}^{p} b_j m_j$				=	2

Table 7.2 Determining Subjective Norm (*SN*) from b_j m_j

7.2.3 Calculating Intention

Having determined a formula for attitude (Expression 7.3) and one for the subjective norm (Expression 7.6), they can now be combined to produce the agent's intention, *I*, as shown in Expression 7.7:

$$I = w_1 \sum_{i=1}^{n} b_i e_i + w_2 \sum_{j=1}^{p} b_j m_j \qquad 7.7$$

where w_1 and w_2 are weightings added to represent that attitudes and subjective norms are not always evaluated equally in the formation of behavioural intentions. In other words, the intention of an agent towards a behaviour, I_B, is the combination of its attitude towards the behaviour and its subjective norm towards the same behaviour. For example, consider the plan *eating at a five star restaurant*, whose attitude and subjective norm were previously calculated. If the agent's attitude about the plan was weighted to be twice as important as the subjective norm (w_1 = 2 and w_2 = 1) the agent's intention toward the plan would be calculated to be 14 (6*2 + 2*1).

The higher the resultant intention the more favourable the behaviour and the more likely that an EMAI agent will choose to act on it. However, the intention value on a single plan or event is meaningless in isolation. It must be compared with intentions about other plans in the same event space. The intention value of 14 for *eating at a five star restaurant* would have to be compared with the intention of the agent's other choices of places to eat in the event space. Calculations of intention might be performed on *eating at a two star restaurant*, *eating at MacDonald's* or *eating at a hotel.* When intentions for each of these plans is calculated only then can the agent's chosen behaviour be predicted. Given the collection of intentions about places to eat, the agent would most likely select the choice with the highest value for intention.

7.3 Event Space Appraisal

Although the intention calculation equation is used in the Theory of Reasoned Action to holistically assess an intention, according to Petty and Cacioppo (Petty and Cacioppo 1996), it can also be used to assess attitudes towards elements (such as people, objects and issues) involved in intentional activities. These elements (as discussed in Chapter 6) collectively define the event space. For each of an EMAI agent's goals there exists an event space that may contain one or more plans of events and each event, *E*, will include an action, *a*, an object, *o*, a time, *t* and a context, *c* as shown in Expression 7.8.

$$E = \{a, o, c, t\} \tag{7.8}$$

The task of the Intention Generator, on receipt of the event space, is to appraise each element of each event and order the plans according to the EMAI agent's preferences. In this chapter, affective decision making is being examined through the use of a single-dimensional appraisal. The Intention Generator achieves a single-dimensional appraisal ordering of an event space by applying the Theory of Reasoned Action to the elements of the events in the event space and calculates the single-dimension of intention for each plan of events. Just as intention (using attitude and subjective norm) can be calculated about a particular behaviour using expression 7.7, intention can also be calculated about the elements of an event.

Let T_E represent the set that is the union of all the elements of event E. Therefore:

$$T_E = a \cup o \cup t \cup c \tag{7.9}$$

where e_E is an element of the set T_E, written as:

$$e_E \in T_E \tag{7.10}$$

The number of elements n, of an event E is:

$$n = |T_E| \tag{7.11}$$

For each element el, in an event E, the attitude is defined as Expression 7.12:

$$A_{el} = \sum_{i=1}^{k} b_i e_i \tag{7.12}$$

where k is the number of beliefs, b, the agent holds about the element and e is the evaluation of the belief.

The subjective norm for each element in the event is defined as Expression 7.13:

$$SN_{el} = \sum_{j=1}^{p} b_j m_j \tag{7.13}$$

where p is the number of beliefs that other agents have about el and m is the agent's motivation to comply with the beliefs of the others.

The intention, I, toward an element of the event can now be defined as Expression 7.14:

$$I_{el} = w_1 A_{el} + w_2 SN_{el} \tag{7.14}$$

where I_{el} is the intention the EMAI agent has about interacting with an element during the performance of a behaviour and the attitude and subjective norm are weighted as before.

The intention toward the event, I_E, is further calculated by summing the intentions toward all the elements involved with the event as shown in Expression 7.15:

$$I_E = \sum_{k=1}^{n} I_{el_k} \tag{7.15}$$

Furthermore, as the agent is making a decision about performing a particular plan of events which in most cases may not be a single event, the intention toward the entire plan, I_p, should also be calculated using expression 7.16:

$$I_p = \sum_{i=1}^{m} I_{E_i} \tag{7.16}$$

where m is the number of events in the plan.

Once the Intention Generator performs the single-dimensional appraisal on the events in the event space, the EMAI agent can begin executing the plans from most intended to least intended. After the execution of each individual event in a plan, the EMAI agent updates its beliefs and recalculates its attitudes and intentions towards all events, event elements, and plans in the event space. At this time, the plans may be reordered.

Based on the outcome of the event, E, the agent will assign a weighting w, to its attitude toward each of the elements, e. It should be noted at this point that the agent is not updating its intentions toward the elements in a holistic fashion. The reason for this is that the value of intention is calculated from the agent's own attitudes with the addition of the attitudes of others (the subjective norm). However, after an agent has performed an event and been involved with the event elements, the assessment the agent makes is based on the outcome of the event with respect to the agent itself. Therefore, the agent only updates its attitude towards events and event elements having completed and event. The subjective norm is not recalculated[13].

---

[13] While the attitude of others (subjective norm) may change, it is not the purpose of the EMAI agent to speculate how its own interaction with an event or event element may affect another's opinion of it.

As the weighting of an element and the agent's attitude with respect to an event are dynamic, the time, *t,* at which the intention is being calculated, must also be taken into consideration. Therefore, the attitude resulting from an event E, written as $A_{E,t}$ is calculated as Expression 7.17:

$$A_{E,t} = \sum_{e=1}^{n} w_{e,t} A_{e,t} \qquad 7.17$$

where *n* is the number of elements in the event *E*, and

$$0 \leq w_e \leq 1 \text{ and } \sum_{e=1}^{n} w_e = 1$$

Once an event has been completed, each of the elements, involved in the event, have their attitude values updated. These new attitude values can then be used for future calculations of intention toward events, event elements and plans. As the scale on which belief is rated is between *good* and *bad*, each time and event occurs, if it is successful the rating of belief, *b* for the event element is increased by 1 and if the event is not successful the rating is decreased by 1. There is no particular reason why the value of 1 has been chosen to increment and decrement the value for *b* except that it is the simplest way of updating the value. Therefore, the value of *b* at time *t* is equal to the initial value of *b* plus all the past evaluations of *b* as defined in Expression 7.18:

$$b_{el,t} = b_{el,0} + \sum_{i=1}^{m} \beta(b_{el,i}) \qquad 7.18$$

where *m* is the number of times the agent has come in contact with the element *el*, and β is a function that returns 1 if encounter, *i*, with *e* was good and –1 if the encounter was bad.

How the agent evaluates these beliefs, *e*, is also determined at each encounter with the element. The scale on which an evaluation is made ranges from *likely* to *unlikely*. For the purpose of the EMAI model this equates to a certainty rating and therefore, the evaluation of a belief is calculated by averaging the past likelihood of the belief. This is defined by Expression 7.19:

$$e_{b_{el,t}} = \left(\sum_{i=1}^{m} \alpha(b_{el,i}) \right) / m \qquad 7.19$$

where α is a function that returns 1 if the belief occurred on encounter i, with the element and 0 if it did not.

For example, assume an event E with one element A. At the beginning of the event, element A has values for attitude, subjective norm and intention as shown below. Assume for this example the weighting value is 1 ($w_A = 1$) for simplification.

$$A_{A,t} = 7, \qquad SN_{A,t} = 3, \qquad I_{A,t} = 10$$

where the value of $A_{A,t}$ was calculated using Expression 7.12 with two beliefs as follows:

$$\begin{aligned} b_1 &= 6, & e_1 &= 1 \\ b_2 &= 2, & e_2 &= 0.5 \end{aligned}$$

where e_1 indicates that b_1 is always likely and e_2 indicates that b_2 is likely half of the time. This would result in an intention for the event before execution as:

$$\begin{aligned} I_{E,t} &= w_{A,t} I_{A,t} \\ &= 10 \end{aligned}$$

Assuming after the event has occurred, the outcome is successful and A is still weighted the same, the attitude toward A, $A_{A,t}$, can be updated by reevaluating the beliefs involved with the element (in this example b_1 and b_2). As the event was successful both beliefs for the involved element have their values updated by 1 indicating the agent feels more favourably about each belief. In addition to this update, the evaluation of the beliefs will also be updated using Expression 7.18. If the belief about the element was instrumental during the event the value is set to $\alpha(b) = 1$ and if not $\alpha(b) = 0$. Given that during past involvement with b_1 the agent recorded evaluations of 1, 1, 1, 1, 1 and 1 and during the latest occurrence the agent recorded a 0, e_2 would now have a value of 0.86 ((1+1+1+1+1+1+0)/7). This gives b_1 a less likelihood of occurring next time compared with the previous values of 1 where b_1 had occurred each time in past events. Assuming that $\alpha(b_2) = 1$, and past evaluations of b_2 where 1, 0, 1, 0 , 1 and 0 (a previous value for e_2

of 0.5; likely to occur half of the time), e_2 will now equate to 0.57((1+0+1+0+1+0+1)/7) indicating the likelihood of b_2 has slightly increased. The resulting intention value calculated for the event would now be:

$$
\begin{aligned}
I_{E,t+\delta} &= b_{1,t+\delta} \times e_{1,t+\delta} + b_{2,t+\delta} \times e_{2,t+\delta} + SN_{A,t+\delta} \\
&= 7 \times 0.86 + 3 \times 0.57 + 3 \\
&= 6.02 + 1.57 + 3 \\
&= 10.59
\end{aligned}
$$

where the subjective norm for the element has not been changed. Therefore, due to the success of the past occurrence of the event E and the updating of the agent's attitude towards the element involved in E the intention towards E is stronger. This means the next time this event is schedule for execution it will be rated more highly.

7.4 Affective Decision Making

The EMAI architecture implements affective decision making to choose among multiple plans that it may execute in order to satisfy a goal. For each goal there exists an event space that may contain one or more mutually exclusive plans. The EMAI agent needs only to satisfactorily execute one of these plans to satisfy the corresponding goal. When there are more than one plan in the event space, the agent must make a choice. The difficulty in making the choice comes when a collection of plans in the event space can satisfy the goal equally as well as each other. The agent must therefore, make a choice based on its own preference. The concept of preference refers to attitude theory and calculating the intentions the agent has toward each event and subsequent plan. Preference may be different in different EMAI agents as it correlates directly with past experiences.

The Intention Generator prioritises event preferences in an EMAI agent with respect to performing plans that will satisfy an associated goal. Given a goal to satisfy and a number of plans (represented by Activity Digraphs) that will satisfy the goal, the agent calculates its preferred choice of plan from values (measures of intention) that it holds about the elements in the event spaces for those plans using the expressions defined in Section 7.3.

When combinatorial explosion becomes an issue, the event space for a goal is restricted to a specific number of plans. This means the Intention Generator will only determined priorities for plans in the current event space. If no suitable plan is identified after the agent has exhausted its options within the current event space, the event space is repopulated with a new set of Activity Digraphs and the process begins again. Initially the process of finding a suitable activity to achieve a goal can be extensive. However, as more plans are attempted and intentions calculated for each, the agent quickly builds an ordered list of plans that can best achieve associated goals. This process is demonstrated in Case Study 2 of this chapter, where plans are reordered by the agent to the point where the most successful plans for solving a goal are highly prioritised and therefore always attempted before others are considered.

The formulae presented in the previous section have been implemented as part of the EMAI architecture and are performed by the Intention Generator to order the agent's schedule using the Theory of Reasoned Action. This process allows the agent to make intuitive decisions and judge activities as liked or disliked in order to enhance its decision making and reasoning capabilities. As the EMAI agent is making a decision based on the single appraisal of the value of intention, this model is referred to as a single-dimensional model of affective decision making.

The remainder of this chapter presents four case studies to illustrate how attitude theory has been integrated into the EMAI architecture and its influence on the agent's reasoning and decision making capabilities. The first two case studies examine the separate parts of the strength of intention equation; attitude and subjective norm, respectively. The third case study examines the use of attitudes implemented in the EMAI architecture for decision making purposes. It compares the results obtained from the EMAI agent against the results collected from a human subject who is given the same decision making activity. The fourth case study examines a decision making dilemma proposed by Picard (1998) and demonstrates the use of the Theory of Reasoned Action via an EMAI agent to solve it using affective methods.

7.5 Case Study 1: Decision Making and Reasoning with Attitude

According to Velasquez (1998a) most theories about human performance in reasoning and decision-making can be classified into two different positions. Firstly, humans reduce all decisions to a list of possible outcomes and use logic to conduct a feasibility study that will give a list of best possible choices. Secondly, new decision situations are compared to past experiences and people tend to make choices that follow consistent patterns of previous choices. This case study examines the use of a combination of decision making techniques.

In many situations, it is impractical to analyse all possible courses of action and make a decision based on the measured plausibility of each. Picard (1998) suggests a model of decision making using intuition as a guide to reasoning. In her affective decision making scenario, high level or general decisions are made using measures of *bad* or *good* to assess options. While choices deemed to be bad or negative overall are not dismissed as alternatives, choices that are good or positive overall are explored further.

By the given definition, the weightings on the choices mentioned above are the attitude towards the options. Given a set of choices, if a direct attitude for the choice does not exist, the attitude towards each of the options can be calculated from other attitudes and beliefs of associated concepts involved with the choice. In this way, attitude measurements of atomic concepts can be used as internal states that determine the attitude value of complex situations. The method described above for determining overall attitude can also be implemented in emotional agents to enhance their emotion synthesis process. This topic will be explored further in Chapter 9.

The aim of this case study is to provide a worked example of the agent making an affective decision using the car example as introduced in Section 7.1. Assume that the agent has the Goal Hierarchy as shown in Figure 7.1 and a Type Hierarchy as shown in Figure 7.2.

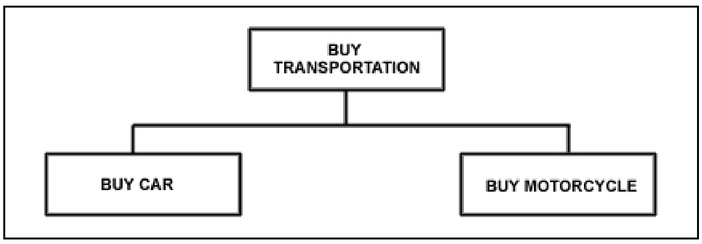


Figure 7.1 Goal Hierarchy for BUY TRANSPORTATION

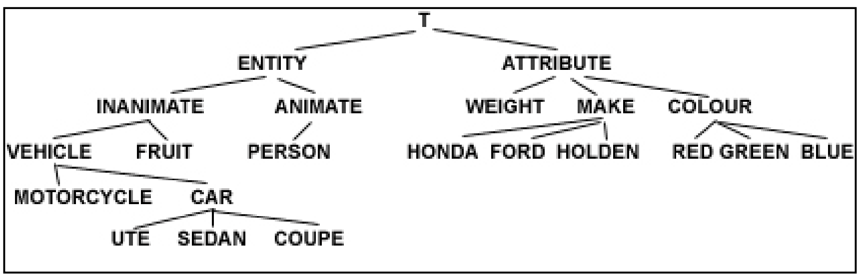


Figure 7.2 Type Hierarchy Showing CAR and COLOUR Types

When the BUY TRANSPORTATION goal becomes active so too will the BUY CAR and BUY MOTORCYCLE subgoals. To illustrate the way in which multiple activities are ordered by preference, the BUY CAR subgoal will be the focus of this example as there are multiple CAR subtypes for the agent to choose among. At this point it is important to remember that either purchasing the UTE, SEDAN or COUPE would all satisfy the agent's goal to BUY TRANSPORTATION. The conceptual graph that explains the associated Activity Digraph for the BUY TRANPORTATION goal may look similar to this:

```
[BUY] –
    (AGNT) -> [PERSON: EMAI]
    (SRCE) -> [PERSON: Joe]
    (OBJ) -> [VEHICLE] -> (ATTR) -> [COLOUR]
    (INST) -> [MONEY]
```

By restricting this graph to specialisations from the Type Hierarchy, the conceptual graph for the BUY CAR goal would look like this:

```
[BUY] –
    (AGNT) -> [PERSON: EMAI]
    (SRCE) -> [PERSON: Joe]
    (OBJ) -> [CAR] -> (ATTR) -> [COLOUR]
    (INST) -> [MONEY]
```

As this subgoal is the last in the agent's Goal Hierarchy, the agent will treat this graph as atomic and representative of the exact activity that it is to perform. This graph and the Type Hierarchy are passed to the Event Space Generator where event spaces are constructed using atomic elements from the Type Hierarchy to further specialise the subgoal BUY CAR into specific events. The Event Space Generator will produce conceptual graphs describing every different combination of atomic types. In this example, the Event Space Generator will produce nine graphs, three of which are shown below:

```
[BUY] –
    (AGNT) -> [PERSON: EMAI]
    (SRCE) -> [PERSON: Joe]
    (OBJ) -> [UTE] -> (ATTR) -> [GREEN]
    (INST) -> [MONEY]
```

```
[BUY] –
    (AGNT) -> [PERSON: EMAI]
    (SRCE) -> [PERSON: Joe]
    (OBJ) -> [UTE] -> (ATTR) -> [BLUE]
    (INST) -> [MONEY]

[BUY] –
    (AGNT) -> [PERSON: EMAI]
    (SRCE) -> [PERSON: Joe]
    (OBJ) -> [SEDAN] -> (ATTR) -> [RED]
    (INST) -> [MONEY]
```

The agent will continue producing graphs until it has created events for each combination of CAR type and COLOUR type. In this case, the nine separate graphs will constitute the event space.

To keep this example simple, the agent's decision on buying a car will be determined by its attitude toward each car based on the car's colour and make only. Assume the agent is programmed with the information that the car's make is weighted as 4 times more important than its colour when performing the activity of BUY CAR. Table 7.3 provides an example of possible weights, *w*, and attitudes, *A*, with which the agent may be programmed, that could be used to determine an attitude for each combination.

Cars											
Colours				*w*	*A*		*w*	*A*		*w*	*A*
	w	*A*	Honda	4	5	Ford	4	3	Holden	4	2
			Overall Attitude			**Overall Attitude**			**Overall Attitude**		
Red	1	5	(1x5) + (4x5) = **25**			(1x5) + (4x3) = **17**			(1x5) + (4x2) = **13**		
Green	1	2	(1x2) + (4x5) = **22**			(1x2) + (4x3) = **14**			(1x2) + (4x2) = **10**		
Blue	1	1	(1x1) + (4x5) = **21**			(1x1) + (4x3) = **13**			(1x1) + (4x2) = **9**		

Table 7.3 Hypothetical Weights and Attitude Values Assigned to 9 Car and Colour Combinations

In this example, the red Honda is the favourite car and the blue Holden is the least favourite car. When these values are used by the Event Space Generator to order the cars for the BUY CAR activity, the red Honda will appear first in the agent's schedule and the agent would try to perform the activity of buying a red Honda. Should this activity be unsuccessful, the agent would go to its next most favourite choice of the green Honda. If the agent had car types in its Type Hierarchy that it had no attitudes about, these would be given the value of $A = 0$ and overall attitudes would be calculated for them and they would be included in the order of activities. If the agent had a negative attitude towards a type of car this would result in an overall negative attitude and the car type that the agent had no attitude about would be ranked higher in the activity list.

The next case study compares human attitudes used in decision making against a simulation run using EMAI.

7.6 Case Study 2: Human versus Computer in Attitude Choice Making

7.6.1 Introduction

For this attitude experiment, the agent has been given one simple goal. The agent is preprogrammed with the knowledge that there are 10 different ways in which to satisfy this goal. The agent is then asked to satisfy this goal 100 times. Data is gathered about the choices the agent makes. This same goal was also given to a human subject. At this point it may be simpler to insert the narrative given to the human subject at the outset of the experiment in order to explain it.

> *You are in a plane crash and find yourself stranded in the Jungle. It is 9am on a clear sunny morning. You have to find your way back to civilization. In the back of the plane you find 10 survival packs. Each survival pack contains a number of items to help you with your journey. Each pack weighs the same and is comfortable to carry. You have been given a list of the packs and their contents. Each item in the pack can be responsible for your success or failure. If just one item fails when you try to use it then your mission fails.*
>
> *When asked, select a pack by entering the number. Each item in the pack will be examined to see if you can use it.*
>
> *To complete a game you have to keep trying to escape the Jungle by selecting a mission survival pack.*

Each survival pack represents one of ten ways to attempt a task that will satisfy the goal of escaping from the Jungle (alive). The list given to the human subject inventories the ten packs and their contents. The packs randomly contain up to five items named *a, b, c, d* and *e*. For example, survival pack 1 may contain items *a*, *c* and *e* and survival pack 5 may contain items *a* and *b*. Each item's alphabetical designation refers to the same item in each pack. For example, item *a* in pack 1 is the same as item *a* in pack 5. It was chosen not to associate the items with real world objects in an attempt to eliminate any preconceived attitudes that could influence the human subjects decision making process.

The agents (both artificial and human) had to try and escape from the jungle 100 times. Each of these times was called a *game*. To complete a game the agents had to find a pack that helped them to successfully escape. Before the start of the experiment, each item in the pack was given a static uniformly random probability rating of failure. When either agent chose a pack, the success or failure of that pack was calculated by assessing a success or failure on each item in the pack. If one item in the pack produces a failure rating, the whole task was rated as a failure. The agents were then made aware of which items were responsible.

As the experiment ran, data was gathered about the number of packs chosen in a single game before a success was found, the order in which the packs were chosen and how the agents felt about the items. The aim was to determine if the attitude model implemented in the agent performs the task in a similar manner to a human.

7.6.2 Results

The initial results were encouraging. Figure 7.3 is a box plot showing a 95% confidence interval for the mean number of survival packs chosen in a single game before a successful one was found. As can be seen there is statistically no difference between the EMAI agent and the human subject.

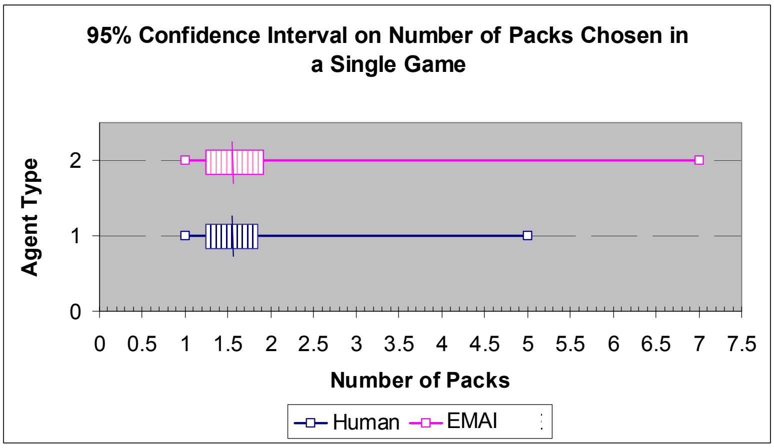


Figure 7.3 95% Confidence Intervals on Mean Number of Pack Choices per Game

Further investigation showed different results for the choice patterns made by the two agents. Figures 7.4 and 7.5 display the sequence of choices the agents made in selecting a survival pack. The dark points show each pack chosen and the white dots represent when the chosen pack was successful. Although the packs chosen were quite different, similar patterns can be seen in each graph, as explained below.

Firstly, plateaus are apparent where the agents selected a pack and the pack was successful. This caused the agents to select this same pack first for the very next game. For the human this is mostly evident for packs 7, 2 and 1. For the EMAI agent they were packs 9, 4 and 3. For each of the agents, these were the packs that were successful most of the time and therefore, the agents returned to use them again.

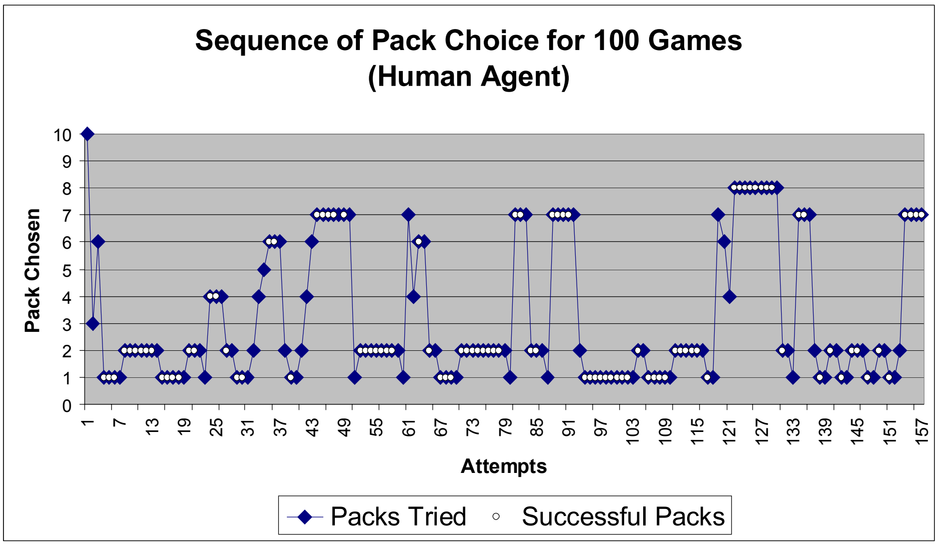


Figure 7.4 Human Agent Pack Choice Sequence

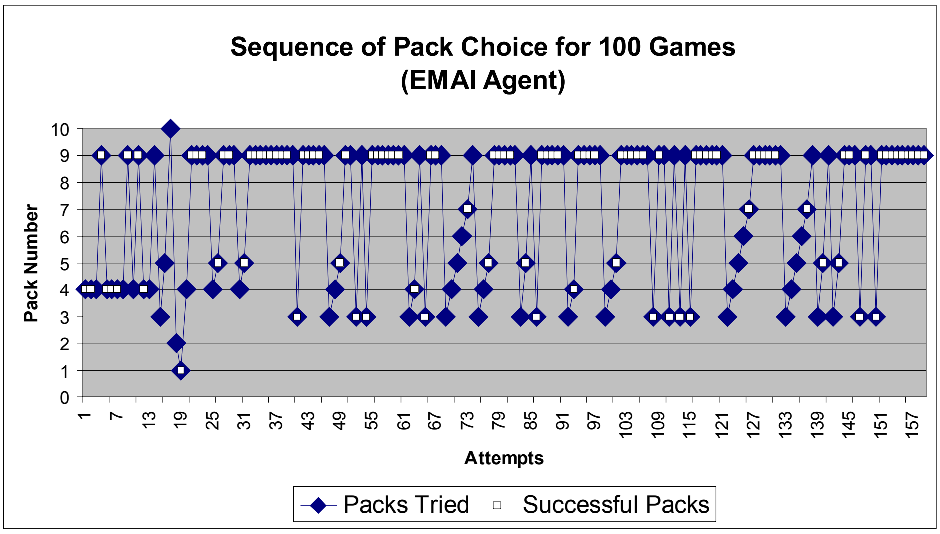


Figure 7.5 EMAI Agent Pack Choice Sequence

Secondly, there were set sequences of pack choice. Should a chosen pack have failed, each agent had a *backup or alternative* pack from which to select. For the human agent this sequence consisted of the packs identified in the graph plateaus. One emerging pattern was to select pack 1 should pack 2 fail and vice versa. Pack 7 was also chosen several times when packs 1 or 2 failed or both and a sequence of packs 2 then 1 were chosen should pack 7 fail.

In the case of the EMAI agent the patterns were more obvious. This should be expected, as the agent is a logical machine basing its reasoning on discrete values. This agent's *favourite* sequence was 9, 3, 4, 5, 6, and 7.

Thirdly, at the initial stage each of the agents tried several tasks before establishing a pattern. These initial tries also had an effect on future choices. In the case of the human agent, at the beginning of the graph it can be seen that several tasks were tried. The failure of these tasks had an impact on other pack choices made by the human subject. Packs 10 and 3 were never chosen again and pack 6 was chosen again only after a succession of other failures.

The graph produced by the EMAI agent was slightly different. In the beginning the agent was fortunate enough to select packs 4 and 9, which had a high success rate, so it continued using them. Once they both failed in succession around the 15th game, it can be seen that the agent tried a number of other packs before finding one that worked. During this time it tried three packs that it never chose again. These were packs 2, 10 and 1 (even though 1 was successful).

Finally, in each case there were packs that were not chosen at all. The human agent did not select pack 9 and the EMAI agent did not choose pack 8.

7.6.3 Discussion

While computers can outperform humans in many logical problem-solving situations, there are often times when optimal solutions to problems cannot be enumerated with the available resources or within a finite time. Humans have the remarkable ability to respond in situations with limited knowledge, limited memory and comparatively slower processing speeds (Picard 1998).

This case study shows the results of using attitude prediction theories to produce humanistic reasoning patterns in the EMAI architecture. In this experiment neither agent knew the success rates of the survival packs. Only by using the packs could the agents form an opinion about them. In both agents, the packs that succeeded most often had a higher attitude formed towards them. This caused these *favourite* packs to be chosen time after time.

In the human agent, a popular pack did not lose favour with the subject after several failures. This too can be seen in the artificial agent. The information integration formulae used to update EMAI's attitude towards the packs increases the attitude value for a pack each time it succeeds. After many successes, the agent can have a high attitude value for a pack. This attitude value may be much higher than any other pack. When such a pack fails, the agent decreases its attitude value towards the pack. However, when the attitude value is very high, one failure of a favourite pack will not cause it to become unpopular. Depending on its attitude value relative to other packs, the pack may continue to remain the favourite.

The exercise in this experiment was to find a pack that helped the agent escape from the jungle. The agents did not have to identify the optimal pack. For *interest*'s sake, the pack with the least probability of failure was pack 7. The human subject identified this pack although it was ranked as third most used, as can be seen in Figure 7.6. The EMAI agent chose pack 7 rarely. Agents identified a small number of packs that could help them achieve their task in the least number of choices.

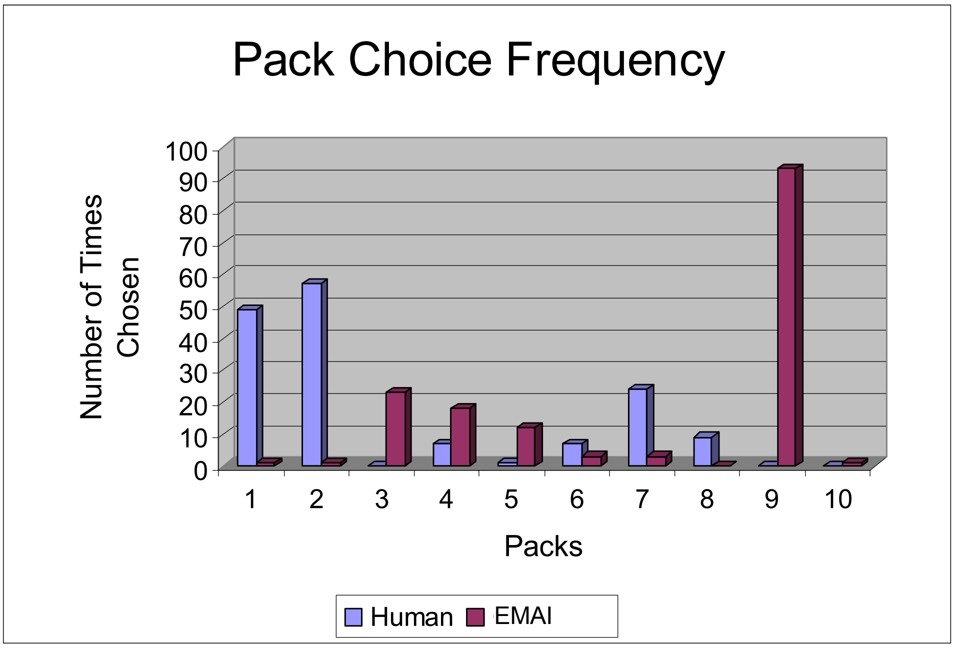


Figure 7.6 Pack Choice Frequency for Agents

In many real life situations, human's tend to select a solution that works within specified parameters, and if it is successful, there is often no need to look for another solution. For example, a person may drive to work the same way everyday. The route may not be the optimal route (the shortest, a better road or with less traffic) but for one reason or another, this route has worked successfully for the person day after day. Imagine that one day the route is closed due to road works. The person will find another route that works for then. When the road works are finished, they may return to the old route, or if they found the newer route to be better in some sense of the word, they may stick with it.

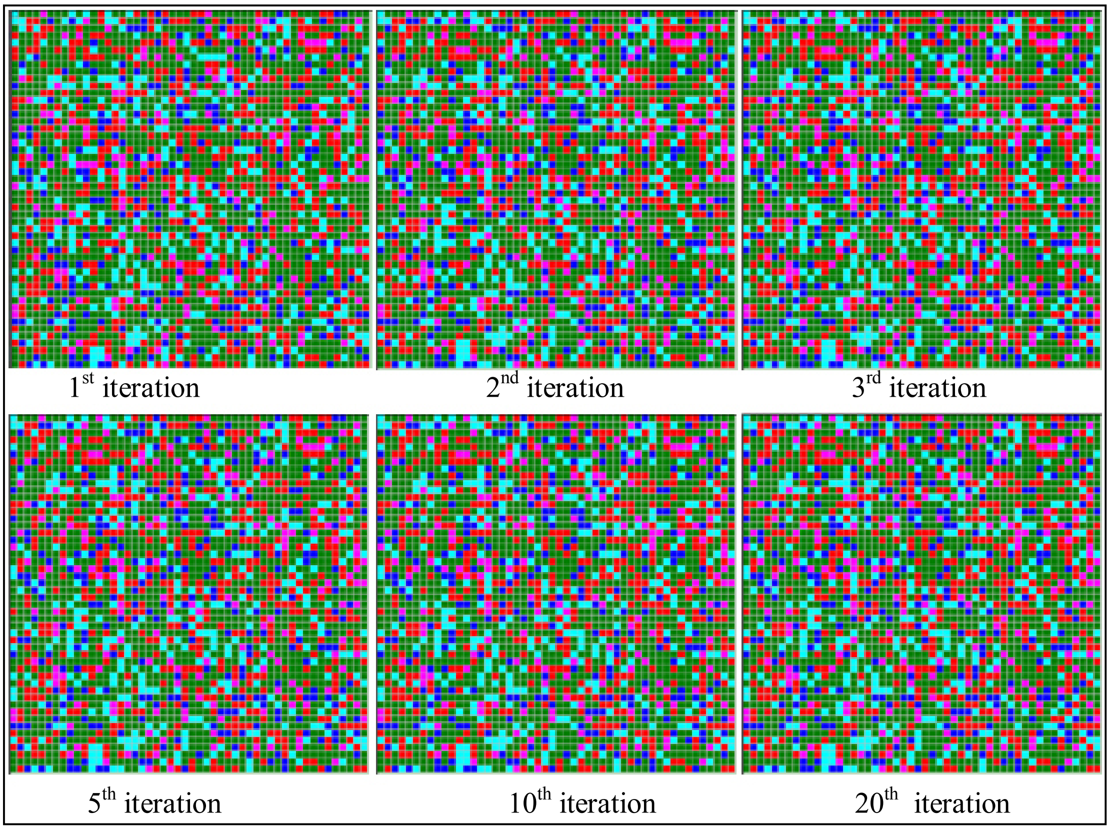
In this experiment, it was not possible for either agent to initially access the success probabilities of the packs without using them. Only through experience can the agent form an attitude toward a pack. The more games the agents played, the more information they had about a pack's success probability. This is an example of a problem where it was not possible to calculate an optimal solution before applying one. Each agent was not given the opportunity to examine each of the packs before making a selection. Therefore, the agents did not know which pack would have had the highest probability of success for each game. In each game they could make only one decision and as the games progressed they were able to use their experiences in past games to make a decision.

The next case study continues to examine attitude formation and its use in selecting an activity but also includes the use of subjective norms.

7.7 Case Study 3: Restaurant Selection via Peer Pressure

In order to demonstrate how attitude values and subjective norms can be used by an agent to make an affective decision, this case study examines a simple cell automata simulation. In this simulation there are 2500 cells that represent 2500 individuals. Each decision making event in the simulation is referred to as an iteration. At each iteration, each individual makes a decision among five choices. The outcome from each decision has the same probability of success. The decision making process is the same one that is integrated in the EMAI architecture's Intention Generator.

To visualise the experiment consider a population that has a choice of five restaurants. Initially the members of the population have not eaten at any of the restaurants and therefore, they have no preconceived attitudes about them. In each iteration, an individual selects a restaurant based on how it *feels* about that restaurant. On the first iteration there only exists neutral attitudes about each and so the individuals just pick one based on the FIFO method used by EMAI's Intention Generator (see Section 3.5.1). Having chosen a restaurant at which to eat, the simulation is programmed with a 10% chance that an individual will have a bad experience. If this occurs the attitude held by the individual about the restaurant decreases otherwise its attitudinal beliefs about the restaurant are reconfirmed and its attitude becomes stronger. Figure 7.7 displays progressive iterations of the simulation with each cell's restaurant preference represented by a different colour. In this first simulation the individual cells change their preferences based on failure or success of the restaurant only. Subjective norms have not been calculated.



Notice that in this simulation, favourite restaurant selection is randomly distributed among the cells and very little fluctuation takes place when individuals change their restaurant preferences.

As intentions are not only governed by attitudes, in the second simulation, each individual makes a choice of which restaurant to eat at as well as gathering the attitudes about the restaurants from each of the individuals in the neighbouring cells and uses these to calculate an overall attitude. Figure 7.8 displays several iterations of the simulation that includes subjective norms.

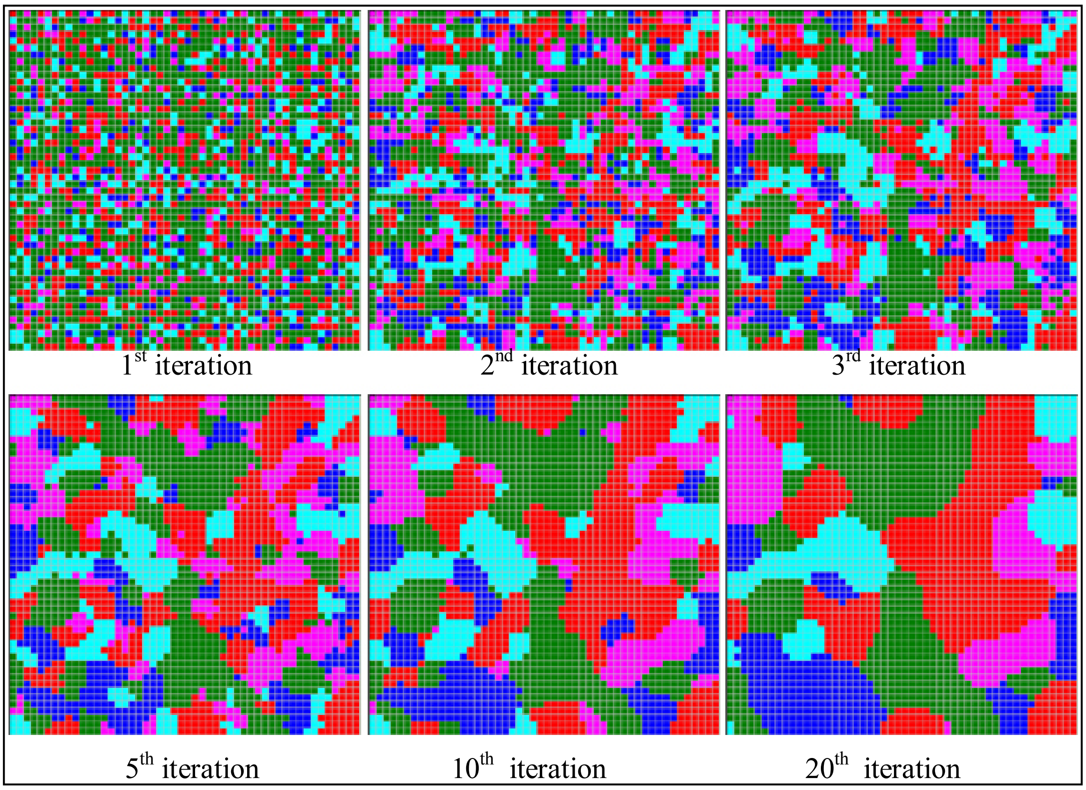


Figure 7.8 Cellular Automata of Favourite Restaurants with Social Norms

Not unexpectedly, from random beginnings the individuals in close vicinity formulate the same attitudes about their favourite restaurant. Eventually just one restaurant wins out as the favourite of the whole population. In the above simulation, each member of the population had an equal influence over other members. This of course does not occur in a real population and the Theory of Reasoned Action allows for some individuals to have a stronger influence over others. The simulation was run again and this time several individuals with a stronger influence over the surrounding population were added. Black cells in Figure 7.9 represent these individuals.

The number of influential individuals had a significant impact on the stability of favourite restaurants. As with the previous simulation, where cells with the same restaurant preferences began to group together, the influential individuals have groups form around them. Groups without influential individuals also form in the beginning of the simulation. However, in the end the groups with an influential individual as a member grew the largest and instead of one restaurant winning out overall, as was the case in the previous simulation, the influential individuals added stability to the population and separate groups remained.

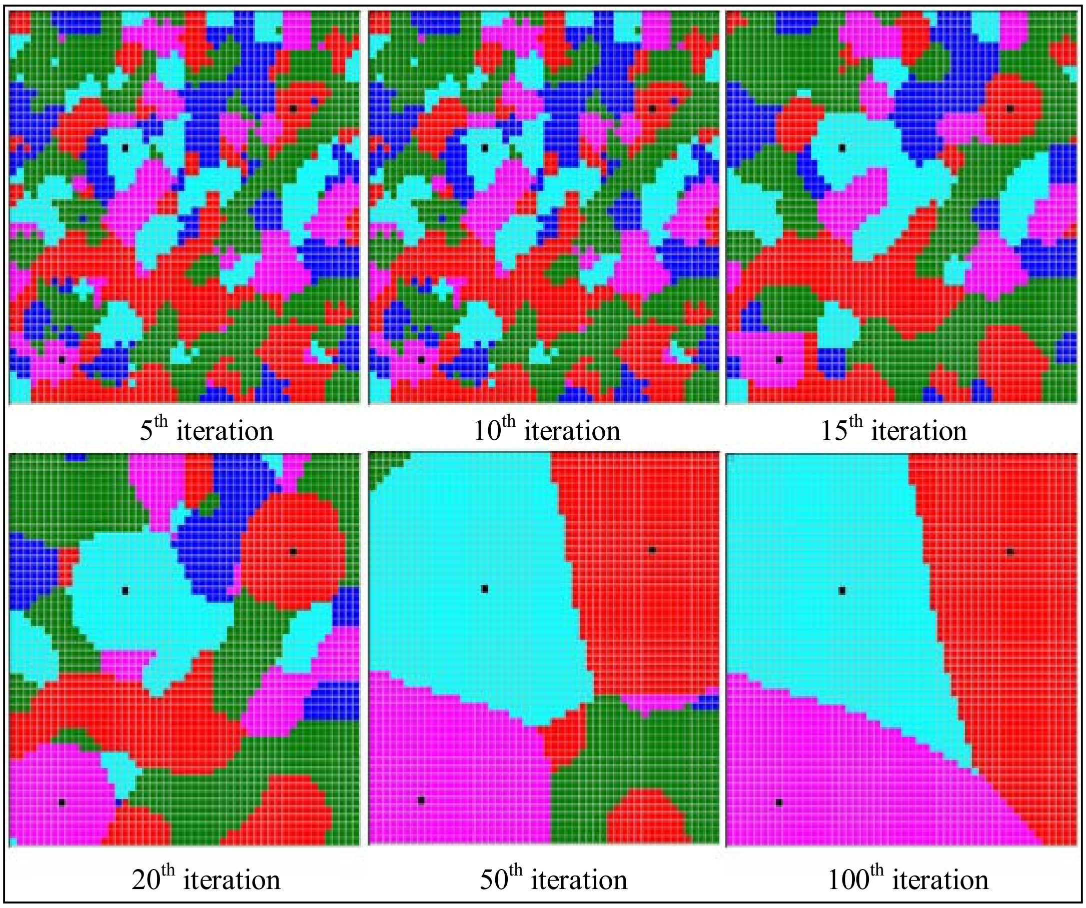


Figure 7.9 Favourite Restaurants with Influential Individuals

This case study has illustrated the mechanisms that are part of the EMAI architecture that update attitude values based on success and failure of activity choices with social influences. The last case study in this chapter examines a *challenge* initiated by Picard in her book *Affective Computing* (Picard 1998). How can attitude theory be used to produce an artificial agent acting within the parameters she supplies?

7.8 Case Study 4: Picard's Albert

7.8.1 Introduction

This case study whose results are published in (Baillie et al. 2000a), examines an example of affective decision making suggested by Picard (1998):

'Albert, a very busy scientist, has a beloved eight-week-old boy, and is trying to decide how to provide for his son while he works during the day. He does not know any family members or friends who could help. He acquires lists for three kinds of

day care providers: a list of ten nanny referral services, a list of 145 licensed family car providers, and a list of 24 day care centers located nearby. He contemplates posting notices in newspapers and on bulletin boards. Albert loves his son, and want's to choose the best care for him. He needs a care-provider within a month. Albert is a highly rational man; how does he decide what to do?'

In Picard's proposed solution there is a narrative that illustrates how each of the options available to Albert are weighed against each other by his valenced reactions towards them. Albert's feelings towards the options are based in associated concepts that Albert already has experienced. Picard sums up Albert's decision making process thus:

'Albert did not search all the possibilities, but he searched until he ran into either negative feelings or logical constraints, and then he stopped and tried something else. He continued this strategy – exploring possibilities that felt reasonable and good, and modifying his criteria as he accumulated more information............Before his time ran out, he arrived at a decision that combined logical constraints with weighing the good and bad valences associated with his options. Emotions played an integral part not only in his final decision, but also in his process of gathering information.' (Picard 1998)

From the point of view of the research in this dissertation, the valenced reactions that Picard discusses are attitudes and each of the possibilities represents an event space. To evaluate his options, Albert forms an attitude about the new event based on overlapping events about which he has already formed attitudes. For example, one of Albert's options is to advertise, however, Picard states that Albert believes that advertising attracts weirdoes. From this, it may be assumed that Albert (or someone he knows) has had a *bad* experience with advertising and he is applying that belief to this new situation. The following is an examination of one of Albert's options and an explanation of how he formulated an intention toward it.

One of Albert's options is to choose from a list of nannies. In order to calculate Albert's intention toward this option two parts need to be calculated; his attitudes and subjective norms toward the behaviour options. Table 7.4 displays Albert's beliefs (taken from Picard's narrative solution) about using a nanny for day care and his evaluation of each of these beliefs. The beliefs are extracted from the elements that Albert considers to be in the event space for this behaviour. A seven point scale, similar to the one used in (Petty and

Cacioppo 1996) for assessing evaluations and beliefs, (such as +3 for good or likely and -3 for bad or unlikely) has been applied to give hypothetical values.

Consequences of Hiring a Nanny for Day Care for Son.	Belief (b) 3 = likely -3 = unlikely		Evaluation (e) 3 = good -3 = bad		$(b_i)(e_i)$
1. nanny would come to the house during the day.	3	×	3	=	9
2. A nanny would be expensive.	3	×	-1	=	-3
3. nanny-referral services want fees up front	3	×	-2	=	-6
4. nanny could be with us for many years	2	×	3	=	6
5. nanny may not be available to start within four weeks - same newspaper ads for nanny positions repeated over several weeks*	2	×	-3	=	-6
6. nannies can be abusive - a documentary he saw on TV*	-2	×	-3	=	6
$A_B = \sum_{i=1}^{n} b_i e_i$ * how these beliefs were formed.				=	6

Table 7.4 Determining Attitude (A) from b_i and e_i

Given these are all of Albert's beliefs that are relevant to the hiring of a nanny and if this is how he assessed them, Albert's attitude towards the option of hiring a nanny is calculated to have a value of 6. To fully assess Albert's intention to hire a nanny, his belief regarding others opinions about this action and his motivation to comply with them must also be examined. Picard's narrative is not clear on the influence that other's opinions have on Albert's decision making, so hypothetical values are given to calculate his subjective norm in Table 7.5. Belief is again rated on a 7 point scale where 3 infers the person is likely to approve of Albert’s hiring a nanny and -3 infers they do not. Motivation to comply is rated in a similar manner where 3 implies Albert generally does what this person wants him to do and -3 implies Albert hardly ever complies with this person's wishes.

From the examples given, it is calculated that Albert's subjective norm about hiring a nanny has a value of 8. Assuming that Albert weights his attitude to be twice as important as his subjective norm ($w_1 = 2$ and $w_2 = 1$), his intention to hire a nanny is calculated to be 20. Although this would suggest a high motivation in Albert to perform this intention, the other options that Albert has available to him should also be assessed and attitudes compared. A similar method would be applied to his other options for determining their attitude values. Ranking these options in order of intention would give further insight into an improved prediction of Albert's choice of behaviour regarding the search for a day care provider.

Important Referents	Belief (b) 3 = likely -3 = unlikely		Motivation to Comply (m) 3 = always -3 = never		$(b_i)(m_i)$
1. his mother Anna	3	×	2	=	6
2. his friend Joe	-1	×	-2	=	2
3. his brother Phillip	-3	×	0	=	0
$SN_B = \sum_{j=1}^{p} b_j m_j$				=	8

Table 7.5 Determining Subjective Norm (SN b_j m_j

The next section examines the EMAI agent while simulating this decision making process.

7.8.2 Simulating Albert

The goal, type and relation hierarchies necessary to produce a partial simulation of Albert have been programmed into EMAI agent. Instead of simulating all of Albert's choices as goals, a partial selection of three goals has been chosen as these provide a complex simulation and illustrate the nature of the agent without the need for including all of Albert's choices. *Please note that different colours have been used throughout to distinguish different event graphs.*

Initially Albert is given the goal of finding day care for his son. Since the agent that is simulating Albert's behaviour has never performed a plan to solve this goal before, it has all of its attitude values set to zero. This means the agent has no preferences toward any of the subgoals and related process plans from which it can choose. As the agent gathers information about each process by analysing the elements of the associated event, it sorts and stores its preferences about each process choice in its belief system. As can be seen in Figure 7.10, the agent is simulating the FIND CARER FOR SON goal as a choice of three subgoals, each of which is mutually exclusive, to achieve the highest goal. At atomic goal level, each of the three choices becomes a two or three-fold plan. Firstly, the agent will perform the ASK ABOUT DAY CARE, ASK ABOUT NANNY and ASK FRIEND activities and secondly, the agent attempts a procurement type process. This is represented in the PLACE IN DAY CARE goal as the subgoal INQUIRE ABOUT VACANCY.

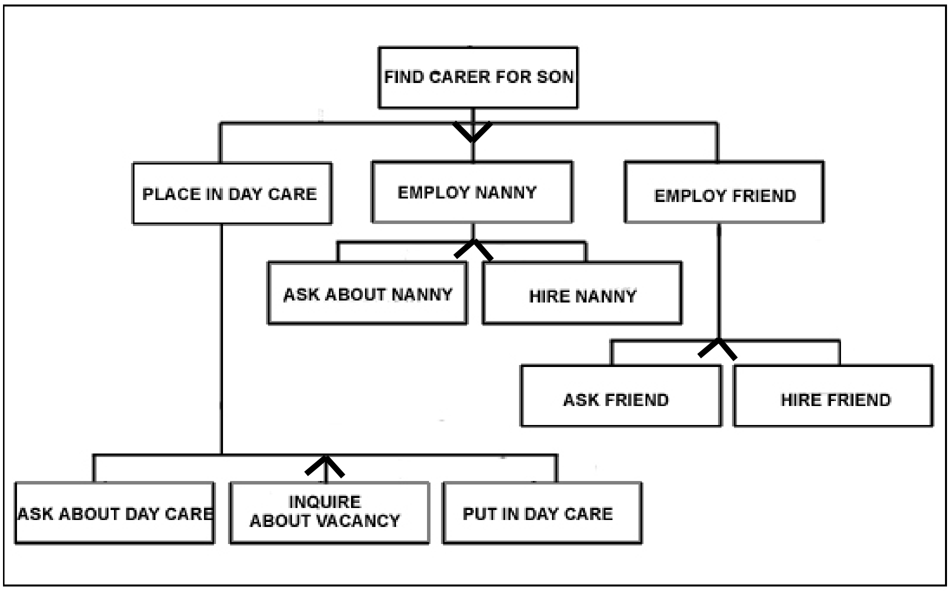
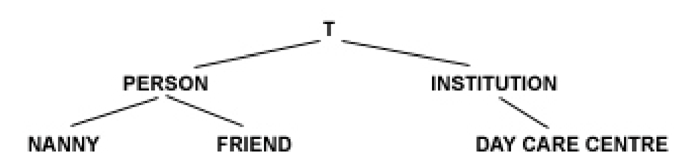


Figure 7.10 A Partial View of Albert's Goal Hierarchy

When the FIND CARER FOR SON goal is triggered the Activity Digraphs that explain the process plans that satisfy this goal are sent to the Event Space Generator along with the agent's hierarchy of concept types. The Type Hierarchy used for this simulation is shown in Figure 7.11. This hierarchy provides the agent with sufficient knowledge to find concepts that can be used by the Event Space Generator for conceiving new event spaces to explore with regards to the FIND CARER FOR SON goal.



Type	Individual
NANNY	Joe
	Jenny
FRIEND	Mum
	Harry
	Mary
DAY CARE CENTRE	Little Pumpkins
	Sunny Day

Figure 7.11 Albert's Type Hierarchy for Caregiver Types

For example, the conceptual graph constructed from the agent's set of Activity Digraphs (see Figure 7.12) that defines the goal is shown below:

```
[ASK ABOUT NANNY] –
    (AGNT) -> [AGENT: ALBERT]
    (OBJ) -> [NANNY]
    (INST) -> [TELEPHONE]
    (RCPT) -> [FRIEND]
    (FBS) -> [HIRE NANNY] –
                (AGNT) -> [AGENT: ALBERT]
                (OBJ) -> [NANNY]
```

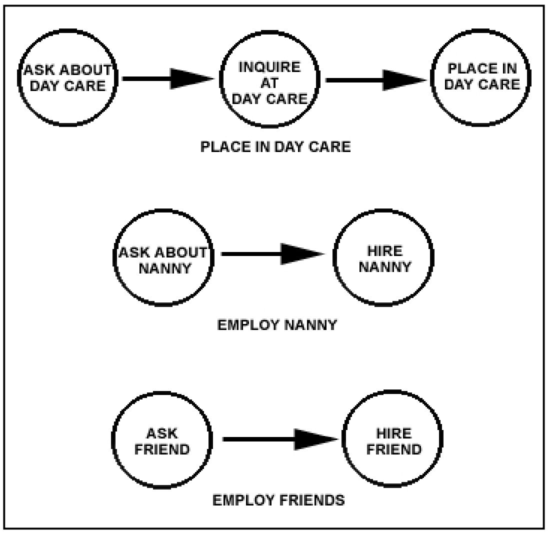


Figure 7.12 Albert's Activity Digraphs for the FIND CARER FOR SON Goal

Given the Type Hierarchy of Figure 7.11, the Event Space Generator would create two events that could satisfy the EMPLOY NANNY goal with appropriate conceptual graphs built by restricting the concept NANNY to the individual Joe for the first graph and Jenny for the second graph thus generating the graphs below:

```
[ASK ABOUT NANNY] –
    (AGNT) -> [AGENT: ALBERT]
    (OBJ) -> [NANNY: Joe]
    (INST) -> [TELEPHONE]
    (RCPT) -> [FRIEND]
    (FBS) -> [HIRE NANNY] –
                (AGNT) -> [AGENT: ALBERT]
                (OBJ) -> [NANNY: Joe]
```

and

```
[ASK ABOUT NANNY] –
    (AGNT) -> [AGENT: ALBERT]
    (OBJ) -> [NANNY: Jenny]
    (INST) -> [TELEPHONE]
    (RCPT) -> [FRIEND]
    (FBS) -> [HIRE NANNY] –
                (AGNT) -> [AGENT: ALBERT]
                (OBJ) -> [NANNY: Jenny]
```

Furthermore, the Event Space Generator would recognise the type concept of FRIEND and its associated referents and would generate further events where each of the FRIEND referents in the EMPLOY NANNY graph are specialised to the individuals Mum, Harry and Mary. This would lead to the creation of six separate Activity Digraphs (two NANNY referents by three FRIEND referents) for the goal of EMPLOY NANNY. They would look similar to the following graph:

```
[ASK ABOUT NANNY] –
    (AGNT) -> [AGENT: ALBERT]
    (OBJ) -> [NANNY: Jenny]
    (INST) -> [TELEPHONE]
    (RCPT) -> [FRIEND: Harry]
    (FBS) -> [HIRE NANNY] –
                (AGNT) -> [AGENT: ALBERT]
                (OBJ) -> [NANNY: Jenny]
```

Each of these Activity Digraphs would consist of two distinct events. For both the EMPLOY NANNY and EMPLOY FRIEND goals there exist two subgoals that are satisfied by processes in the Activity Digraphs as shown in Figure 7.12. The EMPLOY NANNY goal can be satisfied by executing ASK NANNY followed by HIRE NANNY and the EMPLOY FRIEND goal can be satisfied by executing the ASK FRIEND followed by EMPLOY FRIEND. Due to the way in which the agent selects a process, an Activity Digraph may not be followed through to its completion. The agent could complete the first part of a digraph and reevaluate its attitudes toward that part. This in turn could influence the agent's attitudes toward the rest of the digraph. If the agent considered the rest of the digraph to have a lower attitude value than parts of other digraphs, the agent would abandon the current digraph for another. This is in keeping with Picard's narrative solution for Albert's decision-making process where she states:

> *'Albert did not search all the possibilities, but he searched until he ran into either negative feelings or logical constraints, and then he stopped and tried something else.'*

To produce the type of behaviour mentioned above, the agent evaluates each event in an Activity Digraph individually. In the case of the EMPLOY NANNY goal and associated Activity Digraphs, that in this simulation number six, the Event Space Generator takes each of the six plans, specialises them into graphs that represent individual events and determines the referent (if any) to create specific events. As there are six plans that could satisfy the EMPLOY NANNY goal, each with two generic events, the associated event space would contain 12 specific events thus:

$$E_1 = \{[ASK],[NANNY : Joe],[FRIEND : Mary],[NULL]\} -> E_2 = \{[HIRE],[NANNY : Joe],[NULL],[NULL]\}$$
$$E_3 = \{[ASK],[NANNY : Joe],[FRIEND : Harry],[NULL]\} -> E_4 = \{[HIRE],[NANNY : Joe],[NULL],[NULL]\}$$
$$E_5 = \{[ASK],[NANNY : Joe],[FRIEND : Mum],[NULL]\} -> E_6 = \{[HIRE],[NANNY : Joe],[NULL],[NULL]\}$$
$$E_7 = \{[ASK],[NANNY : Jenny],[FRIEND : Mary],[NULL]\} -> E_8 = \{[HIRE],[NANNY : Jenny],[NULL],[NULL]\}$$
$$E_9 = \{[ASK],[NANNY : Jenny],[FRIEND : Harry],[NULL]\} -> E_{10} = \{[HIRE],[NANNY : Jenny],[NULL],[NULL]\}$$
$$E_{11} = \{[ASK],[NANNY : Jenny],[FRIEND : Mum],[NULL]\} -> E_{12} = \{[HIRE],[NANNY : Jenny],[NULL],[NULL]\}$$

where sequential events that make up the individual Activity Digraphs are joined with the characters '->'. The 12 events, however, only consist of eight unique events as the events for *hiring joe* and *hiring jenny* are repeated three times each. As previously mentioned the dimension of explicitness of each element of the event is inconsequential. In this case, the temporal component is nondescript. Not all elementary factors need to be present in the event.

Once the event space has been generated, it is passed to the Intention Generator. In the Intention Generator, each event in the event space is assigned an attitude value calculated by summing the attitude values the agent has about each of the elements in the event. From these assessments, the agent's intentions toward each event and associated Activity Digraph are formalized and the value of the intention is used as a prioritising mechanism that determines the agent's behaviour. Once each option has been assessed, the agent's highest intended Activity Digraph becomes its behaviour and the agent begins work on that plan.

7.8.3 Results and Discussion

The partial narrative output from one simulation run with the EMAI agent can be seen in Table 7.6. The agent begins by executing the Activity Digraph for the EMPLOY NANNY goal. As the agent does not have any preconceived attitudes, the events are executed in a FIFO order. The agent executes the ASK ABOUT NANNY$_{\text{Jenny}}$[14] event. The agent receives positive feedback about Jenny. However, the agent does not follow this process with the subsequent event HIRE NANNY$_{\text{Jenny}}$, instead it executes an activity for the ASK ABOUT NANNY$_{\text{Joe}}$. The agent selected this activity next because the positive feedback from the ASK ABOUT NANNY$_{\text{Jenny}}$ event updated the agent's attitudes about the event elements and in turn affected the overall attitudes of other events that included some or all of the same elements.

```
EMAI ask about Nanny Jenny
EMAI hears GOOD things about Jenny.
EMAI ask about Nanny Joe
EMAI hears BAD things about Joe.
EMAI tries to hire Nanny Jenny
EMAI hires Nanny Jenny.
```

Table 7.6 Narrative Output from Albert Simulation

The ASK ABOUT NANNY$_{\text{Jenny}}$ event includes the elements of the action *ASK ABOUT NANNY* and the object *Jenny*. The ASK ABOUT NANNY$_{\text{Joe}}$ event includes the same action element as the previous event. The intersection of these events can be seen in Figure 7.13.

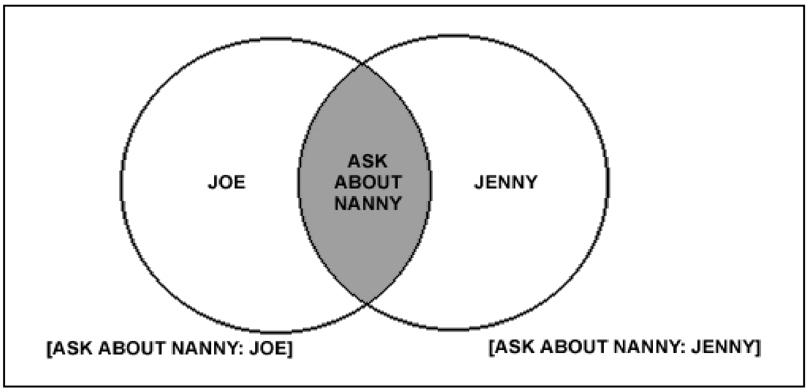


Figure 7.13 Intersection of ASK ABOUT NANNY Events

[14] As there are similar events that involve different referents, for the rest of this discussion an event that involves a particular referent will be written as EVENT NAME$_{\text{Referent}}$. For example, the ASK ABOUT NANNY event involving the referent Jenny will be written ASK ABOUT NANNY$_{\text{Jenny}}$.

Therefore, when the event ASK ABOUT $\text{NANNY}_{\text{Jenny}}$ is performed, the resultant attitude affects the ASK ABOUT $\text{NANNY}_{\text{Joe}}$ event. In other words, the agent is simulating Albert asking about the hiring of a nanny, in particular one called Jenny. When Albert receives good feedback about the prospect of not only Jenny but also the hiring of any nanny, Albert decides to ask about other nannies. The narrative continues to show that the agent receives negative feedback about the Joe referent and makes the decision to continue with the employment of the Jenny referent.

The results of this simulation are similar to those in Case Study 2 in this chapter. This is not a surprise as the mechanisms within the agent are identical. The only difference lies in the agent's Ontology. The same processing is occurring, however, for this simulation, rather than selecting an abstract *pack* concept the concepts relate to a specific domain (in this case day care selection). Figure 7.14 displays the agent's attitudes toward each of the events during a simulation run. In a similar method used in Case Study 2, the agent is asked to satisfy the FIND CARER FOR SON goal a number of times. Each time the agent satisfies the goal, the goal is triggered again. This allows the agent to build up the attitude values of its events with each success and failure. The agent is also able to use these attitude values to make decisions the next time it has to satisfy the goal. Notice the second process for each of the PLACE IN DAY CARE subgoals has been used because the final processes in the PLACE IN DAY CARE Activity Digraphs are not dependent on chance determining their success or failure, as is the case with each of the other processes. They were added for illustrative and narrative purposes only. Subgoals such as GIVE SON TO FRIEND and GIVE SON TO NANNY could have also been employed under the appropriate graphs but are unnecessary to illustrate Albert's decision making process. The graph in Figure 7.14 displays the agent's attitude toward each of the events with respect to time. The attitude value towards each event rises when the execution of the event is successful and falls when it fails. The attitude toward an event is also affected by other events that contain similar elements such as actions. This was previously illustrated in Figure 7.13 where the goals ASK ABOUT $\text{NANNY}_{\text{Joe}}$ and ASK ABOUT $\text{NANNY}_{\text{Jenny}}$ have the action ASK ABOUT NANNY in common.

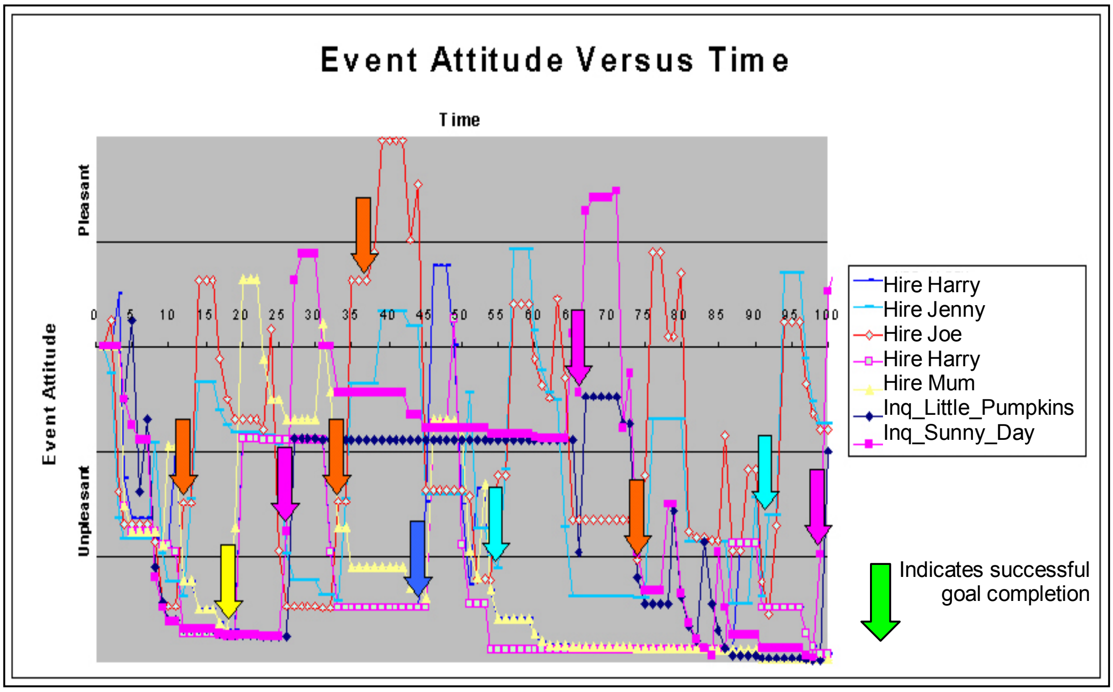


Figure 7.14 Attitude Values for Terminating Activity Subgoals

In contrast to Case Study 2 where only one subgoal needed to be completed successfully to achieve the agent's goal, this simulation employs two and three subgoals. This makes the data more complex to display in a simple graph showing choice selection order (as was the case in Case Study 2). However, similar results can be shown by displaying the attitude values of the events involved in the decision making process and the associated successful processes as shown in Figure 7.14. This graph displays how pleasant or unpleasant the agent assesses the events of each of the final processes for an Activity Digraph of the FIND CARER FOR SON goal.

As was the case in Case Study 2, the agent did not need to attempt all of its options before finding a successful solution. In the graph shown in Figure 7.14, areas can be seen where the attitude value for a particular event forms a plateau over a number of time intervals. This illustrates a time period when the agent's attitude value for this event did not change because the agent did not attempt this event or any other event that had overlapping elements. The large coloured arrows indicate when an attempted event was successful and thus satisfied the goal. Due to the event being successful, the attitude value of the associated element is increased. This can be seen by an immediate increase in the attitude value for a successful event shown on the graph. Increases in attitude value also occur when there has been a failed event execution. This also happens where elements involved in one event have their attitudes updated and thus influence the other events. The way in which the elements overlap for this simulation can be seen in Figure 7.15. For example the individual Jenny is present in both ASK ABOUT $\text{NANNY}_{\text{Jenny}}$ and HIRE $\text{NANNY}_{\text{Jenny}}$ events. If the agent's attitude toward Jenny changes as a result of executing the goal ASK ABOUT $\text{NANNY}_{\text{Jenny}}$ this in turn will influence the overall attitude for the event HIRE $\text{NANNY}_{\text{Jenny}}$ without the goal having to be executed.

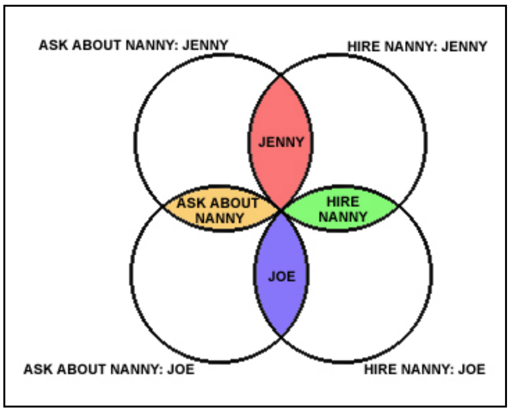


Figure 7.15 Event Intersections for the EMPLOY NANNY Goal

This interaction is difficult to see when the simulation is examined in its entirety and so the graph in Figure 7.16 has been produced to isolate and demonstrate this event interaction. The graph in Figure 7.16 takes a slice of the simulation run between time indexes 35 and 45 and shows the attitude values for the events generated from the EMPLOY NANNY goal.

Keeping in mind that the attitude value is updated after the event and that it is not shown on the graph until the following time index, the effects of the positive response the agent receives about Joe at time index 36 are shown at time index 37. The event for ASK ABOUT $\text{NANNY}_{\text{Joe}}$ affects all events containing the individual Joe and all events containing the action ASK ABOUT NANNY. Thus, the attitude values for the events ASK ABOUT $\text{NANNY}_{\text{Joe}}$, ASK ABOUT $\text{NANNY}_{\text{Jenny}}$ and HIRE $\text{NANNY}_{\text{Joe}}$ all receive an updated attitude value. This can be seen in Figure 7.16 at time index 37. The event for HIRE $\text{NANNY}_{\text{Jenny}}$ is not affected as it does not intersect with the ASK ABOUT $\text{NANNY}_{\text{Joe}}$ event.

At time index 37 the agent receives negative feedback about the event ASK ABOUT $\text{NANNY}_{\text{Jenny}}$ and positive feedback about the event HIRE $\text{NANNY}_{\text{Joe}}$. This results in the attitude value changes displayed at time index 38. All events including ASK ABOUT NANNY are decreased and the events containing the action HIRE NANNY are increased. However, as positive feedback was received regarding an event containing the referent Joe, the attitude toward the event ASK ABOUT $\text{NANNY}_{\text{Joe}}$ decreases less than the attitude for the event ASK ABOUT NANNY$_{Jenny}$. The reverse is true for the HIRE NANNY$_{Joe}$ event whose attitude value increases by a greater amount than the event ASK ABOUT NANNY$_{Jenny}$ as Joe had a positive influence on the event and Jenny had a negative one. Further interactions of this type can be seen after time index 41.

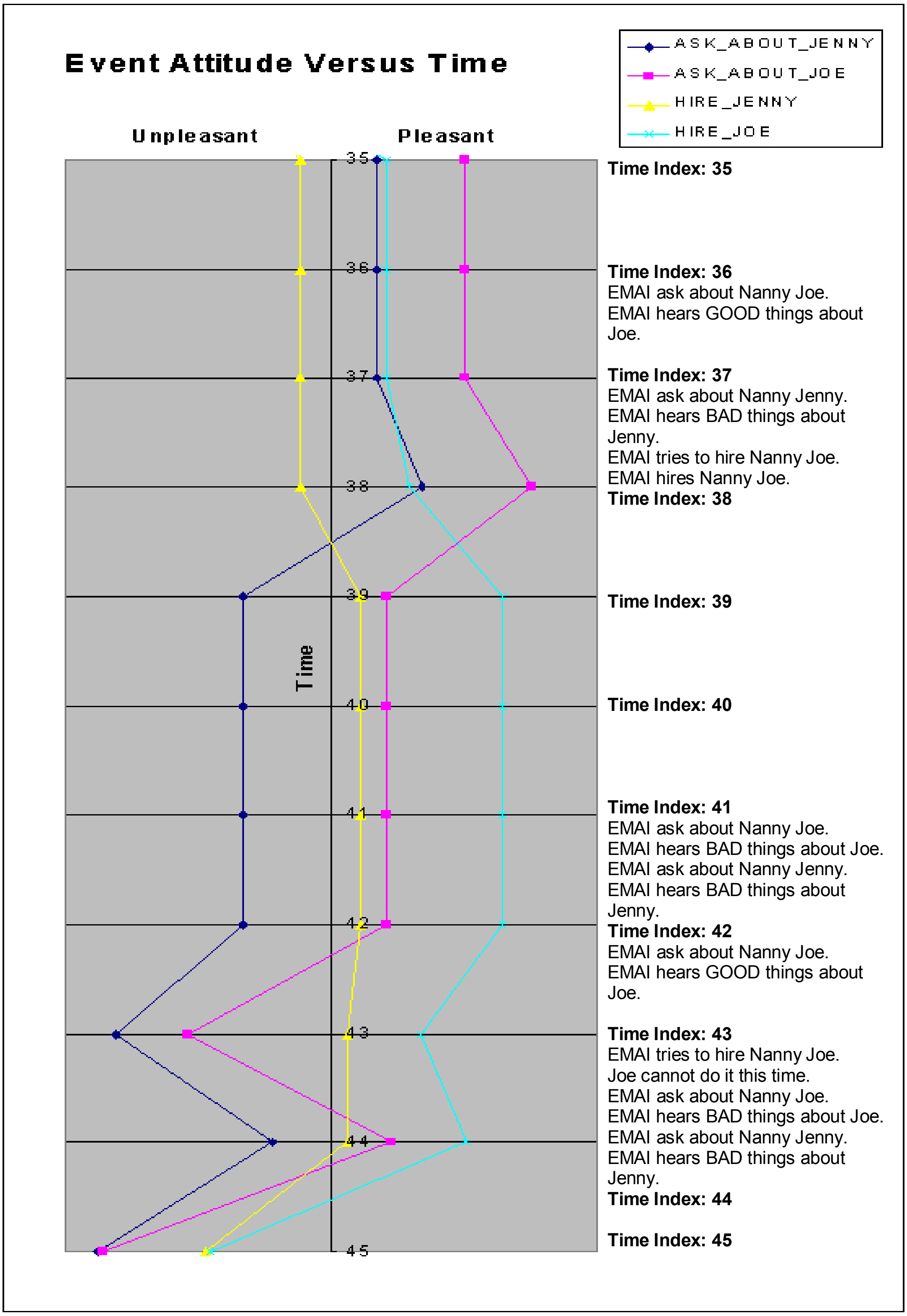


Figure 7.16 Event Attitude versus Time Showing Interrelated Event Effects

In the time between time index 39 and 42, the attitude values on all events remain constant. This result is caused when the agent does not interact with these elements and therefore, there are no new negative or positive influences over the elements for these events.

7.8.4 Conclusions

In Picard's suggested solution to Albert's problem, she explains that in making his decision, Albert does not evaluate all possible solutions but works along a path of *feel good* options until a negative feeling arises at which point he changes tack and begins evaluating other options. In the simulation presented in this case study the agent performs the same kind of process: it selects a process from one Activity Digraph and proceeds to work on that as a plan to achieving its goal. However, at any time, the agent can decide it does not like the current plan it is working and can abandon it for another. The agent can, of course, return to the abandoned plan at any time if it evaluates it as being more *liked* or *better* than the other plans.

The attitude values stored by the agent are based on the agent's experiences within its environment. In the case of the child-care problem the solution is neither right nor wrong. Any of the agent's subgoals and associated digraphs, when completed successfully, would satisfy the agent's initial goal. By setting attitude values the agent can form opinions about which process it *likes* the best and can use these values to make a final decision about its choices without exploring all the processes that it has available to it through its hierarchy.

7.9 Summary

Attitude theory provides a firm foundation for integrating single-dimensional emotion-inducing mechanisms into an agent architecture. In this chapter, the foundations for development of an affective reasoning agent have been examined. The concept of attitude has been used to calculate values for an affective reasoning system. By using attitude as a measurement, the EMAI agent is programmed with the means of appraising an event as either *pleasurable* or *displeasurable*. This pleasure rating is calculated for the events of processes that the agent is to perform and acts as a device for the prioritising of an agent's intentions with respect to the agent's current set of goals.

Fishbein and Ajzen's Theory of Reasoned Action fits well with goal-orientated agent architectures such as the EMAI architecture and the discrete nature of computational AI systems. The Theory of Reasoned Action can not only be used to determine an attitude about an activity from a holistic viewpoint, but it can also be used to assess atomic elements of an event, namely the action being performed, the target object, the context and the time of the action. This makes it ideal for integration into an agent architecture to simulate affective decision making across a series of processes and the associated events where event elements may overlap. This gives the agent the ability to assess its attitude towards a process it has not performed before where it has had contact with several of the elements associated with an event.

Attitude theory provides the agent with a one-dimensional emotion system. This may be adequate for simulating simple decision making scenarios where the complex effects of such decisions on the agent's state are not necessary. However, when an agent is required to exhibit emotional human-like behaviour, it is essential to allow activities and the environment to have a complex influence on the agent's internal state, which will in turn influence the agent's outward behaviour. Attitude ratings of liked and disliked may allow a shallow implementation of emotions that are highly correlated to this dimension, such as *happiness* equating with liked and *sadness* equating with disliked. However, complex emotional states are beyond the definitions of only liked and disliked. For example the question may be asked, *"Is pride more or less liked than happiness?"* This illustrates that a wide range of emotional states cannot adequately be represented on a one-dimensional scale.

The next chapter builds on the one-dimensional aspect of attitude theory presented in this chapter in order to create a mechanism within the agent architecture that is capable of generating complex emotional states. This is achieved by the inclusion of a component in the EMAI architecture called the *Affective Space*.

8. Affective Space

The foot less prompt to meet the morning dew, the heart less bounding at emotion new, and hope, once crush'd, less quick to spring again.
- *Matthew Arnold, 1888.*

8.1 Introduction

Due to their discrete and categorising structure, appraisal theories have been the popular foundation and theoretical concept on which emotional agent architectures have been built (Picard 1998). The most popular to date has been the OCC model discussed in Section 2.1.3. Appraisal theories make assessments on particular aspects of an entity, be it an event, object, a piece of music, another person and so on, in order to group it into an emotional category. Using just one appraisal dimension to determine complex emotional states, such as pleasantness, is a difficult, if not an impossible task. For example, *pride* is a pleasant emotion, but is it more or less pleasant than *surprise*? The Theory of Reasoned Action and attitude theory in general collectively refers to a number of appraisal dimensions. Petty and Cacioppo (1996) tabled a survey conducted by the White House Chief of Staff, where the White House staff were asked questions that relate to a number of different appraisal dimensions such as effort, quality, confidence and frequency, to gather a clearer picture of the subject's attitudes regarding other staff members. A simple, "*I like this person"*, or "*I do not like this person"*, would not be adequate for evaluating the staff member. Therefore, more specific appraisals need to be gathered. This is also the case with determining an emotional state, for example, distinguishing *pride* from *surprise*.

The number of appraisal dimensions that could be used to categorise emotional states is almost limitless and varies depending on the person, object or event being assessed. An emotional state could be evoked from the assessment of smell, sound, sight or on more calculated assessments of predictability, effort and legitimacy. Healey et al. (1998) described a wearable computer that selects music to play, based on the assessment of the user's emotional state calculated from the appraisal of physiological variables that indicate the user's present mood. Appraisals will differ in different context. The objective of this chapter is to examine event appraisal. An emotional reaction to a piece of music may be assessed with different criteria. An emotional reaction toward a person based on how they look may also be assessed with different criteria.

Developers of the agents that have used the OCC model as a basis for the generation of emotions such as El-Nasr's PETEII (El-Nasr 1998) and Reilly's emotional agents (Reilly

1996) have developed complex environments where the agents not only make their own decisions but are also affected by events occurring within their environment that are beyond their control. Among others, these agents use appraisal dimensions that make abstract appraisals such as the liking, disliking, attractive or unattractive. These appraisals are not only conjectural but also require emotional judgments themselves. Having calculated an emotional state, these agents then use a set of production rules to determine their behaviour. In the case of PETEEI, which simulates a pet dog, there exists the rule *if angry and bowl taken away then bark*. These types of static production rules do not allow for dynamic performances from the agents. Although adaptive behaviour may not have been the basis for these types of agents, it makes the transition of these architectures into other domains difficult. For example, given the aforementioned production rule involving the *angry* dog and the bowl, this implies that the *bowl taken away* is a bad action and if the dog is already *angry* then its behaviour should be to bark angrily thus signalling to the user the dog is displeased. However, it could be possible through the experiences of such an agent that the *bowl taken away* action becomes a good action and therefore, the production rule would no longer apply.

To overcome this type of problem, the EMAI architecture does not include a set of production rules but instead implements an *Affective Space* that associates emotional assessments to all concepts known to the agent. The agent's emotional assessment of a concept is dynamic and changes as the agent gains experience through interaction with the concept. This emotional assessment is then coupled with the agent's mood in order to determine an appropriate behaviour. To remove the difficulty of using judgments based on emotions as appraisal dimensions another psychological theory was sought that eliminated such decisions. The six-dimensional appraisal model proposed by Smith and Ellsworth in (Smith and Ellsworth 1985) was chosen for this task.[15] This model proposes six cognitive appraisal dimensions by which to differentiate emotional experiences and will be introduced in the next section.

Another problem with the previous emotional agent models is the use of gauge like mechanisms for modelling individual emotions (Reilly 1996, El-Nasr 1998, Padgham and Taylor 1997, Velasquez 1998). These gauges must be individually added to the agent, one

[15] Other appraisal theories exist that present the analysis of emotion using a multidimensional (Scherer 1982, Roseman 1990). While the Affective Space is designed such that other appraisal dimensions can be used, empirical data for Smith and Ellsworth's model was readily available.

for each emotion, and their own set of production rules and integrative programming incorporated into the existing agent. The *Affective Space* created from the six-dimensional appraisal model overcomes this problem as all emotions are integrated into the Affective Space and do not require individual sets of production rules to determine behaviour. They are not represented as gauges but rather as single locations or points within the space.

The appraisal theory chosen needs to be complex enough to include emotional blending and pliable enough to allow the creation of different personalities. The six-dimensional appraisal model utilises the assessment of physical characteristics rather than emotional or nonspecific assessment criteria. This model is not concerned with emotions of others but rather an assessment of an event. The model implemented in the EMAI architecture allows an EMAI agent to make an emotional decision about the performance of future behaviours. Using this emotion-based assessment the agent can change its emotional state on the execution of a behaviour.

8.2 Six Dimensions of Appraisal

Smith and Ellsworth's experimentation and analysis outlined in (Smith and Ellsworth 1985), identifies six orthogonal dimensions: *pleasantness, anticipated effort, certainty, attentional activity, responsibility* and *control*, across 15 emotions: *happiness, sadness, anger, boredom, challenge, hope, fear, interest, contempt, disgust, frustration, surprise, pride, shame* and *guilt.*

In the previous chapter, the value of intention was used to make a one-dimensional assessment of a situation. This appraisal dimension merged a rating on a scale from good to bad with a rating on a scale from likely to unlikely. Although, the dimension of intention does not appear in the six dimensions mentioned above, the ratings are still used and therefore, the premise of intention still exists in the multi-dimensional model. The dimension of pleasantness is synonymous with the good to bad scale and the dimension of certainty is synonymous with the likely to unlikely scale. Figure 8.1 displays two of the orthogonal dimensions (*pleasantness* and *control*) and where the 15 emotional states positioned with respect to them[16]. There is no particular reason why these dimensions were chosen for display except that the pure emotion points are spread most across the

[16] The shape and colour of the emotional point markers are immaterial in this plot.

pleasantness axis. This illustration will be used in Chapter 9 to demonstrate the EMAI architecture's computational operations on emotional states. The empirical values for the positions of the pure emotions comes from Smith and Ellsworth's original study (Smith and Ellsworth 1985)[17].

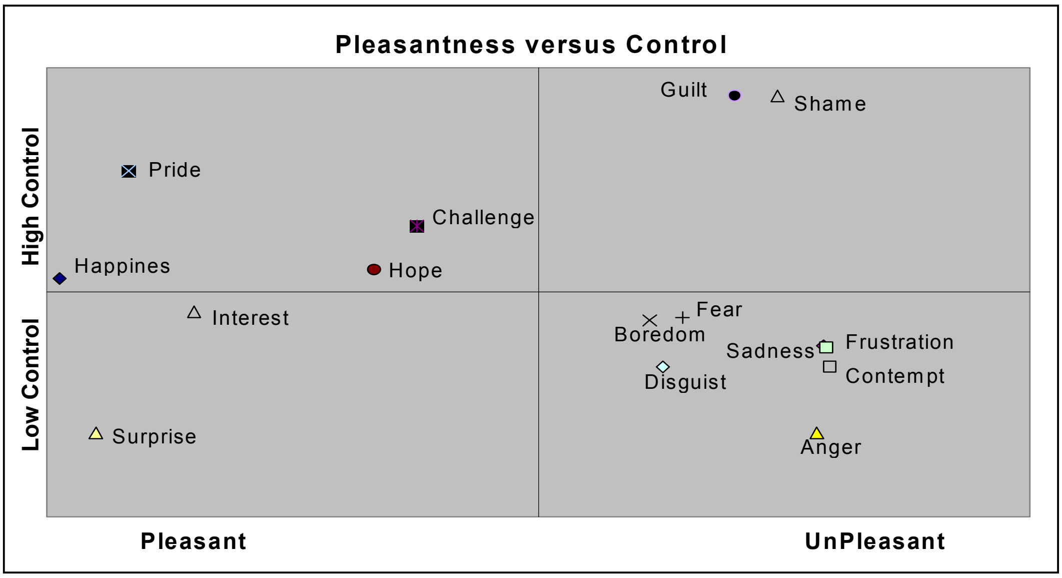


Figure 8.1 Empirical Location of Emotional States with Respect to the Pleasantness and Control Dimensions

The values of the pure emotion points (as determined by Smith and Ellsworth (1985)) for each of the 15 emotional states in the model are shown in Table 8.1.

EP	P	R	C	A	E	O
Happiness	-1.46	0.09	-0.46	0.15	-0.33	-0.21
Sadness	0.87	-0.36	0	-0.21	-0.14	1.51
Anger	0.85	-0.94	-0.29	0.12	0.53	-0.96
Boredom	0.34	-0.19	-0.35	-1.27	-1.19	0.12
Challenge	-0.37	0.44	-0.01	0.52	1.19	-0.2
Hope	-0.5	0.15	0.46	0.31	-0.18	0.35
Fear	0.44	-0.17	0.73	0.03	0.63	0.59
Interest	-1.05	-0.13	-0.07	0.7	-0.07	0.41
Contempt	0.89	-0.5	-0.12	0.08	-0.07	-0.63
Disgust	0.38	-0.5	-0.39	-0.96	0.06	-0.19
Frustration	0.88	-0.37	-0.08	0.6	0.48	0.22
Surprise	-1.35	-0.97	0.73	0.4	-0.66	0.15
Pride	-1.25	0.81	-0.32	0.02	-0.31	-0.46
Shame	0.73	1.13	0.21	-0.11	0.07	-0.07
Guilt	0.6	1.13	-0.15	-0.36	0	-0.29

Table 8.1 Mean Locations of Emotional Points as Compiled in Smith and Ellsworth's Study

[17] Note: Though the empirical location of the emotional states shown in Figure 8.1 are the result of Smith and Ellsworth's original study, it is possible that other studies could locate these states in different quadrants. For example, it is possible in other contexts to locate surprise in the unpleasant region.

The points are the mean position of the emotional state within the Affective Space as collected from 16 participants (Smith and Ellsworth 1985). As each human's individual perspective on an emotion and a situation are different, it would be incorrect to conclude that emotional state points be equally positioned for all people. From this it is assumed that an Affective Space can be calibrated for an individual, thus producing a personality perspective on how each of the six dimensions affects that person and triggers different emotional states.

The term *Affective Space* has been coined in this dissertation to describe this space. Due to the continuous nature of the space with further research it would be possible to pinpoint the location of an infinite number of emotions. It could also be done without the need to add further variables, gauges or dimensions to the model (as would be the case with previously mentioned models).

Smith and Ellsworth (1985) selected the appraisal dimension from a larger group of appraisals that has been previously identified in popular appraisal theories. The appraisal dimensions chosen were those, which presented themselves during investigation, as the most highly correlated with emotional states within the set of human subjects examined. These dimensions will now be defined with respect to their use in the EMAI architecture.

8.3 Analysis of Appraisal Dimensions

The six appraisal dimensions are orthogonal, and no single emotion can be identified without taking into account each of these six appraisal dimensions. Each of these dimensions will now be reviewed.

Pleasantness. This dimension relates to an individual's expression of liking or disliking towards an event, object or another agent. This is the scale that is referred to in the Theory of Reasoned Action as *attitude*. For the purposes of the EMAI architecture, this traditional definition of pleasantness is quite vague and meaningless in the assessment of an element by an artificial being. Whether or not the agent finds something pleasing or displeasing could be programmed directly into the agent. However, this would defeat the purpose of building an agent capable of adapting its assessment of this dimension. For example, the agent should be able to find an element pleasing at one moment and displeasing at another as a result of a negative experience.

An EMAI agent assesses and updates the pleasantness dimension as an assessment of a task event and its outcome. If the agent fails to complete a task the pleasantness dimension is decreased in value. If the agent experiences success in relation to a task, the pleasantness dimension is increased in value.

Responsibility. This dimension correlates with an individual's sense of personal involvement and amount of blame or credit attributed towards the self when interacting with an event, object or another agent. The two extremes of measure are self-responsibility and others-responsibility. In the EMAI architecture, responsibility is a measure of the agent's relationship and attachment toward an event element. Note the agent's attachment towards an event element should not be measured in terms of emotional attachment. Rather it should be measured in terms of other forms of attachment like how long the agent has had the item, how much it cost, what is the agent's relationship within the event elements, and so on. For example, the agent will calculate a high responsibility for an object where the agent considers it has ownership over it. More precisely, if the agent were in a team situation and within the team there were different ranks (for example, team leader, second in charge and third in charge), and the team was performing some task and the task failed then the team leader would feel most responsible, the second in charge less responsible and so on through the ranks.

Effort. The values along this dimension are gauged from an agent's exertion with respect to an event, object or another agent, either mental or physical. In the EMAI architecture, effort is a function of the depletion of resources used when interacting with an object or performing an event divided by the length of time spent interacting. For an EMAI agent, at the beginning of performing an event the agent's physical state is recorded. During the event, the agent's physical state may be affected by the elements in the event space. At the completion of an event, the change in the agent's physical state is used to calculate effort. For example, if the agent were to perform the task of digging a hole, during the execution of the event the agent's physical state may deteriorate because the agent may get *tired* and *hungry*. When the agent finished digging the hole, the change in *tiredness* and *hunger* during the task performance will be directly related to the effort involved in the event.

Attention. This dimension is the rating of an individual's regard for an event element (that is an event, an object or other agent) with respect to the level of concentration exerted towards it during interaction. Each element has an associated value indicating the level of attention required to successfully accomplish the task. All EMAI agents are programmed with a maximum attention capacity, and each of these agents can perform one or more tasks, that utilise its maximum capacity. In EMAI, attention is measured as the total amount of an agent's attention that is utilised in performing one or more events concurrently. For example, consider an EMAI agent that is programmed with maximum capacity value of 100 units, and the following three events that may be performed concurrently by an agent: *drive car*, *eat ice cream* and *drink coffee*. The *drive car* event may have associated with it an attention value of 70 units, the *eat ice cream* event may utilise 40 units and the *drink coffee* event may have attached to it a value of 20 units. Here, if this agent has decided to *drive car*, then it can only perform either *eat ice cream* or drink *coffee*, but not these two at the same time.

Control. This dimension refers to an agent's authority to manipulate and direct an event, object or another agent. Examining the relationship of the agent toward the entities involved in an event, including the assessment of past episodes of these events and their outcomes, assesses control in the EMAI architecture. Each EMAI agent is initially programmed with its control value over other elements as 0. Over time, as an EMAI agent evolves, its control values toward another element (such as an event, an object or another agent) changes (either increases, or decreases) according to the outcome of event processing. Almost always there will be one or more elements involved in an event. So initially, the EMAI agent's control over this event is 0, and the control over each of the individual elements involved in this event is also 0. Based on the outcome of this event (that is, success or failure), the control of the agent towards each of the elements involved in the event will be either increased by 1 if the event is successful or reduced by 1 otherwise. The overall control over the event is then the average of the control over each individual element in the event. For example, assume that an agent is driving a car on the Princes Highway from Sydney to Melbourne. Here, there are four elements involved in this event. They are: the car, the Princes Highway, the source city Sydney, and the destination city Melbourne. Initially, the control the agent has over these four elements is 0. Now, assuming that this event was successful, the agent will increment the control value of each of these elements by 1, and calculate the overall control value towards this event to be the average control over all the elements in the event. In this case, it will be 1 ((1+1+1+1)/4). Further, assume the agent now performs the same event successfully at another time. Now, its control over each of the elements will be increased by 1, and the overall control towards the event will be 2. But, if on the third trip, the agent drives the car along Pacific Highway from Sydney to Brisbane, the initial control over Sydney and car is 2, respectively, but the control over Pacific Highway and Brisbane is 0, respectively. So, the overall control over this event will be 1 ((2+2+0+0)/4). If this event was successful, the overall control will be 2 ((3+3+1+1)/4).

Certainty. This dimension refers to an individual's assessment of an event, object or person as to the reliability that its outcomes or behaviours can be predicted. This is calculated in an EMAI agent by considering the degree of success or failure the agent has had in executing past episodes of similar events. For example, if the agent has carried out 50 attempts at driving from Sydney to Melbourne on the Princes Highway, and it succeeded in 30 of its attempts, and failed the other 20, then the certainty of success on the 51st attempt will be 0.6 and the certainty of failure will be 0.4.

To summarise, of all the dimensions the dimension of pleasantness is by far the most dominant. As can be seen in Figure 8.1, emotions lie near both extremes of the pleasantness axis. However, as many emotional states such as *frustration* and *contempt* have similar assessments in the dimension, as was mentioned previously, pleasantness alone cannot be used to distinguish between the emotions.

Measuring the spread of values of the emotional points (shown in Table 8.1) can identify the weight of effect that a dimension has on an emotion. Based on the values collected by Smith and Ellsworth (1985), the influence of each dimension listed from most to least dominant would be: pleasantness, responsibility, effort, attention, control and certainty.

The Affective Space, Φ, is defined as a six-dimensional space occupied by points that ascertain the current emotional state of the agent. The emotion that is closest to this point is the most influential emotion in the agent. The following formalisation of the Affective Space appears in (Baillie and Lukose 2001b). Each emotional state is a point, Ω, in that space defined by a six point coordinate based on the attitudinal value of the six cognitive appraisals toward an event, object or another agent. Ω can be defined as:

$$\Omega = \{P, E, C, A, R, O\} \qquad 8.1$$

where *P, E, C, A, R* and *O* are the six appraisal dimensions of *pleasantness*, *effort*, *certainty*, *attention*, *responsibility* and *control*, respectively. Each of these values is determined by the analyses given earlier in this section. From these descriptions the following equations have been formulated.

Pleasantness, *P*, is the average summation of the rating of pleasantness the agent has given to an element or event each time the agent has come into contact with it:

$$P = \left(\sum_{i=1}^{m} p_{e_i} \right) / m \qquad 8.2$$

where *m* is the number of times the agent has come in contact with the element *e* and p_e is the pleasantness rating of *e*. For example, assume the agent has driven the same car five times. The first time the car performs as expected and the agent is pleased with the car. In this first instance the agent may set the pleasantness rating to 8 on a scale from 1 to 10 where 1 is unpleasant and 10 is very pleasant. The second time the agent drives the car it breaks down. For this instance the agent rates the pleasantness of the car as 2. For the next three contacts with the car the agent rates the pleasantness as 2, 7 and 9. After these five contacts with the car, the agent, using Expression 8.2 assesses the overall pleasantness rating of the car to be (8+2+2+7+9)/5 which equates to 5.6.

Effort, *E*, is the averaged summation of the effort associated with an element or event each time the agent has been in contact with it:

$$E = \left(\sum_{i=1}^{m} f_{e_i} \right) / m \qquad 8.3$$

where *m* is the number of times the agent has come in contact with the element *e*, and *f* is the amount of effort involved with *e*.

Certainty, *C* is calculated by determining the probability of success for the agent's involvement with an element or event. Given that *S* is a function that returns the number of times that *e* has been used by or involved with the agent for a successful event, *C* can be calculated as:

$$C = S(e) / m \qquad 8.4$$

where *m* is the number of times the agent has come in contact with the element *e*.

Attention, *A,* is determined by averaging the summation of all the attention ratings the agent has assigned to an element or event. Each time the agent has an involvement with the element or event, the agent records how much concentration was exerted during the encounter and uses these values to calculate *A*:

$$A = \left(\sum_{i=1}^{m} a_{e_i} \right) / m \qquad 8.5$$

where *m* is the number of times the agent has come in contact with the element *e*, and *a* is the amount of attention required by the agent when involved with *e*.

Responsibility, *R,* is calculated using the function *r* that returns the level of responsibility related to an element or event. The function works by determining the nature of the relationship between the agent and *e.* For example, it the agent were the owner of or in charge of *e*, *r(e)* would return a high value. Responsibility is calculated as follows:

$$R = r(e) \qquad 8.6$$

In Expression 8.6, the average of the summation of *r(e)* is not considered because the assumption that is adopted in this dissertation is that the responsibility an agent has towards an element or event does not change over time. Although it is not believed this is the case in the real world, the issue of responsibility is complex and has many influencing factors. As this model is already complex enough, a dynamic responsibility element will not be explored.

The final dimension, Control, *O,* is calculated by averaging the summation of the amount of control the agent has had over an element or event during every encounter the agent has had with the element or event:

$$O = \left(\sum_{i=1}^{m} o_{e_i} \right) / m \qquad 8.7$$

where *m* is the number of times the agent has come in contact with the element *e*, and *o* is the amount of control the agent had over *e*.

The values of these six dimensions form the foundation for describing the emotional aspects of an event. Combined, they form an Affective Space.

8.4 Assigning an Emotional State

In the EMAI architecture an emotional state can be assigned to any item be it an event element, an event or a plan. By assessing an item using the six appraisal dimensions a six coordinate point can be determined. This point, when plotted in the Affective Space can be compared to the 15 pure emotion points that exist within.

For any item, *i*, an $\mathbf{\Omega_i}$ value can be determined that represents the emotional state evoked by the item. The P_i, E_i, C_i, A_i, R_i and O_i values are calculated for an item using the appraisal methods explained in Section 8.3. The representation for $\mathbf{\Omega_i}$ is shown in Expression 8.8 and will be further discussed in Chapter 9.

$$\Omega_i = \{P_i, E_i, C_i, A_i, R_i, O_i\} \tag{8.8}$$

Given $\mathbf{\Omega_i}$ for an item, the emotional state can be deduced, for the purpose of expression in natural language, by determining the distance that $\mathbf{\Omega_i}$ is from each of the 15 pure emotion points. To do this a simple linear distance function is applied. While more complex distance functions could be implemented and examined, for simplicity this will not be examined further in this investigation of the EMAI architecture.

To determine a word that best describes the emotion point of $\mathbf{\Omega_i}$ the distance between the item's emotional state point and each of the pure emotions ($\mathbf{\Omega_1 \ldots \Omega_{15}}$) is calculated using Expression 8.9.

$$\Delta_{\Omega_i} = \sqrt{(P_i - P_j)^2 + (E_i - E_j)^2 + (C_i - C_j)^2 + (A_i - A_j)^2 + (R_i - R_j)^2 + (O_i - O_j)^2} \tag{8.9}$$

where 15 values will be calculated for $j = 1..15$. The pure emotion for the item expressed as a written word, Em_i, closest to $\mathbf{\Omega_i}$ is then determined by using Expression 8.10.

$$Em_i = emotion\_name(\min(\bigcup_{j=1}^{15} \Delta_{\Omega_i})) \tag{8.10}$$

where the function *min* returns the pure emotion point in closest proximity to $\mathbf{\Omega_i}$ and the function *emotion_name* converts the pure emotion point into a string.

For example, assume an item with an emotional state point of $\mathbf{\Omega_i}$ =[15, 87, 35, -30,10,-50]. To find the name of the emotion that best describes the item's emotional state, the first step is to find the distance between this point and the 15 pure emotion points in the Affective Space using Expression 8.9. The results are shown in Table 8.2.

Emotion	Δ_Ω
happiness	2.08
sadness	2.22
anger	2.18
boredom	2.15
challenge	1.59
hope	1.46
fear	1.70
interest	2.11
contempt	1.68
disgust	1.74
frustration	1.93
surprise	2.68
pride	1.64
shame	0.80
guilt	0.84

Table 8.2 Distance Between Item's Emotional State and Pure Emotions in Affective Space

It can be seen from Table 8.2 that the pure emotion closest to the item's emotional state is *shame*. Therefore, the item's emotional state is referred to as *shame*.

8.5 Emotion Blending

Emotion blending is a concept devised by psychologists such as Plutchik (see Section 2.2.3) where it is stated that there exist several basic emotions and that more complex emotions can be created by blending different combinations of the basic emotions. The problem with any emotion synthesis in affective computing will be in determining which emotions when blended together produce other emotions. In affective agent architectures such as Em (Reilly 1996) and PETEEI (El-Nasr 1998) emotions are represented by individual variables or gauges within the architecture. The problem occurs when the gauges need to be coordinated. This requires complex programming within the agent to describe the relationships that exist between the individual variables. For example, if two variables existed where one represented a level of *happiness* and the other *sadness*, the agent would have to be programmed in a way that made certain the agent was not experiencing pure *happiness* and pure *sadness* at the same time. The more emotions represented within the architecture, the more individual variables that would exist and the more complex the relationships between them.

The EMAI architecture eliminates the need for defining the complex relationships and programming coordination between emotion state variables because of the nature of the Affective Space. The relationships that exist between the pure emotions of the Affective Space are simply defined by the distance between them along each of the six appraisal dimensions. Rather than having individual variables that represent the emotions integrated into the architecture and using these same variables to record the strength of the emotion of the agent, the EMAI architecture separates the two. The 15 pure emotions are fixed points within the Affective Space. They represent the relationship between the pure emotion and the appraisal dimensions but do not record the agent's current emotional state (to be examined in Chapter 9) or the emotional state of any assessed item (as determined in the previous section). The emotional state of an item or the agent is separate from each pure emotion. Because of this representation, theoretically, an infinite number of pure emotions could be added to the architecture with very little effort. Furthermore, as the emotional state point for an item (or the agent) is independent of the pure emotion points, the value for the emotional state point can freely change and not affect the pure emotion points.

Although discrete points within the Affective Space represent the pure emotions, they exist within an affective continuum. Unlike the other affective agent architectures that model pure emotions as discrete values with complex rules for blending these values, the Affective Space defines blended emotions as the space that exists between the pure emotion points. While the emotional state for an item may fall close to the proximity of a pure emotion within the Affective Space, most often it will not fall exactly at the same spot as a pure emotion.

The location of the emotional state point for the item represents a blended emotional state. For example, an emotional state point that has a distance of 2.15 from the point of pure *happiness* and 2.15 from the point of pure *surprise* will represent a blended emotion with equal parts of *happiness* and *surprise*. Such a state would be described as *surprised happiness*. Table 8.3 displays the results of blending equal parts of two of the 15 pure emotion points with each other. Where the blended value falls close to another pure emotion point the name of this emotion is used to describe the blended state.

The way in which emotional states are blended together is defined and examined in detail in Chapter 9.

	Sadness	Anger	Boredom	Challenge	Hope	Fear	Interest	Contempt	Disgust	Frustration	Surprise	Pride	Shame	Guilt
Happiness	hope	contempt	bored-happiness	interest	interest	hope	interested-happiness	hope	disgusted-happiness	hope	surprised-happiness	happy-pride	hope	pride
Sadness		frustration	bored-sadness	fear	hopeful-sadness	fearful-sadness	hope	frustration	disgusted-sadness	frustrated-sadness	hope	hope	shameful-sadness	shame
Anger			disgust	challenging-anger	contempt	frustration	contempt	*angry*-contempt	contempt	*angry*-frustration	contempt	contempt	contempt	contempt
Boredom				disgust	hopeful-boredom	disgust	hope	disgust	bored-disgust	disgust	hope	bored-pride	guilt	guilty-boredom
Challenge					challenging-hope	fearful-challenge	challenging-interest	frustration	challenging-disgust	challenging-frustration	interest	challenging-pride	challenging-shame	shame
Hope						hopeful-fear	hopeful-interest	hopeful-contempt	hopeful-disgust	hopeful-frustration	hopeful-surprise	hopeful-pride	shame	shame
Fear							hope	frustration	fearful-disgust	fearful-frustration	hope	hope	fearful-shame	shame
Interest								hope	hope	hope	interesting-surprise	happiness	hope	hope
Contempt									disgusting-contempt	frustrating-contempt	hope	hope	shameful-contempt	shame
Disgust										contempt	hope	disgusting-pride	shameful-disgust	disgusting-guilt
Frustration											hope	hope	frustrating-shame	shame
Surprise												happiness	hope	hope
Pride													guilt	guilty-pride
Shame														guilty-shame

Blended Emotions

Table 8.3 Resultant Emotions from Blending Pure Emotions

8.6 Emotional State Decay

The same characteristics of the Affective Space that allow for a natural emotion blending process also address the issue of emotional state decay. Emotional decay is the mechanism by which an agent that is experiencing an emotional state will, overtime, begin to have the strength of that emotional state reduced. This ensures the agent does not stay in the same emotional mood constantly. Contemporary affective agent models (PETEEI (El-Nasr 1998), Em (Reilly 1996), Silas (Blumberg 1996), Yuppy (Velasquez 1999)) address the issue of emotional state decay using decay rate functions. Because these models represent emotions and their associated strengths as individual variables, emotional decay must be dealt with manually. This means coordinating the rate of decay of an emotion variable with other emotion variables. For example, an agent that is experiencing *happiness*, will need to have the value of the *happiness* variable decreased as the value of the variable representing *sadness* is increased, otherwise the agent will be experiencing two opposing emotions at the same time.

As the Affective Space is a continuous multi-dimensional space and the pure emotions are segregated from the affects of an emotional state point mapped into the space, the decaying of emotional states is not necessary. A single point in the Affective Space represents an emotional state. If the point is in proximity to a pure emotion, the emotional state point is assigned that emotion name. The closer the emotional state point is to a pure emotion, the stronger the emotion being experienced. The closest that an emotional state point can get to the pure emotion is to be positioned at the exact same location in the Affective Space as the pure emotion. When the emotional state of the item being assessed changes, the position of the emotional state point in the Affective Space also changes. Therefore, the emotional state point for the assessed item cannot be experiencing two opposing emotional states at the same time. For example, an item cannot be assessed as being in an emotional state of pure *happiness* if the emotional state point is close to pure *sadness*.

The Affective Space, by its very nature, generates naturally decaying emotions without any extra programming or coordination effort. Figure 8.2 illustrates the strength of emotion associated with the emotional state point of an item as the emotional state point moves along the appraisal dimension axis of pleasantness from the values of pleasant to unpleasant. As can be seen, when the emotional state point for the item is at its most pleasant, the strength of the emotion *happiness* is the greatest. This is because the point for the pure emotion of *happiness* is closest to the pleasantness axis at the pleasant end. As the emotional state point moves from the left to the right in the graph in Figure 8.2 the strength of the *happiness* emotion with respect to the item's emotional state point reduces.

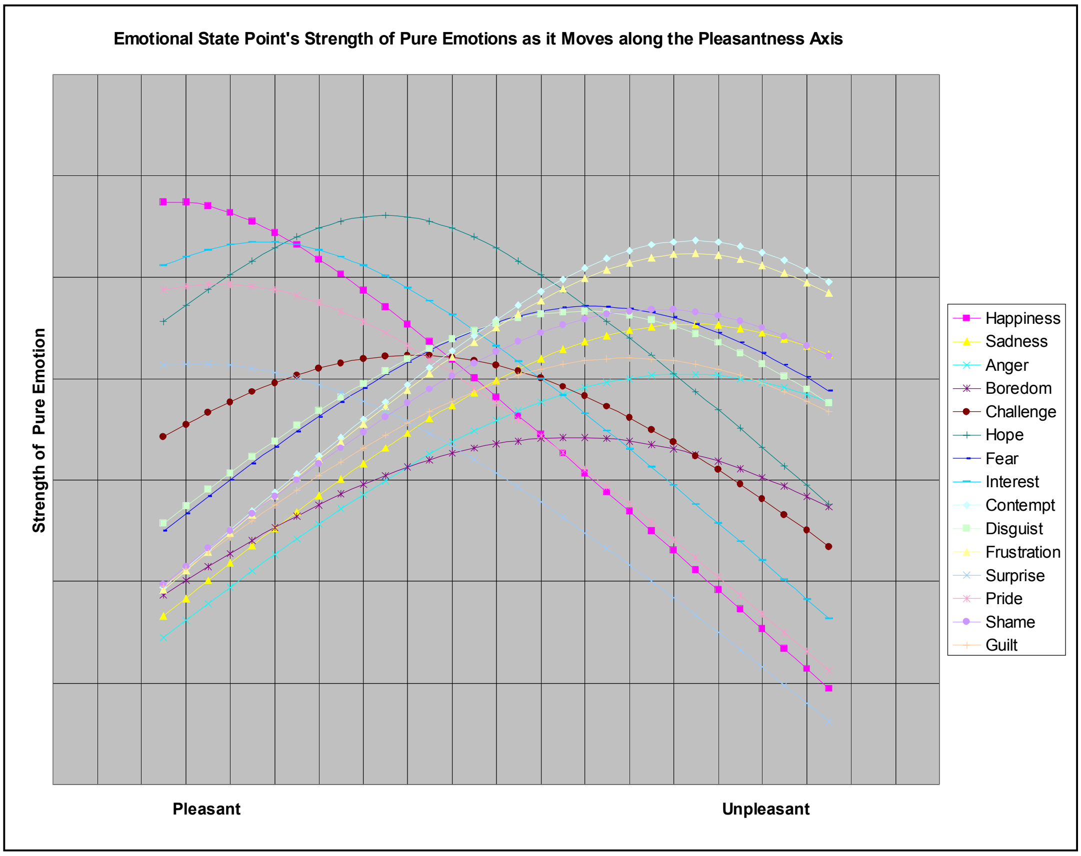


Figure 8.2 The Strength of Each Pure Emotion with Respect to an Emotional State Point Moving Along the Pleasantness Axis through the Affective Space.

It is important to note as Koestler did (see Section 2.2) that a pure emotion is never experienced. Although the EMAI architecture uses the name of the pure emotion point closest to an item's emotional state point to describe the emotion associated with the item, it will still contain amounts of the other 15 emotions. Wherever an emotional state point is placed in the Affective Space, it can never be infinitely far away from any of the pure emotions and therefore, they all have an influence on the emotional state point.

8.7 Summary

The Affective Space, introduced in this chapter, is a new and unique concept in the domain of affective computing. Due to its multi-dimensional and continuous structure it eliminates the need for several time consuming and complex programming tasks that are necessary in other affective agent models (PETEII (El-Nasr 1998), Em (Reilly 1996), Silas

(Blumberg 1996) and Yuppy (Velasquez 1999)). These include emotion state assignment, emotion blending and emotion state decay.

The EMAI architecture allows an agent to determine the emotion of an event element, event or plan by assessing the appraisal dimensions to determine the item's emotional state and then comparing this value with the pure emotions that exist within the Affective Space. The pure emotion closest in proximity to the assessment item's emotional state is said to be how the item makes the agent feel.

As initially discussed in Chapter 2, emotions are a strong factor in determining the decision that a human-like agent will make. This concept is also important in the EMAI architecture where an agent can make decisions based on how it feels and how items or events will make it feel. This is the basis for the multi-dimensional appraisal model for affective decision making and is discussed in the next chapter.

9. Multi-Dimensional Affective Decision Making

Everyone should carefully observe which way his heart draws him,
and then choose that way with all his strength.
- Hasidic Saying

9.1 Introduction

In this chapter, the concept of affective decision making using multi-dimensional appraisals is examined. In Chapter 7 a single-dimensional appraisal was used in decision making about the ordering of plans in the event space. The dimension used was *intention*. While the case studies reviewed in Chapter 7 give favourable results for creating human-like decision making mechanisms for choosing among behaviours the applications for creating human-like personalities and learning from complex affective states are limited. Choices are essentially rated as *liked* or *disliked*. This one-dimensional scale gives very limited scope for distinguishing the effects of complex emotional states on decision making. For example, on the scale of *liked* to *disliked* it might be quite simple to say the emotion *happy* exists on one end of the scale and *sadness* on the other end. However, where would the emotions of *surprise* and *anger* fit on the scale? *Anger* may be a disliked emotion like *sadness* but the way in which they differ cannot be determined on a one-dimensional scale. Furthermore, the one-dimensional appraisal model presented in Chapter 7 does not allow the agent to update its beliefs about an event based on anything more than whether the event was successful or not. An event that was executed successfully may have also been quite strenuous and therefore, not rated as a liked activity. In this case, although the event was successful, the agent may not want to perform it again and therefore, the agent's intention toward the event needs to fall appropriately. This cannot be achieved with a one-dimensional model (Baillie and Lukose 2002).

To these ends, this chapter further examines the Affective Space described in Chapter 8 and illustrates its use in providing a multi-dimensional model of affective decision making.

9.2 Event Space Appraisal

The multi-dimensional affective decision making model of the EMAI agent uses an emotional prioritising procedure to select behaviours. This process occurs in the Deliberate Area (see Chapter 3) where the agent schedules all events for execution. However, before the prioritising can begin, an event must be assigned an emotion. The emotion assigned to the event is calculated by considering the agent's emotional responses to each of the elements in the event space. As defined in Chapter 6, an event E is made up of a set of actions a, a set of objects o, occurs at a time t, and has context c, as in Expression 9.1:

$$E = \{a, o, c, t\} \tag{9.1}$$

For each element e, in an event E, the emotional state, $\mathbf{\Omega}_e$ is defined as Expression 9.2:

$$\Omega_e = \{P_e, E_e, C_e, A_e, R_e, O_e\} \tag{9.2}$$

Based on the outcome of event E, the agent will assign a weighting w, to the emotional state of each of the elements, e. As the weighting of an element and resulting emotional state with respect to an event are dynamic, the time, t, at which the emotional state is being calculated must also be taken into consideration. Therefore, the emotional state resulting from an event E, written as $\Omega_{E,t}$ is calculated as Expression 9.3:

$$\Omega_{E,t} = \sum_{e=1}^{n} w_{e,t} \Omega_{e,t} \tag{9.3}$$

where n is the number of elements associated to event E, and

$$0 \leq w_e \leq 1 \qquad \text{and} \qquad \sum_{e=1}^{n} w_e = 1$$

Once an event has been completed, each of the elements involved in the event have their emotional states updated with respect to the change in the emotional state of the agent (or the agent's mood) evoked by the outcome of the event, $\mathbf{\Omega}_{O,t+1}$. $\mathbf{\Omega}_{O,t+1}$ represents the emotional state of the agent after an event has occurred where O stands for the outcome emotion and $t+1$ is the time at which the event ended. This value is not the same as the emotional state assigned to the event after it has been executed. This is calculated later in this section.

A change in the emotional state of the agent occurs when the values for each of the appraisal dimensions (P, E, C, A, R, O) are updated during and after an event. While each of six appraisal values of an individual element involved in the event influence how these values are changed for the agent's emotional state, the final emotional state cannot be determined before the event occurs. The agent can only speculate. For example, an event the agent believes will make the agent *happy* may fail during execution, may take longer to complete than initially thought or may require extra effort. These factors would change the values of the appraisal dimensions independently of any influence over these values by the elements of an event or the event itself. The resulting emotional state in this example may be *sad* rather than the expected *happy*. Therefore, $\Omega_{O,t+1}$ cannot be calculated by combining the appraisal dimensions of the elements of an event, but can only be determined after the event has been executed. Only then can an analysis of the appraisal dimensions take place. This would include values from the appraisal dimensions of the elements in the event and also takes into consideration changes in the agent's physical and mental states. The new emotional state of the agent is used to update the values of the appraisal dimensions for each of the event elements. The agent attributes a change in its emotional state to be the result of the event and therefore, updates the emotional state of the event and its elements accordingly.

The change in the emotional state of an event is calculated using Expression 9.4:

$$\Delta_\Omega = \Omega_{O,t+1} - \Omega_{E,t} \tag{9.4}$$

After this has been calculated the emotional states for each element in the event set can be updated as Expression 9.5:

$$\Omega_{e,t+1} = \Omega_{e,t} + w_{e,t+1}\Delta_\Omega \tag{9.5}$$

Instead of the element taking on the final emotional state of the event, the previous emotional state of the element is taken into account along with the effect the element had in the resulting emotional state for the event. If the event's resulting emotional state is the same as its initial state and $w_{e,t} = w_{e,t+1}$ then the emotional state for the element will not change.

For example, assume an event *E* with two elements *A* and *B*. At the beginning of the event, elements A and B both have emotional states of *happiness*[18], that is:

[18] Values have been scaled up by a factor of 100 purely for illustrative purposes.

$$\Omega_{A,t} = [-146, -33, -46, 15, 9, -21]$$
$$\Omega_{B,t} = [-146, -33, -46, 15, 9, -21]$$

and weightings of 0.2 and 0.8 respectively, thus:

$$w_{A,t} = 0.2$$
$$w_{B,t} = 0.8$$

This would result in the emotional state for the event before execution as:

$$\begin{aligned}\Omega_{E,t} &= w_{A,t}\Omega_{A,t} + w_{B,t}\Omega_{B,t} \\ &= 0.2 \times [-146, -33, -46, 15, 9, -21] + 0.8 \times [-146, -33, -46, 15, 9, -21] \\ &= [-146, -33, -46, 15, 9, -21]\end{aligned}$$

In other words, a *happy* event. Assuming that after the event has occurred, the outcome results in an emotional state of *happiness* and *A* and *B* are still weighted the same then both *A* and *B* can be updated using Expression 9.5 as shown below.

$$\begin{aligned}\Omega_{A,t+1} &= \Omega_{A,t} + w_{A,t+1}(\Omega_{O,t+1} - \Omega_{E,t}) \\ &= ([-146, -33, -46, 15, 9, -21] + 0.2([-146, -33, -46, 15, 9, -21] \\ &\qquad - [-146, -33, -46, 15, 9, -21]) \\ &= [-146, -33, -46, 15, 9, -21]\end{aligned}$$

$$\begin{aligned}\Omega_{B,t+1} &= \Omega_{B,t} + w_{B,t+1}(\Omega_{O,t+1} - \Omega_{E,t}) \\ &= ([-146, -33, -46, 15, 9, -21] + 0.8([-146, -33, -46, 15, 9, -21] \\ &\qquad - [-146, -33, -46, 15, 9, -21]) \\ &= [-146, -33, -46, 15, 9, -21]\end{aligned}$$

When the same event's emotional state needs to be calculated in the future it will again be evaluated, using Expression 9.8, to be *happy*. However, if the final emotional state of the event changes from its initial state, the emotional states for the elements will change.

For example, imagine the same event is to occur again. This time the elements *A* and *B* are weighted initially as before (0.2 and 0.8 respectively) but after the event the emotional state is *sad* and the weightings of *A* and *B* are now 0.4 and 0.6, respectively. Figure 9.1 graphically represents the elements *A* and *B* and the event *E* before execution with a vector representing the change in emotional state of the agent (the agent's mood) after the execution of the event E.

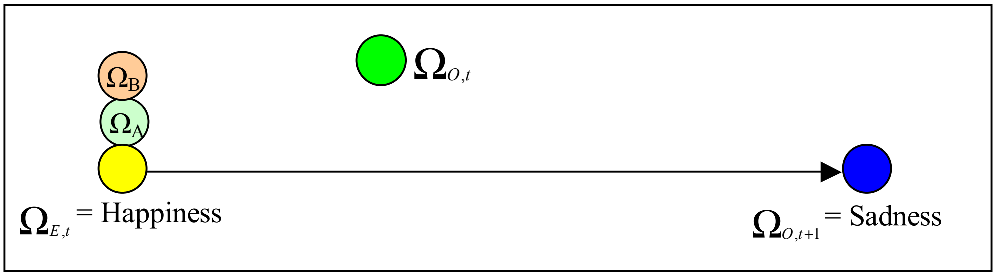


Figure 9.1 Emotional State of Event Elements before Event Execution

Given the new emotional state of *E* and the weighting of *A* and *B* after the event, the new emotional states for *A* and *B* can be calculated as:

$$\Omega_{A,t+1} = \Omega_{A,t} + w_{A,t+1}(\Omega_{O,t+1} - \Omega_{E,t})$$
$$= ([-146,-33,-46,15,9,-21] + 0.4([87,-14,0,-21,-36,151] - [-146,-33,-46,15,9,-21])$$
$$= [-146,-33,-46,15,9,-21] + [93.2,7.6,-18.4,-14.4,-18,68.8]$$
$$= [-58.8,-25.4,64.4,6,-9,47.8]$$
$$\Omega_{B,t+1} = \Omega_{B,t} + w_{B,t+1}(\Omega_{O,t+1} - \Omega_{E,t})$$
$$= ([-146,-33,-46,15,9,-21] + 0.6([87,-14,0,-21,-36,151] - [-146,-33,-46,15,9,-21])$$
$$= [-146,-33,-46,15,9,-21] + [139.8,11.4,-27.6,-21.6,-27,103.2]$$
$$= [-6.2,-21.6,-73.6,-6.6,-18,82.2]$$

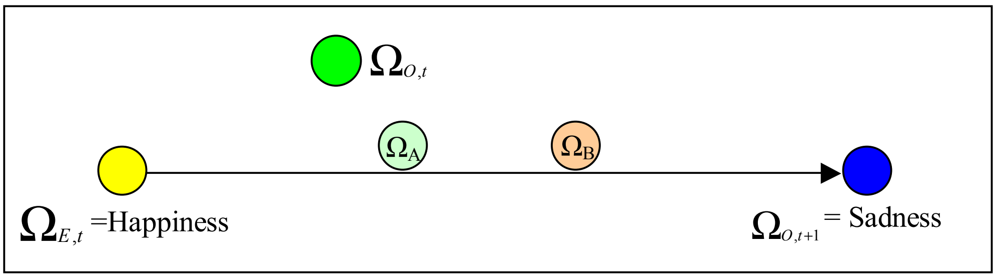


Figure 9.2 Emotional State of Event Elements after Event Execution

Essentially, this moves the emotional point for each of the elements closer to the final emotional state for the event by using the weightings to add a portion of the vector from E_t to E_{t+1} onto the emotional states for A and B. The result is represented graphically in Figure 9.2. Once each element of the event has had its emotional state updated, the emotional state for the event itself is updated using the new values for A and B and Expression 6.15 thus:

$$\begin{aligned}\Omega_{E,t+1} &= w_{A,t+1}\Omega_{A,t+1} + w_{B,t+1}\Omega_{B,t+1} \\ &= 0.4\times[-58.8,-25.4,64.4,6,-9,47.8] + 0.6\times[-6.2,-21.6,-73.6,-6.6,-18,82.2] \\ &= [-27.24, -23.12, -18.4, -1.56, -14.4, 68.44]\end{aligned}$$

This results in a new emotional state that lies between the elements of A and B as shown in Figure 9.3.

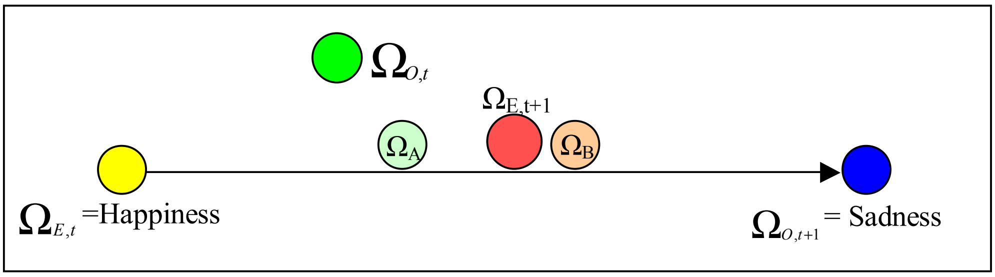


Figure 9.3 Emotional State of Event after Event Execution

Once the agent has calculated the emotional points for each of the events and in turn each element involved in an event, the agent can use this emotional point in its affective decision-making process.

9.3 Affective Decision Making

As described in Chapter 3, the EMAI's decision-making procedure is twofold. Firstly, the agent prioritises its behaviours by ordering events in the Deliberate Area based on the number of times an event has been triggered by a goal and thus representing the urgency of a event. Secondly, the agent further orders the events by calculating the resulting emotional effect that performing the event will have on the agent's emotional state.

The agent's emotional state is stored as a six-dimensional coordinate in the agent's Emotional State Register (see Section 3.3.4). This coordinate can be projected into the

Affective Space in the same way that an event's affective coordinate can be projected into the space to determine an emotional state. The agent's emotional state Ω_{EMAI} is the result of weighing the importance of an event w_E and summing the m number of resultant emotional points for all episodes of all events performed by the agent E and defined by Expression 9.6:

$$\Omega_{EMAI} = \sum_{E=1}^{m} w_E \Omega_E \qquad 9.6$$

where $0 \leq w_E \leq 1$ and $\sum_{E=1}^{m} w_E = 1$

While the emotional state is stored internally as a six-dimensional coordinate, it can also be expressed as a word by calculating the distance of the emotional state from each of the 15 discrete emotion points as defined in Table 8.1 (see Chapter 8) and selecting the name of the discrete emotion to which EMAI's emotional state is closest.

Given a number of events that have the same priority, the agent will select an event that will most likely update the agent's emotional state to a more preferred emotional state. For example, if the agent had two events of equal urgency from which to select, the agent would further prioritise these events emotionally. The agent calculates the emotional point for each event and then interpolates how this event, when performed, will update the agent's emotional state. If the agent would prefer to have an emotional state closer to *happy* it would select the event that would, when combined with its current emotional state, make the agent *happy*. Of course, the agent cannot predict exactly how a task will influence its emotional state. For example, a task that is calculated to make the agent *happy*, may fail during execution, the emotional point for this episode would shift and the effect on the agent would be different. It could instead, result in an emotional state of *angry*.

As it is impossible to visually represent in two dimensions this process in six dimensions, for the sake of explanation the emotional space will now be examined as a two-dimensional space with the dimensions of *pleasantness* and *control* as shown in Figure 9.4. For this demonstration, the agent's emotional state has been set close to *happy*.

Due to the discrete nature of the location of the emotional states within the Affective Space and as the agent's current emotional state (mood) could be located anywhere within the space, it is unlikely the agent's emotional state will commonly fit the exact location of an emotional state point. For this reason, the distance from the agent's emotional state to all of the emotional state points is calculated. The emotional state point to which the agent's mood is closest is assigned to the agent to be expressed as the agent's current emotional state. In Figure 9.4, the agent's emotional state is not at the exact location of any of the emotional state points, but it is closest to *happiness*, therefore, the agent is said to be in a *happy* state or mood.

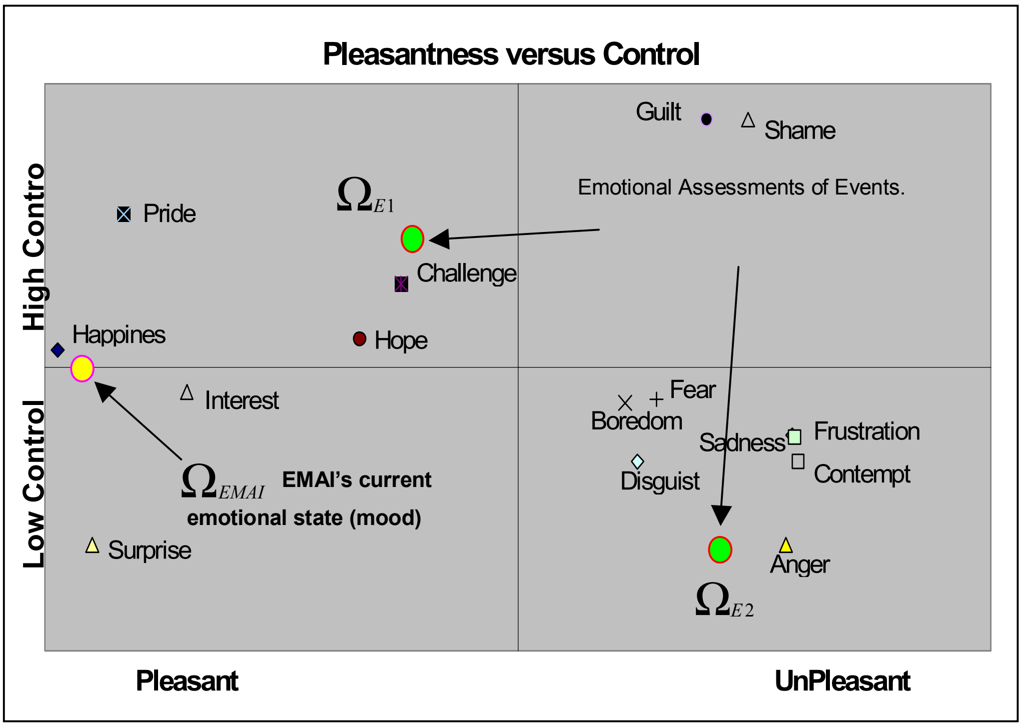


Figure 9.4 Empirical Location of Emotional States with Respect to the Pleasantness and Control Dimensions Displaying Agent's Mood (Ω_{EMA}) and Two Events' States (Ω_{E1}, Ω_{E2})

Figure 9.4 also shows the emotional location of two events (Ω_{E1} and Ω_{E2}). If these events have been given the same priority for execution, the agent needs to make an emotional decision about which event to execute. One event is closest to the emotion *challenge* and the other is closest to *anger*. The agent assesses each of the events by calculating how the execution of the event will influence the agent's mood using Expression 9.7. The equation is the same as 9.6 except the agent's mood and events are replacing the events and the event elements, respectively.

$$\Omega_{EMAI,t+1} = \Omega_{EMAI,t} + w_{E,t+1} \Delta_{\Omega,EMAI} \tag{9.7}$$

where $w_{E,t+1}$ is the weighted influence that an event has on the agent's mood and $\Delta_{\Omega,EMAI}$ is the difference between the agent's current mood and the emotional state after event *E*.

Returning to the example in Figure 9.4, if the agent is in a *happy* mood and the agent is programmed to prefer its state to be *happy*, then the agent will select the event that will keep itself closest to a *happy* mood. In this example, if the emotional points of E_1, with a weighted effect of 0.6, and E_2, with a weighted effect of 0.4, are added to the agent's current emotional state, the resulting state would be as shown in Figure 9.5 as Ω_{EMAI1} and Ω_{EMAI2}. By using a linear distance function[19], the successful execution of the challenging event E_1 would result in a mood closer to *happiness* than E_2 would. In this example the agent would select to execute E_1 before executing E_2.

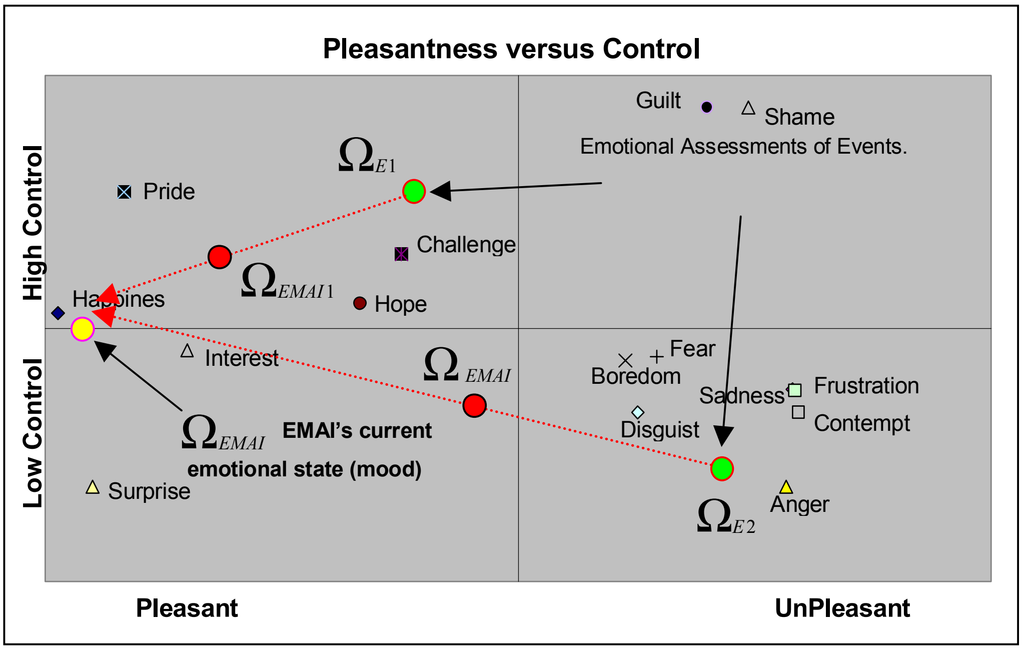


Figure 9.5 Resultant Agent Moods Generated by Combining the Agent's Current Mood (Ω_{EMA}) and Two Events' States (Ω_{E1}, Ω_{E2})

In the above example, the agent is making a choice based on its emotional perception of events as they have occurred in the past. The agent can never be certain that when it executes an event it will be successful or the same values for the six appraisal dimensions will remain the same. For example, each time an episode of an event is executed, the six appraisal dimensions are reevaluated. If the values for these dimensions remain the same, the emotional state associated with the event will not change. However, if an event that has an emotional state of *happiness* results in an outcome of *sadness* then the emotional

[19] For the initial design of this architecture, a linear distance function has been chosen for simplicity. Further investigations into this aspect of the model are needed to determine the best method for calculating emotional distances, however, this will not be explored here.

state of the event is updated to a value somewhere between the emotional states of *happiness* and *sadness*. This is shown in the example of Figure 9.4. This illustrates that although the agent can select an event based on how it predicts it will make it feel, once the event has been executed, how it actually changes the agent's mood could be quite different. This in turn updates the agent's emotional perception of the event and its elements and is used in future affective decisions.

9.4 Simulation of a Multi-Dimensional Affective Decision Making EMAI Agent

9.4.1 Objective

The objective of this simulation study is to evaluate the effectiveness of an EMAI Agent that is configured using the multi-dimensional affective decision making model described earlier in this chapter. This case study does not examine the complex behaviours of a dog nor attempt to explain the real emotions or goals of such an animal. The study is anthropomorphic[20] in the sense that it evaluates a model for human emotion via a simplistic computerised character. The same method of evaluation is used in Blumberg (1996) and El-Nasr (1998).

The simulation will attempt to demonstrate the affective decision making process based on the multi-dimensional model. A detailed evaluation of the EMAI architecture, based on this Fido configuration is described in Chapter 10.

9.4.2 Fido's Knowledge Area

Ontology

The Ontology is the very first component configured in an EMAI based agent like Fido. The Ontology is composed of the Goal Hierarchy, Type Hierarchy, Relation Hierarchy, Event Graphs, and Activity Digraphs. The Goal Hierarchy, Type Hierarchy and Relation Hierarchy are depicted in Figure 9.6, 9.7 and 9.8, respectively. The Goal Hierarchy shown in Figure 9.6 illustrates Fido's goals and subgoals. Beneath the ultimate goal, *U*, Fido has

[20] An interpretation of what is not human in terms of human or personal characteristics

three primary goals: EAT, PLAY and SLEEP. These goals decompose into a number of subgoals, some which are AND-ed and some with are OR-ed. As the goals are decomposed they become more specialised including specialist entities from the Type Hierarchy. The illustration in Figure 9.6 takes the Goal Hierarchy to third level.

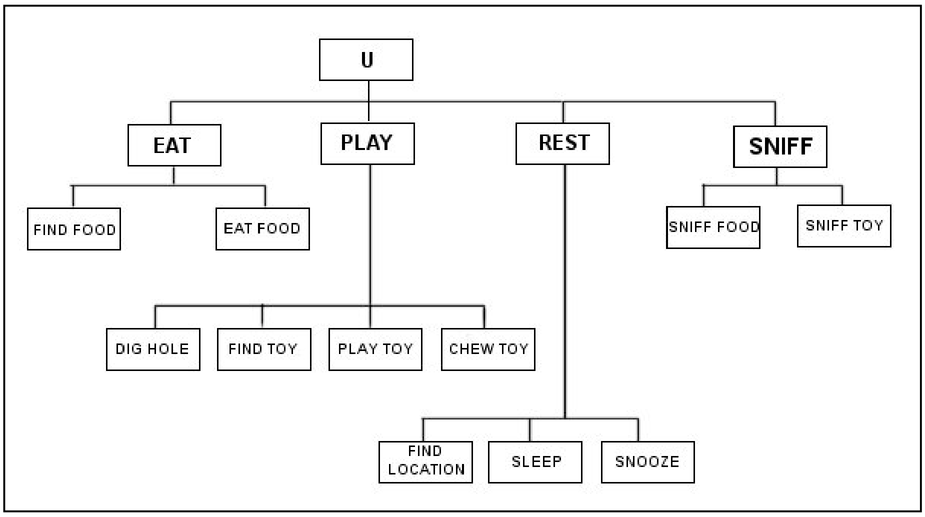


Figure 9.6 Fido's Goal Hierarchy

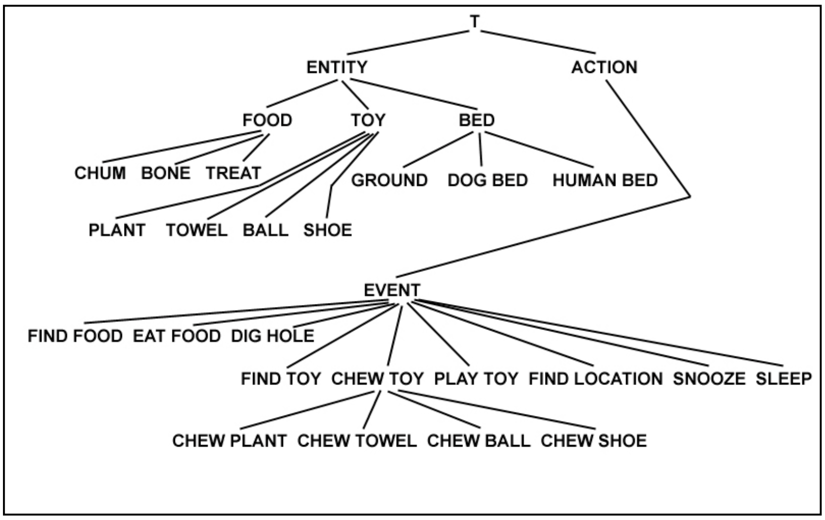


Figure 9.7 Fido's Type Hierarchy

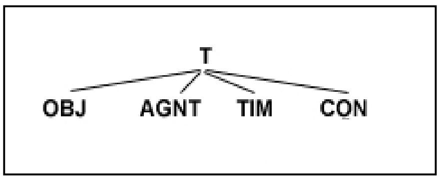


Figure 9.8 Fido's Relation Hierarchy

The decomposition of Fido's Goal Hierarchy goes beyond this, however as it is difficult to show the extent of the Goal Hierarchy to its most extreme decomposition in a diagram, the primitive subgoals are not included in the illustration. Each of the lowest level subgoals shown can be specialised further with concepts from the Type Hierarchy as shown in Figure 9.7. For example, the FIND FOOD goal can be decomposed into the subgoals of FIND BONE and FIND CHUM where the FOOD element of the goal is specialised using subtypes of the concept type FOOD from the Type Hierarchy. Furthermore, the EAT FOOD goal will decompose to the subgoals: EAT BONE and EAT CHUM; the FIND TOY goal will decompose to the subgoals: FIND PLANT, FIND TOWEL, FIND BALL and FIND SHOE, and so on.

As there are 35 Event Graphs in this Ontology, they will not be listed here. However, as an illustration, the Event Graph for FIND FOOD and the Activity Digraph for the EAT goal are shown below:

```
The Event Graph for [FIND FOOD] is
    [FIND] –
          (OBJ) -> [FOOD: BONE]
          (AGNT) -> [AGENT: FIDO]
          (TIM) -> [TIME: #now]
          (LOC) -> [BACKYARD]

Precondition:
    [AGENT: FIDO]<-(AGNT)<-[ACCESS]->(OBJ)->[FOOD: BONE]
Postcondition:
    [AGENT: FIDO]<-(POSS)->[FOOD: BONE]
Delete List:
    [AGENT: FIDO]<-(AGNT)<-[ACCESS]->(OBJ)->[FOOD: BONE]

The Activity Digraph for [EAT] is
    [FIND FOOD] -> (FBS) -> [EAT FOOD]
```

Motivational Drive Generator

Fido's Motivational Drive Generator consists of the EMAI architecture's three drive mechanisms (Homeostatic, Cyclical and Default) and includes three Internal State Registers: *hunger, fatigue* and *intrigue*. The *hunger* register depicts the agent's need for food level, the *fatigue* register depicts the agent's need for sleep and the *intrigue* register depicts the agent's need for exercise. Each register is periodically updated by the Cyclical Drive Mechanism at different rates. For example, Fido will begin to feel *hungry* before it feels sleepy. The Homeostatic Drive Mechanism monitors the values of each register. At certain threshold values, the Homeostatic Drive Mechanism sends an appropriate signal to the agent's Sensory Processor to cause goal activation. For example, when the *hunger* register reaches a particular high value the drive mechanism will inform the Sensory Processor that the agent is *hungry* and the Sensory Processor will activate a goal of EAT in the agent's Goal Hierarchy.

As it is possible for the agent to have no goals active, the Default Drive Mechanism is used to ensure the agent is never devoid of behaviour. For the Fido agent, the Default Drive Mechanism monitors the values of the registers and determines when the agent is not active. When it finds the agent to be in such a state, the Default Drive Mechanism sends a message to the Sensory Processor indicating that a default goal should be triggered. In the case of Fido, the default goal is SNIFF as shown in Figure 9.6.

When the agent completes a behaviour in response to a triggered goal, the Sensory Processor receives external sensory data. The sensory data is analysed by the Sensory Processor and if required passes relevant data onto the Homeostatic Drive Mechanism in order to update the values on the Internal State Registers. For example, when the agent finishes the event associated with the EAT goal, the Sensory Processor will receive data that represents that the agent has eaten and how much it has eaten. This data is passed to the Homeostatic Drive Mechanism so that it can reduce the agent's *hunger* by reducing the value on the *hunger* register.

Affective Space

Fido's Affective Space is configured with 15 pure emotion points from Smith and Ellsworth's experimentation (Smith and Ellsworth 1985) as outlined earlier in Chapter 8. The values of the emotion points from Smith and Ellsworth's original study have been scaled up by a factor of 100 purely for illustrative purposes.

9.4.3 Fido's Interface

Figure 9.9 The Fido Interface

A simple interface is set up for the Fido simulation. The interface provides the user with the means to interact with the Fido agent as well as observe its behaviour, motivational urges and emotions. The interface is shown in Figure 9.9.

Motivational Urges

The motivational urges section of the interface consists of three urge gauges. These can be seen at the top of Figure 9.9. These gauges are visual representations of the agent's Internal State Registers: *hunger*, *fatigue* and *intrigue*. The registers are controlled by the simulation and the user can not adjust the values.

Environmental Display

The Environmental Display is the large black window in the center of the interface as shown in Figure 9.9. In this area, an image of Fido is displayed. The image and simple animations of Fido illustrate Fido's behaviours and emotions. Along with the image of Fido, the environmental display also shows which items the user has placed in Fido's environment. These are displayed as small individual icons.

Simulation Narration

The textbox under the Environmental Display (in Figure 9.9) contains a narration of the simulation. This keeps a record of Fido's activities, his emotional states and the user's interaction. The user can use the scrollbar at the side to review their interactions with Fido.

Tool Bar

The tool bar consists of a series of buttons that sits below the simulation narrative textbox. The first set of buttons allows the user to put items in and take items out of Fido's environment. When the user adds an item to the environment, a corresponding icon will appear in the Environmental Display. The second part of the tool bar is two buttons that control the running of the simulation. The pause button allows the user to pause the simulation at anytime to review the narration and the start button initiates the simulation.

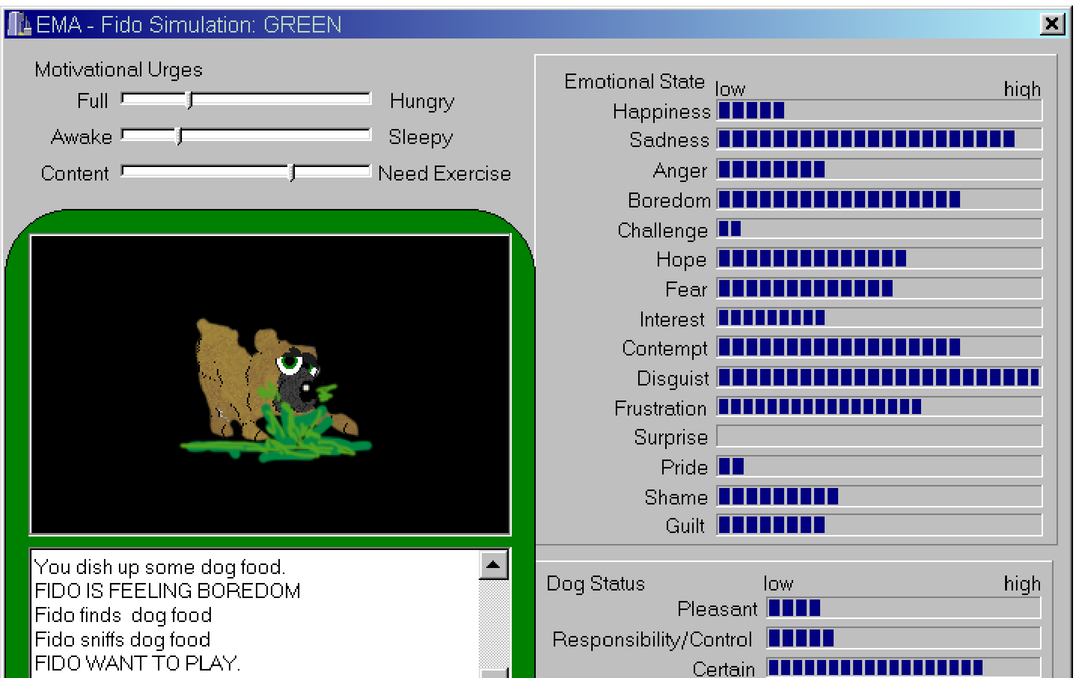
Discipline Buttons

The final items on the Fido interface are the two large buttons that display a *sad* and a *happy* yellow face. These buttons can be pressed by the user to inform Fido of whether or not he is bad or good. The buttons can be pressed any number of times. For example, if Fido is chewing the shoe, the user can click on the BAD DOG button several times to emphasize to Fido just *how* bad he is.

Expressing Emotions

As the assessment of Fido is carried out by an evaluator (who would assess FIDO's emotional behaviour), the interface needs to include a way of expressing Fido's emotional state. Figure 9.10 includes a series of gauges that represent the values of the Affective Space's six appraisal dimensions and the strength of Fido's emotional state with respect to the 15 pure emotions. This interface clearly shows the agent's state with respect to the six appraisal dimensions of its Affective Space and the agent's emotional state. The emotional state is presented as a series of 15 gauges, one for each pure emotion. The higher the blue line on the gauge, the closer the agent's mood is to that pure emotion.

After initial evaluations with the extended interface, it was found that the evaluators were not taking any notice of the gauges and insisted they provided too much information. Therefore, they have been hidden for the duration of the simulation presented in Chapter 10.



The evaluators are able to determine the emotional state of Fido through the simulation narrative and the agent's animation. A different animation sequence exists for each of the 15 pure emotions as shown in Figure 9.11. As the evaluators are evaluating the emotional behaviour of the Fido, they do not need to know about the complex internal workings of the architecture. If Fido is to receive favourable feedback from its outward emotional behaviours, the evaluators need to determine if Fido is demonstrating reasonable emotional behaviour given the environment and Fido's motivations. The internal workings of the agent need to be disguised. If the EMAI architecture were used to generate an intelligent and believable character in a virtual environment, the external behaviour of the agent would have to be reasonable enough to create a suspension of disbelief in an onlooker. Making the internal computations visible to a viewer defeats this purpose.

Emotion	**Fido Animation Frames**	**Emotion**	**Fido Animation Frames**
Happiness		Sadness	
Anger		Boredom	
Challenge		Hope	
Fear		Interest	

9.4.4 Fido's Affective Decision Making Process

Fido's affective decision making process begins when any of the Internal State Registers of the Motivational Drive Generator reach their threshold values. When the drive mechanism senses that an Internal State Register has reached a threshold value, they send appropriate signals to the Sensory Processor indicating the state of the agent. The Sensory Processor determines the most suitable goal to activate in the agent's Goal Hierarchy. This goal is passed to the Event Space Generator where plans of behaviour are devised that will satisfy the agent's goal.

Before making a decision, the agent generates an event space for each choices that it can make. From this event space, an emotional state is evoked based on the assessment in the six appraisal dimensions of the elements in the event space by using the action, object, time and context of the event. Given the emotional states generated by each choice, the agent selects the event that it perceives will make it the happiest. Happiness has been chosen as the value for the Fido agent's Emotional State Bias Register. This of course could have been set to make the agent gravitate toward any emotion, preferred cognitive state or any combination of these. However, for simulating a pet dog, it is envisaged that *happiness* would be an appropriate preferred emotional state.

As the agent may have more than one goal active at any time, each goal is given a priority rating that is used by the agent to determine the urgency of the corresponding event. The priority rating of a goal is proportional to the value on the Internal State Register/s that triggered it (see Section 5.2.4). When there exists a number of plans (represented as Activity Digraphs, refer to Section 4.3.2) for a particular goal, the agent uses its emotional decision making abilities to make a choice.

Here is a typical simulation run of Fido. Initially, all of Fido's goals are activated, one of which is the PLAY goal. The events that Fido determines will satisfy the PLAY goal are DIG HOLE, CHEW TOY and PLAY TOY. The elements involved in each of these events are shown below.

$E_{DIG\ HOLE}$={DIG, GROUND}
$E_{CHEW\ TOY}$={CHEW, TOY}
$E_{PLAY\ TOY}$={PLAY, TOY}

The Event Space Generator specialises these events using subtypes from the Type Hierarchy to produce the following events in the event space for the PLAY goal.

$E_{DIG\ HOLE}$={DIG, GROUND}
$E_{CHEW\ PLANT}$={CHEW, PLANT}
$E_{CHEW\ TOWEL}$={CHEW, TOWEL}
$E_{CHEW\ BALL}$={CHEW, BALL}
$E_{CHEW\ SHOE}$={CHEW, SHOE}
$E_{PLAY\ PLANT}$={PLAY, PLANT}
$E_{PLAY\ TOWEL}$={PLAY, TOWEL}
$E_{PLAY\ BALL}$={PLAY, BALL}
$E_{PLAY\ SHOE}$={PLAY, SHOE}

At the beginning of a simulation, Fido will have no attitudes toward any of these events or there elements. Therefore, the emotional state (Ω) for each will be calculated as shown below using the event PLAY BALL as an example. In this example, it is assumed the weightings on objects are 0.8 and the weightings on actions are 0.2.

$$\Omega_{PLAY,t} = [0,0,0,0,0,0]$$
$$\Omega_{BALL,t} = [0,0,0,0,0,0]$$

$$w_{PLAY,t} = 0.2$$
$$w_{BALL,t} = 0.8$$

This would result in the emotional state for the PLAY event (calculated using Expression 9.6) before execution as:

$$\begin{aligned}\Omega_{PLAY\_BALL,t} &= w_{PLAY,t}\Omega_{PLAY,t} + w_{BALL,t}\Omega_{BALL,t} \\ &= 0.2 \times [0,0,0,0,0,0] + 0.8 \times [0,0,0,0,0,0] \\ &= [0,0,0,0,0,0]\end{aligned}$$

The emotion closest to the emotional state of [0,0,0,0,0,0] is *hope*. As all of the elements in the event space for PLAY have not been interacted with before, they will all have emotional states of *hope* and in turn so will the events of which they are a part. Therefore, when the agent comes to make a decision about which event to choose based on the one that will make the agent most *happy* each event is equally as capable of achieving this. Assuming the agent's current mood is *happiness* and the PLAY BALL event is weighted as 0.8, by using Expression 9.9 the agent will predict the outcome emotional state to be:

$$\begin{aligned}\Omega_{FIDO,t+1} &= \Omega_{FIDO,t} + w_{PLAY\_BALL,t}\left(\Omega_{PLAY\_BALL,t} - \Omega_{FIDO,t}\right) \\ &= [-146,-33,-46,15,9,-21] + 0.8 \times ([0,0,0,0,0,0] - [-146,-33,-46,15,9,-21]) \\ &= [-146,-33,-46,15,9,-21] + [117,26,37,-12,-7,17] \\ &= [-27,-7,-9,\ 3,\ 2,-4]\end{aligned}$$

The predicted emotional state for Fido, $\Omega_{FIDO,t+1}$, is closest in the Affective Space to the emotion of *happiness*. Therefore, the agent will predict that the PLAY BALL event will keep the agent in a *happy* state but with not as much strength as before as the predicted emotional state point is further away in proximity to the pure *happiness* point than the agent's current mood.

If the weighting of the other events were also 0.3, the predictions of how they would make Fido feel (using the same calculation as above) would also result in Fido remaining *happy*. However, if for example, the DIG HOLE event had a weighting of 0.9 the predicted emotional state would be:

$$\Omega_{FIDO,t+1} = \Omega_{FIDO,t} + w_{DIG\_HOLE,t}\left(\Omega_{DIG\_HOLE,t} - \Omega_{FIDO,t}\right)$$

$$= [-146, -33, -46, 15, 9, -21] + 0.3\times([0,0,0,0,0,0] - [-146, -33, -46, 15, 9, -21])$$

$$= [-146, -33, -46, 15, 9, -21] + [44, 10, 14, -5, -3, 6]$$

$$= [-102, -23, -32, 10, 6, -15]$$

Compared with 15 emotion points in the Affective Space, this value is also in closest proximity to the *happiness* pure emotion, however, it is further away from it than the emotional state determined by the PLAY BALL event. The distance from the predicted outcome of PLAY BALL to *happiness* is 30.16 and the distance from the predicted outcome of DIG HOLE to *happiness* is 109.51. Therefore, the agent would choose to perform the PLAY BALL event before the DIG HOLE event.

In the situation where a stalemate occurs between competing events (when each predicted outcome is the same distance from *happiness*), it will not matter which event the agent selects and hence it uses a FIFO method for selection. The event that is placed in the agent's schedule first will be the one executed.

Assuming that Fido selects the event PLAY BALL, the weighting remains at 0.8 and the actual[21] resulting emotional state outcome of the event is *frustration* (for example, the ball is thrown and Fido can not find it), the new mood of the agent can be calculated using Expression 9.10 thus:

---

[21] The *predicted* emotional state outcome of an event is determined *before* the event is executed. There can be no guarantee or way of calculating what the *actual*

$$\Omega_{FIDO,t+1} = \Omega_{FIDO,t} + w_{PLAY\_BALL,t+1}\Delta_{\Omega,FIDO}$$
$$\Omega_{FIDO,t+1} = [-146,-33,-46,15,9,-21] + 0.8([88,-37,-8,60,48,22]-[-146,-33,-46,15,9,-21])$$
$$\Omega_{FIDO,t+1} = [-146,-33,-46,15,9,-21] + 0.8\times[234,-4,38,45,39,43]$$
$$\Omega_{FIDO,t+1} = [-146,-33,-46,15,9,-21] + [187,-3,30,36,31,34]$$
$$\Omega_{FIDO,t+1} = [41,-36,-16,51,40,13]$$

Although this point is not at the exact location of the point for pure *frustration*, compared with the other 15 emotion points, the agent's new mood is closest to the pure *frustration* point and therefore, the agent is said to be feeling *frustration*. After the event has occurred and the new agent mood has been determined, the agent update's the emotional states of the events and elements involved in the event. This may or may not influence the emotional states of other events depending on overlapping elements.

Depending on the outcome of the event (success or failure), the agent updates its Internal State Variables to reflect the associated behaviour that has taken place. If the case occurs where an Internal State Variable is returned to a normal level inside the thresholds, the Homeostatic Drive Mechanism will send data to the Sensory Processor that will deactivate the associated goal. The event space will subsequently be removed from the agent's schedule.

The next time a goal is triggered and the agent generates an event space, the updated values for the elements in the events (calculated after the agent's last behaviour) are used in predicting the emotional state of the events. The affective decision making process begins again using emotional experiences gained from previous behaviour.

The multi-dimensional affective decision making process implemented in the Fido agent, built upon the EMAI architecture, is evaluated in the next chapter.

9.5 Summary

As the one-dimensional model of affective decision making as examined in Chapter 7 does not provide the necessary depth of representation needed to build a complex agent that can make decisions based on multiple appraisal dimensions, the Affective Space concept was included in the EMAI architecture and introduced in Chapter 8. This chapter described how the Affective Space is used for multi-dimensional affective decision making and belief revision in an artificial intelligence.

This multi-dimensional approach to synthesising emotions in artificial beings is a powerful concept. In a world where many decision-making situations have numerous outcomes with an equal quality of solution, this type of reasoning is imperative in humans. If attempts are to be made to produce artificial intelligences with this human-like mentality, the psychological theories of emotion need to be explored and implemented in artificial agent architectures and tested for human-like affective decision-making capabilities.

In the case study presented in this chapter, an EMAI agent modelled a computer character called Fido and demonstrated the use of the Affective Space adapted from a model proposed by Smith and Ellsworth (1985). Fido was programmed to simulate the behaviours of a simple pet dog. The behaviours of a pet dog were chosen as most people would be able to identify with and have expectations of how a pet dog should behave as the behaviours and goals of a pet dog are simple. The simulation presented in the case study examined how the Affective Space can be used to generate different emotional states within the agent and used in the affective decision-making process.

The inclusion of such an Affective Space in the EMAI architecture has removed the need for static production rules in determining an emotional state for the agent and to assist in decision making. The agent is able to assign emotional perceptions to events and elementary components of events and use these perceptions to make choices based on the agent's mood and its prediction of how events will influence the agent's current mood.

During execution of events, the agent is also able to update its emotional perception of events and their elements as episodes of events succeed and fail. This leads to a dynamic process within the agent that allows for emotional perceptions to change and in turn affect future decisions, something lacking in current emotional agent architectures.

The next chapter presents an evaluation of Fido. This simulation is presented to a number of human participants as an experiment in the evaluation of the EMAI architecture's ability to generate reasonable emotional states and the use of these emotions for intelligent decision making.

10. Evaluation of the EMAI Architecture

You cannot teach an old dog new tricks – W.Camden, Remaines Concerning Britaine, 1614.

10.1 Introduction

The previous two chapters examined the central focus of the EMAI architecture: the Affective Space. This concept is used in an EMAI agent to produce multi-dimensional affective decision making capabilities. It allows an agent to make emotional assessments about elements and events for the purpose of using these evaluations in reasoning and decision making.

The ability of the EMAI agent to make reasonable decisions and in turn produce reasonable behaviours based on emotional assessments is important (Baillie 2002). Many of the previous models of emotional artificial intelligence (PETEEI (El-Nasr 1998), EM (Reilly 1996)) were used in the creation of computer based characters and assessed through user interaction. Outward behaviour makes a character appear real and only by assessing this can the internal mechanisms be evaluated. The EMAI architecture was designed with this emotional decision making in mind and therefore, the agent's capability to use emotions as a basis for reasoning, decision-making and behaving is paramount. The concepts of goals (Chapter 5), attitudes (Chapter 6) and an Affective Space (Chapter 8) have never before been integrated into an agent for the purpose of building an emotional artificial intelligence. The extent to which the resulting architecture is emotionally intelligent is a question of evaluation by those who naturally have emotional intelligence, human beings.

For the purpose of the evaluation, an EMAI agent is used to model a pet dog called Fido (see Section 9.4). This is in keeping with the examples already used throughout this dissertation as well as providing the human evaluator with a familiar character with which to interact. Pet dogs have been the characters of choice for other emotional agent evaluations (El-Nasr 1998) as a simple simulation of a pet dog can easily be constructed from a limited number of goals and events. Most people who evaluate the agent's performance will have some idea and expectation about how a pet dog should act.

This chapter describes the experimental procedures used in the evaluation of EMA, the results collected from the evaluator questionnaires and a discussion of the results.

10.2 Evaluation Technique

While some studies have examined the most appropriate types of behaviours that should exist in domain dependent agents (Cohen et al. 1989, Pollack and Ringuette 1990, Hanks et al. 1993), little work has been done on the development of a general methodology for the holistic evaluation of agents. One methodology that has been designed is that of Wallace and Laird (1999). Their methodology evaluates an agent's architecture based on the number of resources in terms of the amount of knowledge required to achieve a specified behaviour and the amount of time required to generate that behaviour. Such a methodology may be appropriate for comparing and analysing complete agent architectures. However, it is not appropriate for evaluating the EMAI architecture because the EMAI architecture is incomplete and it represents a new architectural mechanism for the generation and use of emotional states in decision-making. Therefore, it is more appropriate to evaluate the emotional decision making abilities of the agent.

The methodology for evaluation used by other researches in the affective computing domain (see Chapter 2) has been to implement a simple user interface that allows a user to interact with the agent either through natural language or point-and-click methods. Through interaction with the agent the user is able to evaluate, on a scale of reasonability, the emotions produced by the agent and in turn the behaviours exhibited by the agent. The human evaluator is then asked to complete a questionnaire about their experience with the agent. To ensure that each user is given the same chance to experience similar activities and interactions with the agent, the method of walk-through exercises (Baeker et al. 1995) is implemented. Each exercise guides the user through an interaction with the agent.

The primary use of these agent architectures has been in the development of believable artificial characters, either human or animal-like. The aim in creating such artificial beings is for use in interactive fiction (computer games and virtual reality). In this domain, the agent should produce performances that create a suspension of disbelief for the user. To create this illusion the agent should be able to simulate emotion, motivation, and personality (Pisanich and Prevost 1996). As the generation of motivation and emotion are the primary functions of the EMAI architecture, the same method of evaluation (as above) has been used.

10.3 Outline of Experimental Procedure

The evaluation of the EMAI architecture was conducted using 18 volunteer evaluators who were asked to contribute between one and one and a half hours of their time for the evaluation of the software. The volunteers were recruited through a broadcast email to first year computer science students and staff members at the University of Southern Queensland. The evaluators included nine members of the public and nine first year university students. An introductory verbal explanation and walk-through of the EMAI agent interface was presented to each participant. Evaluators were notified that their responses would be used to evaluate the design of the architecture from which the software was created and their feedback would be used to further enhance the design.

The evaluators were given an instruction sheet that outlined the user interface and explained the function of each of the buttons, textboxes and gauges that appeared on the interface. Once each participant was clear on how the interface functioned, they were asked to begin the simulation of the pet dog *Fido*. The evaluators were given a set of five walk-through exercises to perform with the agent. This ensured that each participant had a similar interactive experience with Fido. While running the simulation the evaluators were asked to fill out the questionnaire to evaluate different aspects of the agent, including how the agent set goals, produced emotions and behaved. The simulation produced a narrative as it ran. Stop and Start buttons on the user interface allowed the participant to stop the simulation and read back over the narrative before answering any of the questions. The instruction sheet and questionnaire are included in Appendix D.

The procedure mentioned above was repeated for a second model of the agent. This second model, without the participant's knowledge, simulated the same pet dog Fido, but this time with random goal setting, random emotion generation and random behaviours. This provided a baseline for the Fido model that used the EMAI architecture. If a participant had not experienced this type of interaction with a computer character before, their initial response to the EMAI model could have been positive to the extent that it biased their judgment and responses on the questionnaire. By including a Fido character with random emotions and behaviours it provided a calibration point for the data collected from the survey about the EMAI model. Due to the random model's randomised behaviours, it was expected that some of the time the agent would act in a manner that the evaluators think is appropriate and other times it would not. Half of the evaluators began their evaluations with the EMAI model. This was followed by an evaluation of the random model. The other half of evaluators began their evaluations with the random model and continued afterwards with the EMAI model. The evaluators were not aware which model was which. Evaluators completed the entire evaluation questionnaire with one agent model before beginning again with the other.

Evaluators were assured of their anonymity and to this end, no personal details were collected.

10.4 Results

The evaluation of the EMAI architecture was performed over a number of weeks as people volunteered to participate in the experiment. On completion of the questionnaires, the evaluations were collected and data was entered for analysis. In all, 18 evaluations were returned.

Evaluators completed a number of walk-through exercises. To prompt evaluators to think about their experience with the EMAI agent and the random model throughout these exercises, at the end of each exercise, the evaluators were asked to complete several questions relating to the exercise. Following this the evaluators were asked a series of questions to obtain an overview of their experience with the EMAI agent and the random model.

The results are given as a collection of two types of evaluator responses. For some questions the evaluators gave simple *yes/no* answers. The computed χ^2 statistic and P-value accompany these results. For other questions the evaluators were asked to rate certain characteristics of the agent on a seven-point scale from 1 to 7. Unless otherwise stated, throughout the following results 1 represents unreasonable and 7 represents reasonable. Included with these rated answers are the statistics for a matched pairs *t* test and associated P-values.

10.4.1 Exercise 1

Exercise 1 asked the evaluators to perform the following list of instructions with Fido:

- ⇒ Run Fido.
- ⇒ Before beginning run the mouse over each part of the screen. Hint boxes will pop up that informs you what each part/button is for.
- ⇒ Click on the Start Button to Begin. Notice that Fido initially has low motivation for all of the motivational urge gauges (at the top). As the values on the motivational urge gauges go up, Fido will begin to do things.
- ⇒ For this exercise just watch what Fido does without interacting with him.
- ⇒ When Fido's *hungry* gauge reaches the extreme right, wait until "FIDO IS HUNGRY" has printed in the narrative box 3 times and press the pause button.
- ⇒ Use your experience in watching Fido and the narration in the text box to answer the questions about exercise 1 on the questionnaire.
- ⇒ Close Fido

The purpose of this exercise was to give the user time to adjust to the Fido software's interface, the simulation of Fido and to become comfortable reading Fido's emotional and physical states as well as his default behaviours without user interaction.

The results for this question where rated from 1 being erratic to 7 being calculated. The question and results are shown in Table 10.1 along with the 95% confidence interval for the data.

Question	**EMAI**		**Random**		***t***	**P**
	Mean	**Interval**	**Mean**	**Interval**		
How did the progression of Fido's behaviour appear to you?	5.33	0.023	3.11	0.027	3.87	<0.001

Table 10.1 Behaviour Progression Evaluations

The next question was asked in order to focus the evaluator's attention on the motivational urge gauges. In the EMAI model of Fido these gauges represented the Motivational Drive Generator (see Chapter 3) of the agent. The values of the gauges (*hunger*, *fatigue* and stimulation) where dynamic and changed with time and the behaviours of the agent. At certain threshold levels the gauges would trigger appropriate goals in the agent's Goal Hierarchy. In the random model the gauge values changed randomly and did not affect the agent's goals. The question asked was:

Did you feel that Fido's behaviour was driven by the values on the motivational urge gauges (the three gauges at the top of the window)?

The answers for this question were given as either *Yes* or *No*. The results showed that 94% of the evaluators were aware that the motivational urges were affecting the behaviour of the EMAI agent. For the random agent, only 44% of the evaluators thought the motivational gauges where affecting the behaviour.

The next two questions asked the evaluator about Fido's changing emotions and behaviours. These questions were asked to prompt the evaluator to take notice of and reflect on the emotions that were being produced by the models and the resulting behaviours.

The results given in Table 10.2 show that the evaluators thought the EMAI model's behaviour with respect to changing emotions and changing behaviours was reasonable. For the random model the evaluators did not entirely rate Fido's emotions and behaviours as unreasonable but did not rate them as highly as the EMAI agent. The random model's changing emotions were regarded as slightly more reasonable than the resulting behaviours.

	6.05	0.012	4.22	0.020	4.34	<0.0005
	6.00	0.012	3.39	0.021	5.61	<0.0005

Table 10.2 Changing Emotional State and Behaviours Evaluation

Exercise 1 was designed to give the evaluators a voyeuristic view of the Fido simulations in action. During this exercise, evaluators were not able to interact with the Fido models: the following exercise introduced user interaction.

10.4.2 Exercise 2

Exercise 2 asked the evaluators to perform the following list of instructions with Fido:

> ⇒ Run Fido.
> ⇒ For this exercise use the BAD DOG and GOOD DOG buttons ONLY.
> ⇒ By using the GOOD DOG button to praise Fido when he performs a good action and the BAD DOG button to reprimand Fido when he is bad try and teach Fido to prefer *playing with the ground* (rolling about) rather than other activities. Continue this until the *hungry* gauge reaches the extreme right, wait unit "FIDO IS HUNGRY" has printed twice. Press the pause button.
> ⇒ Use your experience in watching Fido and the narration in the text box to answer the questions about exercise 2 on the questionnaire. DO NOT CLOSE FIDO WHEN FINISHED.
> ⇒ After answering the exercise 2 questions, continue to Exercise 3 with the same Fido program open.

This exercise built upon Exercise 1 by allowing the evaluator to experience Fido's default behaviours. However this time the evaluators were able to interact with Fido and discipline him. The evaluators were allowed to select either the "GOOD DOG" or "BAD DOG" buttons to praise or discipline Fido respectively. These buttons allowed the user to give positive or negative reinforcement to the behaviours being performed by Fido. The EMAI Fido model associated a pleasant emotional state with the "GOOD DOG" button and an unpleasant emotional state with the "BAD DOG" button. When either button was pressed, EMAI Fido would update the emotional states for each of the elements for the current event. The random Fido simulation applied random emotional states when either of the buttons was pressed.

The first two questions related to ease of training and whether or not the evaluators thought Fido learned anything during the simulation. The goal was to determine if the evaluators thought they could influence Fido's behaviour. The questions were:

Did you find Fido easy to train?

Do you think that Fido learned anything during the training?

94% of the evaluators thought the EMAI model was easy to train opposed to the 11% who believed the random model was trainable. Not such a convincing difference in results was given toward the agents' learning, with 88% of evaluators believing the EMAI model learned and 33% believing the random model learned. Even though only 11% of evaluators thought the random Fido was difficult to train, a third thought the agent was learning throughout the interaction.

The next question asked if the evaluators thought Fido's emotional states where related to the way in which they used the "GOOD DOG" and "BAD DOG" buttons. The question and results are shown in Table 10.3. This data shows that the evaluators thought the EMAI model reacted quite reasonably to the way in which they applied the "GOOD DOG" and "BAD DOG" buttons and the random model reacted quite unreasonably.

techniques?	6.05	0.012	3.44	0.027	4.00	<0.0005

Table 10.3 Emotional State During Training

A second part to the previous question asked each evaluator to briefly explain their training techniques and how they thought Fido responded.

Answers for this question were very similar across evaluators. In summary, they explained that when Fido performed an act they thought was bad, they would click on the "BAD DOG" button and when Fido was good, they clicked on the "GOOD DOG" button. One participant wrote:

> *When he chewed something or raised his leg I click on the BAD DOG button, when he slept or rolled on the ground I clicked on the GOOD DOG button.*

The next exercise involved further user interaction with Fido and asked the evaluators to monitor the agent's behaviours and motivational states. This exercise was continued from Exercise 2 where the Fido models retained the knowledge obtained during Exercise 2.

10.4.3 Exercise 3

Exercise 3 asked the evaluators to perform the following list of instructions with simulated Fido:

- ⇒ Continuing from the last exercise, press the start button to start the simulation going.
- ⇒ Click on the Food Button to give Fido some food.
- ⇒ Once Fido has finished eating the food, take notice of his emotional state and his next behaviour.
- ⇒ When Fido is *hungry* again, click on the Food Button again. Continue this until you have served up about 5 lots of dog food. NOTE: You cannot add more dog food until he has begun eating the lot before.
- ⇒ Click on pause and answer the questions for exercise 3.
- ⇒ Close Fido

The objective of this exercise was to give evaluators experience in monitoring Fido's motivational urges, how the values of these urges affected its behaviour and whether or not Fido retained previously learned knowledge from Exercise 2.

The first two questions were asked to gauge the amount of attention that each evaluator gave to Fido's motivational state and how they thought this state affected its behaviour. The third question asked the evaluators to rate the eating behaviours as related to the motivational urges. The questions and results are shown in Table 10.4. This data shows the evaluators noticed that most of the time the EMAI model ate after it expressed *hunger* and that in most of these instances its next behaviour after expressing *hunger* was eating. However, the evaluators thought the random model ate after it expressed *hunger* around half of the time but very rarely did it eat directly after expressing *hunger*.

Did Fido eat only after he expressed he was hungry?	6.00	0.026	3.78	0.025	3.42	<0.0025
When Fido was hungry and there was food available did Fido eat the food before doing something else?	6.05	0.028	2.00	0.019	7.70	<0.0005
How would you rate this type of behaviour?	6.00	0.013	3.22	0.026	5.61	<0.0005

Table 10.4 Motivational Urge Affects on Behaviour

To further analyse Fido's eating behaviour the evaluators were asked the following question:

During the times when he was hungry, did Fido seem to retain knowledge of his previous training?

It is clear from this data that the evaluators considered the EMAI model to retain knowledge far more than the random model as 83% of the evaluators believed the EMAI model did, against 11% for the random model. In reality, the EMAI model retained all of its knowledge and the random model retained none.

The evaluators rated the EMAI model's behaviour to be more reasonable than the random model's behaviour even though the random model's behaviour was not rated as being entirely unreasonable.

The next exercise allowed the participant further interaction with the agent by allowing the users to add items to the agent's environments and attempt to train the behaviour of the agent with respect to the items.

10.4.4 Exercise 4

Exercise 4 asked the evaluators to perform the following list of instructions with Fido:

⇒ Run Fido.
⇒ Add the ball and towel to the environment.
⇒ Teach Fido that playing with the ball is GOOD and playing with the towel is BAD.
⇒ Do not forget to give Fido dog food when he appears *hungry*!!
⇒ When you think that Fido has been trained, attempt to reverse the Fido's training. See if you can make him think that the towel is GOOD and the ball is BAD.
⇒ When you think you have done this take the towel away. Observe Fido's behaviour.
⇒ After a while, add the towel back into the environment.
⇒ Click on pause and answer the questions for exercise 4.
⇒ Close Fido

The objective of this exercise was to give the user an opportunity to add new items into Fido's environment and to allow the user to train Fido with the use of the objects. The simulation allowed several items to be added to the environment by clicking on appropriate buttons in the tool bar. When items were added to the environment Fido would attempt different actions with the items. For example, when the "BALL" was added to the environment, Fido could choose to "PLAY" with the "BALL" or "CHEW" the "BALL".

For this exercise the evaluators were asked to add a “BALL” and a “TOWEL” into the environment and to train Fido’s actions toward the objects using the “GOOD DOG” and “BAD DOG” buttons. Initially the user had to train Fido to prefer playing with the “BALL” and dislike the “TOWEL” and then the user had to attempt to reverse the agent’s attitude towards each object.

Both agents were programmed with the same set of goals. However, the EMAI model’s goals were triggered by the agent’s motivational gauges and the random model’s goals where triggered at random time intervals during the simulation. The EMAI model would select its behaviours based on the events that would make the agent the *happiest* and satisfy the current set of goals. The random model randomly selected events whether or not they satisfied the goals and whether or not they made the agent *happy*.

The evaluators were asked several questions about the training of Fido. The results are shown in Table 10.5. The evaluators rated the EMAI model as quite easy to train and the random model as almost impossible to train. The results also show the evaluators thought that EMAI model produced reasonable emotional states while the random model produced closer to unreasonable emotional states.

	6.25	0.015	1.78	0.014	9.57	<0.0005
	6.50	0.009	2.61	0.020	10.05	<0.0005

Table 10.5 Fido’s Training

The next question asked the evaluators to reflect on whether or not they thought the agents prioritised their behaviours.

Overall, do you think that Fido prioritises his behaviours?

A unanimous result of *Yes* (100%) was obtained for the EMAI model, while only 13% of the evaluators thought the random model prioritised its behaviours.

The last exercise in the evaluation gave the evaluators even more degrees of freedom with respect to user interaction with Fido. Exercise 5 allowed the evaluators to create their own training regime.

10.4.5 Exercise 5

Exercise 5 asked the evaluators to perform the following list of instructions with Fido:

⇒ Run Fido.
⇒ For this exercise it is your task to train Fido to like an item of your choice for playing with. Add all of Fido's items to the environment including food items and try to teach Fido to prefer playing with and eating certain objects. Get Fido to have several favourite items, so if you take one away he plays with the other.
⇒ Observe Fido's behaviour.
⇒ Click on pause and answer the questions for exercise 5.
⇒ Close Fido

This final exercise allowed each participant to have some free reign over their interaction with Fido. The participant was permitted to use the interactions from past exercises to formulate their own training exercise with Fido. Again the evaluators were asked questions relating to ease of training and emotional states.

The results are shown in Table 10.6. Evaluators thought the EMAI model was relatively easy to train and the random model was near impossible to train. The results show that the EMAI model was far more capable of producing reasonable emotional states with respect to training as compared with the random model.

	6.00	0.016	1.78	0.016	12.84	<0.0005
	6.28	0.011	2.22	0.021	10.14	<0.0005

Table 10.6 Ease of Training Fido in Multiple Tasks

The next question took a slightly different perspective and asked the evaluators to assess the agent's ability to develop attitudes toward objects in its environment. The question asked was:

Do you think that Fido was able to devise a list of favourite and least favourite items?

94% of the evaluators agreed the EMAI model was capable of forming liking and disliking attitudes toward the objects in its environment compared to only 16% who thought the random model was capable of forming attitudes.

Besides being able to produce reasonable emotional states, it is important to the successful evaluation of the EMAI model that the evaluators were able to observe the emotional states affecting the agent's behaviour and learning. To this end the evaluators were asked:

Do you think that Fido learned from his emotional states?

The results were unanimous. All evaluators agreed the EMAI model displayed the ability to learn from its emotional states and the random model did not.

On completion of the exercises and the associated set of questions, the evaluators were asked to reflect on their experience with each model and to give an overview evaluation of the agents. This evaluation focused on four key areas; intelligence, learning, behaviour and emotions.

10.4.6 An Overview of the Evaluation of Fido

The overview evaluation of the Fido models was given to evaluators on the completion of the exercises. The data obtained from this questionnaire was necessary for the evaluation of effectiveness of the EMAI architecture to use emotional mechanisms for intelligence, behaviour and learning. The first item addressed was the agent's intelligence.

Intelligence

For this part of the evaluation, evaluators were given a definition of intelligence within the scope of the EMAI architecture. They were told there are many and varied definitions of intelligence, but for the purpose of this experiment, intelligence was being evaluated from a limited number of perspectives. The first perspective examined was goal orientation: the evaluators were informed that human behaviour is considered as goal orientated. They were told that this means that humans set goals and try to satisfy them and often there are several actions that can achieve the same goal. The evaluators were asked nine questions relating to the agent's intelligence. The answers for each of the questions are in Table 10.7.

From these results it can be seen that the EMAI model performed well in all aspects of goal setting intelligence and outperformed the random model. The evaluators also considered the EMAI agent to have a far better ability to formulate, and in turn use attitudes for decision making. All evaluators thought the EMAI agent had the ability to form emotions and to use them for decision-making. The random model showed the evaluators mostly doubted that the random agent could do the same.

Another evaluation on intelligence was performed on the agents' capabilities for

Question	**EMA** **Yes**	**Random** **Yes**	^{2}	**P**
	78%	11%	16.20	<0.0005
WANTs?	44%	78%	4.21	<0.05
Do you think that Fido can formulate ways in which to satisfy his goals?	83%	11%	18.84	<0.0005
Do you think that Fido tries other plans of action when one fails?	89%	33%	7.20	<0.01
Do you think that Fido shows the ability to form attitudes?	89%	7%	25.08	<0.0005
Do you think that Fido uses his attitudes to make decisions?	89%	11%	21.78	<0.0005
Do you think that Fido shows the ability to form emotions?	100%	28%	20.35	<0.0005
Do you think that Fido's emotions influence his behaviour?	100%	60%	7.89	<0.005
Another definition of intelligence is the ability to adapt to certain environmental changes. Do you think that Fido has this form of intelligence?	94%	22%	13.31	<0.0005

Table 10.7 Intelligence Overview

adaptation. Again the EMAI model performed favourably with 94% of evaluators believing the agent had the ability to use its intelligence to adapt to the environment compared with 22% for the random model could do the same.

As a final evaluation of Fido's intelligence the evaluators were asked to rate the agents' intelligence. The EMAI agent's intelligence was rated to be significantly (t=7.12, P<0.0005) more reasonable (with a mean of 6.00) than the random agent's intelligence (with a mean of 2.28).

Learning

The learning section of the questionnaire focused on the agents' ability to revise its beliefs about its environment, behaviours, objects and emotions. Learning for the Fido simulation was equated with Fido's ability to be obedient. It was assumed that if Fido appeared to be doing what the participant wanted the agent to do based on the way in which the participant had trained Fido, then the participant would consider the agent as obedient. The first question and results are shown in Table 10.8.

How would you rate Fido (with respect to his obedience)?	6.22	0.011	1.89	0.014	14.33	<0.0005

Table 10.8 Fido's Obedience Rating

The results displayed show that the EMAI agent was considered obedient whereas the random agent was considered disobedient. The evaluators were then asked to rate the agent's learning skills as they related to different aspects of the simulation. The four questions asked and the results are shown in Table 10.9.

	72%	0%	20.35	<0.0005
	95%	17%	22.05	<0.0005
Do you think Fido learns about good and bad objects?	100%	11%	28.80	<0.0005
Do you think Fido learns about his behaviours from his emotional reactions?	78%	5%	19.31	<0.0005

Table 10.9 Fido's Learning Abilities

A significant number of evaluators thought the EMAI agent possessed the ability to learn about its environment, its behaviour, objects in its environment and it emotional states. A significant number of evaluators agreed that it was not the same case with the random model.

Behaviour

In this section, the evaluators were simply asked to rate the holistic behaviour of the models in terms of reasonability and predictability. The results are given in Table 10.10 as the mean values where 1 was unreasonable and unpredictable and 7 was reasonable and predictable.

	6.22	0.012	2.28	0.014	11.35	<0.0005
	6.12	0.013	2.11	0.025	5.86	<0.0005

Table 10.10 Predictability and Reasonableness of Model

Emotions

This section of the questionnaire asked the evaluators about the emotional states and responses exhibited by the agents. The first question was:

Do you think that Fido's emotional states were reasonable?

The results obtained indicate that 94% of the evaluators thought the emotional states expressed by the EMAI model where reasonable compared with only 11% for the random model. Furthermore, the mean degree of predictability for the EMAI model was rated higher than the random model for the question shown in Table 10.11.

	5.83	0.023	2.11	0.023	5.95	<0.0005

Table 10.11 Predictability of Emotional Responses

Personality

In this section, the evaluators were asked to express their attitudes toward the personality of the characters created by each of the models. Data was collected about predictability, intelligence, emotional stability and goal orientation. The answers were recorded where 1 was unpredictable, unintelligent, emotionally unstable and erratic and 7 was predictable, intelligent, emotionally stable and goal orientated respectively. The average results of the data collected from the evaluators are shown in Table 10.12.

	6.00	0.012	2.17	0.023	7.67	<0.0005
	5.94	0.015	2.33	0.024	8.00	<0.0005
	5.83	0.017	1.94	0.018	8.03	<0.0005
	5.44	0.023	1.56	0.012	10.28	<0.0005

Table 10.12 Fido's Personality Ratings

These results show that the EMAI model was rated significantly higher than the random model for all of the four personality traits measured.

Expectation

The evaluators were asked four questions relating to how they believed the models should behave and whether or not the agents lived up to their expectations of a programmed computer character. The first three questions and the results are shown in Table 10.13. The answers were recorded with 1 being 'none of the time' and 7 being 'all of the time'. These results show the evaluators thought the EMAI model met their expectations of a computerised character more often than the random model. The results also show the evaluators thought the emotional model implemented in the EMAI model produced emotional states that seemed far more reasonable than the random model.

	5.83	0.015	3.56	0.025	4.26	<0.0005
	6.06	0.015	2.39	0.017	8.57	<0.0005
	2.16	0.019	4.61	0.029	4.22	<0.0005

Table 10.13 User Expectations of Fido

The last question asked:

Did you find that at any time during the simulation that you experienced an emotional attachment to Fido?

This question sought an indication of whether or not the computer character was successful in achieving a sense of 'suspension of disbelief' with the evaluators. 69% of evaluators found they experienced an emotional attachment toward the EMAI model against 22% for the random model. While the EMAI model was rated much more favourably in comparison to the random model, the results for the EMAI model were lower than expected.

This section concluded the overview evaluation of the agents. At this stage, evaluators were given the opportunity to freely comment on the models and to give any suggestions for improvement to the emotional and behavioural mechanisms of the agents. Several of the most common improvements suggested the EMAI model be made a little less predictable. For example,

> *He's too responsive. As soon as you say good dog he'll continue to do whatever it is. This may go on for a while. He needs to be able to do other things as well.*

and

> *I guess this Fido learned too well, perhaps a little bit more naughtiness earlier on.*

and

> *He's a bit too predictable. Fido should be a little bit more erratic to make him more exciting.*

Further comments, while constructive (and even humourous) will not be discussed at this point, but will be invaluable for improving future versions of the EMAI agent.

As the initial questionnaires were filled out in entirety for each agent at different times, the evaluators could only give an indirect comparison of the second agent with the first agent on the second questionnaire. For one final evaluation of the Fido models, on completion of the exercises with both agents, the evaluators were given a set of debriefing questions. This gave the evaluators an opportunity to directly compare both models.

10.4.7 Debriefing Questions

The debriefing questions repeated questions that were asked in the previous sections of the evaluation questionnaire. However, at the time the evaluators completed these questions, they had already had the opportunity to interact with both models. These final questions serve as the evaluators' direct comparison of the models.

Table 10.14 displays the results gathered about the first four questions. The results are given as percentages of evaluators who selected one model over the other in each of the categories shown. For questions 1, 2 and 3 all evaluators agreed the EMAI model possessed better goal setting, learning and intelligence skills and for question four, while not a unanimous vote, the EMAI agent was rated much higher than the random agent.

Which model possessed the best goal setting skills?	100%	0%
Which model learned from your feedback the most?	100%	0%
Which model seemed more intelligent?	100%	0%
Which model seemed more emotionally intelligent?	94%	6%

The final part of the evaluation asked each participant to directly compare the agent models' traits of learning skills, behaviour, emotional states and goal setting. The mean results are displayed in Table 10.15.

Trait	**EMA**	**Random**	**t**	**P**
Learning Skills	6.61	2.44	10.93	<0.0005
Behaviour	6.56	2.39	12.41	<0.0005
Emotional States	6.06	2.56	6.51	<0.0005
Goal Setting	5.78	2.00	11.48	<0.0005

The results in Table 10.15 show the EMAI model was rated as far more reasonable than the random model.

10.5 Discussion

The evaluation of the EMAI and random models was useful in determining how successful the EMAI architecture and the mechanisms designed to emulate emotions were in simulating an intelligent computer character.

The objective of this experiment was to determine if the behaviours and emotional states produced by the EMAI agent were sufficient to create a sense of *suspension of disbelief* in a human subject who was interacting with the agent. To do this an agent needs to produce results that could measure up to a *turing test* and convince the user into thinking they were interacting with a real person or animal. The Fido model that implements the

EMAI architecture is far from being able to live up to the expectations of a turing examination as the interface is very simplistic and makes it obvious to the user that they are interacting with a machine. However, the first step toward using the EMAI architecture to create realistic *artificial life* is to determine if the concepts that are integrated for the use of producing intelligence are reasonable. To this end the EMAI architecture was evaluated on the basis of reasonability for all aspects of the agent's behaviour. To calibrate the data gathered, the same evaluators were asked to also interact with a random agent. The data from the random agent acts as a comparison base to help eliminate any favourable participant responses that could have been due to the novelty and fun of the Fido software. By including a randomly behaving Fido model, the evaluators were forced to examine the behaviour of the models in detail rather than evaluating the animations and user interface. As the random model produced random emotional behaviours, there was always a chance the random Fido would act in a way the user was expecting by coincidence.

The evaluation questions focused on gathering data about four key areas of the EMAI architecture: motivation and goal setting, learning, behaviour and emotions.

10.5.1 Motivation and Goal Setting

The experiment simulated motivation using three Internal State Registers that were represented by gauges located at the top of the user interface. These gauges represented the agent's *hunger*, *fatigue* and *intrigue*. When the Internal State Registers of the EMAI model reached certain threshold values, goals were triggered in the agent that when satisfied would reduce or increase the values of the registers. For example, as the *hunger* register reduced to a threshold value approximately halfway along the gauge, the agent would become *hungry* and goals that related to eating would be triggered. When an eating goal was satisfied the value on the *hunger* gauge would change, thus indicating the agent was no longer *hungry*. In the random model, the values on the gauges did not affect the agent's goals as goals were triggered at random intervals independently of the gauges. The differences between the models were identified by the evaluators with 94% of evaluators agreeing the gauges did influence the EMAI model's goals against 44% who believed the gauges influenced the random model's behaviours. While the goals produced in the EMAI model were deliberate and the results show that the architecture favourably achieves this, due to the very nature of the random model there was always the possibility the agent would by coincidence have the appropriate goal triggered when the gauges were close to threshold values. Therefore, the 44%, of the evaluators who thought the random model was motivationally driven was expected. In fact to produce an even lower result, the random model would have to be programmed to deliberately trigger inappropriate goals at gauge threshold values.

Evaluators remarked of the EMAI model's motivational mechanisms with comments such as:

When he needs exercise, he goes digging the ground.

and

When he was extremely hungry (as shown on the motivational gauge), I fed him and the hungry urge changed.

These showed that the evaluators were aware of Fido's changing urge values and the associated behaviours that were motivated by them. In contrast, evaluators noticed that the random model's behaviour was not related to the values on the motivational gauges with comments such as:

Even though he was hungry, he still wanted to play.

and

Even as the motivational gauges changed the dog still appeared to be active and continued with repetitive actions.

The objectives of motivational mechanisms in the EMAI architecture are to act as triggers for goal setting and thus related behaviours. To gauge the agent's ability to perform this process to a standard that would be rated as reasonable to an onlooker, several other questions were asked. These questions were presented in Section 10.4.6 and are repeated with the results, for convenience in Table 10.16 below.

Do you think that Fido set's goals?	78%	11%	16.20	<0.0005
Did Fido ever perform actions that were not related to his WANTs?	44%	78%	4.21	<0.05
Do you think that Fido can formulate ways in which to satisfy his goals?	83%	11%	18.84	<0.0005
Do you think that Fido tries other plans of action when one fails?	89%	33%	7.20	<0.01

Table 10.16 Goal Setting Intelligence Ratings

The first question asked the evaluators whether they thought the models were able to set goals. The results highly favoured the EMAI model. This showed the evaluators were aware of the EMAI agent's goal processing abilities. The random actions of the other agent proved significantly less convincing. The second question focused on the motivational gauge values and asked the evaluators if they thought the models ever performed actions that were not in line with the agent's urges (as shown on the gauges). The results for the EMAI model were not outstandingly different between evaluators. Although almost half of the evaluators though the EMAI agent performed actions that were not related to its motivations, the result is not as notable as was hoped. This suggests that either the evaluators were not clearly aware of the EMAI agent's wants, the programmed behaviours associated with motivations did not meet the evaluator's expectations or the architecture does not perform as expected. As the EMAI agent was programmed to *eat* when it was *hungry*, *sleep* when it was *tired* and *play* when it was *excited*, the latter is initially rejected. The hypothesis is that this evaluation is the result of a poorly designed interface or badly chosen motivation/behaviour pairs. Another factor that may have also influenced the evaluators is that of the agent's emotional state. Although this question aimed at matching motivational mechanisms with appropriate behaviours, the evaluators may have also assessed motivation with respect to the agent's emotional state. Although a motivation was matched with an appropriate behaviour, the presence of a seemingly inappropriate emotional state may have influenced the assessment. The random model did behave as expected producing a poor result. The behaviours of the random agent were in no way associated with the levels on the motivational gauges.

The next question asked if the agents were able to devise methods that would satisfy their goals. The results suggest the evaluators were aware of the EMAI agent's goals and noticed the agent's attempts to satisfy them. The random model was rated much less favourably.

Finally, a question was asked to discover if the evaluators thought the models were capable of failure-recovery as failure-recovery is a quality of goal orientated systems. The results were favourable with a significant number of the evaluators believing the EMAI agent could recover and try other plans of action after one plan had failed in contrast with a smaller number of evaluators who thought the random model could do the same.

The lower outcome obtained for the question on relating behaviour to the EMAI agent's wants could be due to the wording used. This question is the only one that mentions *wants* rather than referring to them as *goals*. The result obtained for the third question support this view. In order for the evaluator to have noticed the agent formulating ways to achieve its goals, the evaluator would have had to witness the agent's behaviour. The outward behaviour of the agent was the only way in which an evaluator could have understood the actions the agent was attempting to satisfy its goals. Furthermore, the goals of the agent were directly related to the agent's motivational state represented on the motivational gauges. Essentially, these two questions were asking for very similar evaluations to take place. The conclusion is therefore that the evaluators did not associate *goals* and the plans to satisfy them with *wants* and related actions.

In all, the results gathered about motivation and goal setting significantly favoured the EMAI model.

10.5.2 Learning

The focus of the learning element of the EMAI architecture lies in the mechanisms that allow an agent to have valenced reactions toward elements in its environment. In the experiment, the EMAI Fido model learned through user interaction where the evaluator could praise the agent when it was performing a preferred action and scold the agent when it was not. When the agent received praise, it associated the positive emotions with the current event. When it received negative feedback from the evaluator, it associated negative emotions with the event.

In future decision making the agent used its knowledge about the emotions attached to events and the elements of the events, such as the action or any involved objects, to select an event to perform. As discussed in Chapter 9, the EMAI agent initially prioritised its choice of behaviours based on urgency (symbolised by the levels of the Internal State Registers) and then arranged the behaviours on emotional preference. The emotional preference of behaviours is calculated by determining how the behaviour would influence the agent's mood if it were performed. This process was not present in the random model. The random model did not retain any information received from user interaction nor use any knowledge of previous behaviours to make future decisions.

The evaluators were asked to assess both the EMAI and random model's outward behaviour with respect to learning by allowing the evaluators to work through several training exercises with the agents. The evaluators were asked to attempt to train the agents to prefer playing with and eating certain objects. The results obtained were encouraging.

Evaluators explained their training techniques in a similar way. The following comment from one of the evaluators is typical of most comments received. It shows the "BAD DOG" button was used to discipline the agent while the "GOOD DOG" button was used to praise the agent.

> *When Fido ate grass, pushed "Bad Dog" Only did this once. When Fido rolls on ground, pushed "Good Dog" He kept rolling on the ground.*

Initially, the evaluators were asked if they found Fido easy to train and whether they thought that Fido learned anything from the training. Most evaluators thought the EMAI model was easy to train compared to a few who thought the random model was easy to train. Again most of the evaluators thought the EMAI model learned from the training and significantly less thought the random model did. These results were expected as the EMAI agent associated the emotional state of *happy* with a press of the GOOD DOG button. When the evaluator pressed the GOOD DOG button, the EMAI agent associated a happy emotional state with its current behaviour and any associated event elements. This would bring the emotional state value of the event, from which the behaviour derived, closer to happy. In turn, when the event was next placed in the agent's schedule it would have a higher emotional priority for being selected for execution. Therefore, the more praise the EMAI agent received for a behaviour, the higher the likelihood that it would be selected again. This process illustrates affective decision making in the EMAI agent.

The EMAI architecture's mechanism for learning lies with an extension of the Theory of Reasoned Action (see Chapter 6) which provides the agent with the ability to make multi-dimensional affective decisions. This ability allows the agent to develop attitudes about the elements in its environment though interaction with them. Attitudes are based on the outcome of events and the use of objects. A favourable outcome develops positive attitudes and an unfavorable outcome develops negative attitudes. The results show that the evaluators had a sense of this process occurring in the EMAI agent and thought that it produced reasonable behaviour. The opposite is true for the random agent.

Finally, the evaluators were asked to rate the overall intelligence of each model. The mean result obtained for the EMAI model clearly shows the evaluators believed the EMAI agent produced more reasonable intelligent behaviours than the random model. This conclusion was also evident in the evaluator's comments.

Evaluators commented on the EMAI model's obedience aspect of its intelligence as:

> *He was extremely intelligent. If I told him it was good to do something then he did it and continued doing it.*

and

> *He does what you tell him is good to do and he does not do what you tell him is bad.*

Evaluators commented on the random model's intelligence as:

> *Random behaviour, unrelated to motivation or environment.*

and

> *He did not appear to remember anything I tried to teach him. He was extremely erratic, doing whatever he wanted to do.*

These results confirmed that the Theory of Reasoned Action, implemented in the EMAI agent to give the agent the ability to form attitudes, was effective as a learning mechanism.

10.5.3 Emotions

The essential element in the EMAI architecture is the inclusion of the Affective Space and the Theory of Reasoned Action that allows the agent to form emotions about its behaviours and objects with which it comes in contact. It can then use these emotions to make decisions about future events based on the emotional reactions the agent has about these events and how performing these events would influence the agent's mood.

Throughout the evaluation the evaluators were asked to answer questions about the reasonableness of the agent's emotional states and how they thought these emotions were affecting the agents' behaviour. Some initial questions asked the evaluators to rate the agents' emotional states with respect to agent's current behaviour and the evaluators' interaction. The EMAI model rated favourably on both these aspects. With respect to the agent's emotional states and current behaviours, the mean responses obtained for the EMAI model show the emotion producing mechanisms present in the EMAI architecture are working to produce a computer character that is capable of producing quite reasonable emotional states from an onlookers point of view. The random model, while not rated as totally unreasonable, was expected to produce mediocre results. Due to the nature of the randomness of the emotions produced there was always a possibility the agent, by coincidence, would meet the evaluators' expectations.

10.5.4 Behaviour

The emotional mechanism is closely tied to the outward behaviours of the agents. Evaluators had to rely on the outward behaviour of the agent to evaluate the agent's emotion forming abilities. Of course, this is always the case with computer character interaction as the user is not aware of the internal mechanisms driving the character. Although motivation, learning and emotions were evaluated separately from behaviour, the results obtained for the agent's capacities were essentially the evaluator's evaluation of the agent's behaviour. For this reason, several questions about the agent's behaviours were also asked.

Although evaluators were asked questions relating to intelligence, motivation, goals and emotions, the evaluation of these categories relied on how the evaluators viewed the outward behaviour of the agents. In the debriefing the evaluators were asked to rate the agent's behaviours on two scales, one for reasonableness and one for predictability. The evaluators thought the EMAI model produced behaviours that were significantly more reasonable and predictable than the random model. It is ideal that the EMAI model was rated highly for both reasonableness and predictability as the architecture was designed with these characteristics in mind. However, while these results seem in favour of the

EMAI model, the evaluators also needed to be asked if they thought the models met their expectations of computer generated characters. The EMAI model may be able to produce reasonable and predictable behaviours but if these characteristics are not expected in a computer generated character then the architecture may need to be modified when used for this purpose.

Evaluators were given the opportunity in the evaluation to comment freely on their expectations of the models. Typical comments received about the EMAI model included:

> *He responded well to positive and negative reinforcement, acting in a logical fashion with enough randomness of behaviour to discover and evolve new ways of interacting with his world.*

and

> *Sometimes he can be a bit predictable but overall yes. He is motivated by his goals and urges and will try different approaches if he does not get them.*

and

> *From my point of view a computerised character should be more predictable than a real one. So I am very satisfied with Fido's performance.*

Feedback received about the evaluators' expectations in relation to the random model's behaviours were quite different. Some of the comments were:

> *Having had some experience with AI and Alife before, I know they can be contrary, but you can generally sense some structure evolving. This Fido has none.*

and

> *Total unpredictable!*

and

Given the programming of Fido's character to be neurotic, then yes I could see he could meet expectations.

The previous results coupled with the comments from the evaluators are strongly in favour of the EMAI model. Reasonableness and predictability are favourable characteristics to have in a computer generated character. The EMAI model has displayed the ability to successfully fulfill these characteristics.

10.5.5 Direct Comparison of the Models

Having completed the questionnaires for both models separately, the evaluators were given an opportunity to directly compare both agents. This comparison was performed on four characteristics of the agent's learning skills, behaviour, emotional states and goal setting. Each category was rated by the evaluators on the same scale of reasonableness used throughout the evaluation. The EMAI model performed significantly better in comparison with the results obtained for the random model.

10.6 Summary

The results show that the evaluators rated the EMAI model much more favourably than the random model. It was *hope*d these types of results would be obtained as the behaviours and emotional states of the EMAI model were calculated through the mechanisms integrated into the agent's architecture for the very purpose of building a computerised character that could meet a user's expectations. The random model was used against which to benchmark the EMAI model. If the EMAI model could not meet the evaluator's expectations significantly more often than the random model then the affective mechanisms integrated into the EMAI agent could be construed as useless for this purpose.

Not only do the results of the evaluation favour the EMAI model but they also reinforce the success of the new concepts integrated into the agent's architecture. These concepts, discussed in-depth throughout this dissertation, include the Ontology, the Motivational Drive Generator, the Event Space Generator, the Intention Generator and the Affective Space. Each of these new ideas for generating intelligence using emotional theories in artificial beings received exceptional feedback throughout the evaluation. The results reinforce the concepts integrated into the agent architecture and encourage further evaluations of more sophisticated computer generated characters with more complex goal hierarchies and user interaction interfaces.

This chapter concludes the content of the dissertation and discussion of the research herein and the contributions to the field of affective computing and artificial intelligence. The following chapter summarises the content of this dissertation, details the contributions of this work to the domains of affective computing and AI and discusses the future direction of this research.

11. Conclusions and Further Directions

Does the road wind up-hill all the way? Yes to the very end. – Christina Rossetti 1830-1894

11.1 A Piece of the Puzzle

Emotions are a difficult concept to define let alone integrate into the domain of artificial intelligence. Emotions have been studied in fields such as philosophy, physiology, neurology and psychology. All have their own and often distinct ideas and models explaining how emotions are generated and affect behaviour. Investigators in the field of artificial intelligence have only begun their examinations into this domain. Researchers are examining the influence that emotion has on intelligence and investigating ways in which the phenomena can be used to enhance artificial beings.

A small group of researchers has focused their efforts on the integration of appraisal theory models of emotions (see Chapter 2) with agent technologies to produce artificial beings primarily for the entertainment industry. This includes producing believable agents for the field of interactive drama (Maes 1995) and computer games (Stern 1999) that are capable of producing performances from scripts and simulating an illusion of life implementing emotional exaggerations and personalities. Other areas where emotions are beginning to advance artificial intelligence are in the domains of intelligent tutoring systems (Kort et al. 2001), wearable computers (Healey 2000) and user interfaces (Klein et al. 1997, Picard 2000). While these systems may not require the computer to produce emotions or emotional behaviours, models of emotions have been useful in allowing the systems to reason about the emotional states of their students and thus enhancing the interaction and learning experience.

As the integration of emotional mechanisms into artificial intelligences is a research area in its infancy, little work has been performed experimenting with different emotional models. Appraisal theories have been selected for their discrete categorising nature as they are ideal for integration into many agent architectures that are capable of generating appraisals that fit the models. However, much of the work has focused on a small group of appraisal theories and in particular the OCC model (see Chapter 2). While this research has been successful in producing believable emotional agents, the model is restrictive in its categorical nature. Contemporary research in agent models (BDI (Rao and Georgeff 1995), SOAR (Tambe 1997)) does not implement emotional intelligence in decision-making, re-planning, or goal prioritisation. Some new models are emerging that do consider emotions

(PETEEI (El-Nasr 1998), Tok (Reilly 1996), The Affective Reasoner (Elliot 1992)), however, they generally consider the generation of emotional states and not the consequences on the agent's beliefs and goals. Listed here are the limitations of affective agent models that precede EMAI and how the EMAI architecture overcomes these limitations:

1. *Models that do integrate emotions require considerable reprogramming when new emotional categories are added.* The EMAI model overcomes this by replacing the traditional individual gauge method of emotional representation with a continuous Affective Space. This is demonstrated in Chapters 8 and 9.

2. *There is no feedback from the generated emotional state to enable the agent to adapt to its environment.* An EMAI agent uses emotional states to discover how events make it feel and further uses this information in decision making tasks. This is presented in Chapters 5 through 9.

3. *There is a lack of deep knowledge representation and manipulation capabilities.* The integration of conceptual graphs into the EMAI architecture's Ontology allows for a depth of knowledge representation that gives an EMAI agent the ability to construct complex representations of events at varying generalised and specialised levels and to associate emotional states with these. The EMAI's Ontology was described in Chapter 4.

4. *Contemporary research attributes emotional state to holistic events*. The EMAI architecture attributes emotional states at the atomic level of all the entities associated with an event. This was shown in Chapter 9.

5. *Emotional decay in previous models has been dealt with by using constant decay rate functions that act over time.* This is contrary to the EMAI model that uses a continuous Affective Space that allows opposing emotions to counteract each other. This was illustrated in Chapter 8.

6. *Current agent models do not address the issue of dynamic goal interpretation (via dynamic planning).* This differs from EMAI agents (like Fido) that could plan and re-plan, based on changing goals that are associated with changing emotional states. This contribution is demonstrated in Chapters 5 through 9.

7. *Current agent models lack the ability to maintain a dynamically changing experience against an event or element.* In contrast, the EMAI architecture is able to revise its experiences toward during an event or toward and element at every encounter, and its experiences are continuously revised based on the outcome of the encounters. This is also illustrated in Chapters 5 through 9.

8. *Contemporary agent models do not apply emotional bias in decision making.* EMAI agents (like Fido) use a biasing emotional state in its multi-dimensional affective decision making technique to determine its most preferred choice. The biasing mechanism is demonstrated in Chapter 9.

9. *No formal models of emotional intelligence have been described to formalise emergence of emotion in an agent model.* The EMAI model achieves this through the formalisation of event and element appraisal, emotional state generation, emotional state updating and multi-dimensional affective decision making. The formalisations appear in Chapters 5 through 9.

10. *Current agent models represent individual emotional states as separate variables thus making it a complex task to add new emotional states or combine existing ones.* The use of the continuous multi-dimensional Affective Space in the EMAI architecture presents emotional states as related points within the same space. This allows for the easy addition of new emotional states to the space as well as the blending of multiple emotional states to produce a single resulting emotional state. The Affective Space is formally introduced in Chapter 8 and further demonstrated in Chapter 9.

11. *Until now the Theory of Reasoned Action had been applied in the psychological domain and not fully formalised for computational use.* Fishbein and Ajzen (1975) formulae (discussed in Chapter 6) have been extended in this dissertation to calculate the attitudes an agent has towards *events* and *event spaces*. Furthermore, the ideas of formalisation presented in the Theory of Reasoned Action (Fishbein and Ajzen 1975) have been expanded to express the computational nature of Smith and Ellsworth's (1985) appraisal dimensions used in EMAI to generate emotions and perform affective decision making. These ideas are discussed and presented in Chapters 5 through 9.

The contributions the EMAI model makes to the domains of affective computing and AI are made possible via the new and unique mechanisms integrated into the architecture.

11.2 Emotionally Motivated Artificial Intelligence (EMAI)

The EMAI architecture consists of three functional areas; the Knowledge Area, the Constructive Area and the Deliberate Area (explained in Chapter 3). The Knowledge Area of the architecture refers to the agent's Ontology, the Motivational Drive Generator, the Emotional State and Bias Registers, the Sensory Processor and the Affective Space. Within this area resides the knowledge the agent needs in order to function. This includes the agent's knowledge structured in the Ontology (examined in Chapter 4) as the Goal Hierarchy, the Type Hierarchy, the Relation Hierarchy and a set of canonical conceptual graphs.

Discussion on the Motivational Drive Generator (examined in Chapter 5) demonstrated how the Internal State Registers were used to represent the agent's lower emotional states such as *hunger* and *tiredness*. The registers are monitored and updated by three drive mechanisms (Homeostatic, Cyclical and Default). The drive mechanisms are used to trigger associated goals from the agent's Goal Hierarchy by way of the Sensory Processor. These goals represent the *wants* of the agent and the agent strives to satisfy these goals by constructing a series of event spaces in the Constructive Area.

The Constructive Area receives abstract goals from the Knowledge Area. These goals are processed by the Event Space Generator into plans (represented by Activity Digraphs) the agent knows if successfully executed, will satisfy the goals that were triggered earlier from the Goal Hierarchy. The events that constitute the event space (explained in Chapter 6) are constructed from four elemental concept types: actions, objects, time and context. These elements are derived from the Theory of Reasoned Action (Fishbein and Ajzen 1975) and used to describe all aspects of behaviour. Once constructed, the event space is passed to the Deliberate Area for ordering, prioritising and scheduling processing.

The Deliberate Area accepts input in the form of an event space from the Constructive Area. The purpose of the Deliberate Area is to order, prioritise and schedule events. The events in the agent's schedule become the outward behaviour of the agent. The Deliberate Area creates an ordered schedule of events for the agent to perform. The schedule is dynamic and is continually updated and reordered as new event spaces arrive from the

Constructive Area. This operation is performed in the Deliberate Area by the Intention Generator.

The Intention Generator uses the agent's Affective Space to assign emotive values to the elements of an event. These values are then used to prioritise the agent's behaviours from most-liked to least-liked. Initially the process of making an affective decision based on a single appraisal dimension (*intention*) derived from the Theory of Reasoned Action, was examined in Chapter 7.

In Chapter 8, the discussion returned to the Knowledge Area's Affective Space. As this concept forms the heart and soul of the EMAI architecture, the methods implemented for implementing the Affective Space were formalised. The multi-dimensional nature of the Affective Space was also examined along with its inherent nature to deal simply with complex issues such as emotion decay and blending.

In Chapter 9, the process of single-dimension affective decision making was expanded into the realm of multi-dimensional decision making using the Affective Space. Formalisations for generating emotional states, updating emotional states and predicting future emotional states were given.

As the EMAI architecture was designed to evaluate the effectiveness of using concepts of emotion and emotional evaluation of actions, objects and other elements to generate intelligence in an artificial agent, the capabilities of these mechanisms to produce reasonable and believable intelligence in a computer-generated character was evaluated in Chapter 10.

The evaluation of the EMAI architecture determined if the model was sufficiently capable of using the set of highly integrated mechanisms for generating motivation, goal setting, emotional intelligence and event prioritisation and scheduling. The results gave positive feedback about the EMAI architecture's ability to produce reasonable emotional states and associated behaviours. The data also confirmed the agent's ability to set and execute goals using motivational mechanisms related to the agent's physical and mental states.

The emotional concepts implemented in the EMAI architecture are twofold. The functioning of the Motivational Drive Generator represents the lower level or survival emotions (discussed in Chapter 2) and the higher order emotions are represented and formulated using the Affective Space. The Internal State Registers in association with the Homeostatic, Cyclical and Default Drive Mechanisms and the Sensory Processor trigger goals in the agent's Goal Hierarchy. Following this, an EMAI agent will endeavour to satisfy these goals by devising a list of appropriate events. The agent selects events to execute by prioritising them in order of the emotional states the event would evoke if executed.

This functionality has eliminated the need to fix production rules in the agent's knowledge base. Instead, the agent's choice of behaviour is influenced by its motivational and emotional states. The results obtained from the evaluation experiment show this type of behaviour is reasonable in an artificial agent, in particular one used for creating computer characters.

11.3 Research Contributions

The objective behind this dissertation has been to develop a new architecture for an emotionally motivated artificial intelligence. The approach taken has been to move away from the traditional individual variable approach (see Section 1.4) that limits the way in which emotions are represented and processed. The EMAI architecture presented in this dissertation makes a small but significant advance away from this single-dimensional methodology with an innovative mechanism used in agent construction, called the Affective Space. The Affective Space continues to uses appraisal theories, but eliminates fixed formatted production rules and creates a dynamic multi-dimensional emotional state space used for the generation and processing of emotions. The research presented here has led to a number of contributions to the field of Artificial Intelligence and Affective Computing.

Contribution 1: The first contribution has been to develop an agent with the ability to appraise behavioural choices based on emotional assessment. Given a specific goal to achieve and a number of suitable action options from which to choose, the agent can select an option that feels the best. Appraisal theories explain how a person could feel after an event. An EMAI agent determines how an event is considered, emotionally before the behaviour induced by the event takes place. This is generated from experience and beliefs obtained through the agent's interaction with its environment, other agents (including humans) and its internal states.

Contribution 2: Secondly, an EMAI agent has the ability to revise its beliefs based on experience using the emotional states to assess elements of an event and use these evaluations in the assessment of future behaviours. The EMAI architecture extends the single-dimensional appraisal of intention from the Theory of Reasoned Action (Fishbein and Ajzen 1975) to generate a multi-dimensional affective appraisal model that evaluates responses toward events. The values gathered are then used to extrapolate future emotional states.

Contribution 3: Behaviour selection in an EMAI agent is primarily based on an emotional state. The agent, based on its current emotional state and emotional bias, can generate a number of behavioural choices stemming from motivational mechanisms. These emotional states are used by the agent in a debriefing of its choice, the integration of new knowledge into the belief system and its use in the future assessment of events.

Contribution 4: Another concept the EMAI agent introduces is that of a continuous multi-dimensional emotional space. Previous models of emotional agents use a method that employs a number of levels or meters that represent discrete emotions. These levels can rise or fall depending on the agent's state. The EMAI model engages a continuous Affective Space that consists of (initially) 15 emotional states, but is capable of representing infinite emotional states. This suggests the agent can never experience *no* emotion and can support complex or mixed emotional states. Furthermore, mixed emotional states that are generated cannot be formed by an unlikely combination of emotional states due to their positioning in the Affective Space.

Contribution 5: Another inherent characteristic of the continuous emotional state is the natural emotional decay that occurs. As the agent receives input that changes its emotional state and it moves toward another emotional state it will also be moving away from current emotional states. For example, if the agent is *angry*, as it receives *happier* input it will become more *happy* and less *angry*. Emotional decay in previous models has been implemented by using constant decay rate functions that act over time. These require considerable maintenance and coordination. In the EMAI architecture, decay of an emotional state occurs as the result of the mood, represented by a point in the Affective Space, changing. As the appraisal dimensions stored in the Knowledge Area that represent the agent's mood are updated, the location of the mood point in the Affective Space changes. Therefore, as the mood point moves away from one emotional state point towards another, it naturally causes the strength of the previous emotional state to decay.

Contribution 6: A further value that affects the operation of the Affective Space and adds extra functionality to the model that has not been seen in previous models of emotional agents is that of mood or current emotional state. While other models do consider the issue of a mood or current emotional state they do not use this for emotional biasing in decision making. Depending on the agent's mood it can select from tasks that will alter its mood to a more desired state. What this means for the agent is that it will not always select the same task to put it in the desired mood. For example if the agent's preferred mood is *happy* it may select a different task to alter it's mood from *sad* to *happy* than it might from *angry* to *happy*.

Contribution 7: An EMAI agent also possesses a deep knowledge representation through its use of conceptual graphs and its interrelated knowledge of goals and known concepts. Using this knowledge, the agent can extrapolate emotional states about new situations that are related through its Ontology. Although an agent may not have had an emotional experience involving a particular concept, it can estimate its emotional response using related concepts. For example, the agent may have experienced an *angry* incident with a particular teacher. If the agent had to calculate its emotional response toward another teacher, it could use the past *angry* response to extrapolate an emotion. Introduced in the EMAI's knowledge representation were several new abstractions that were used to enhance the use of conceptual graphs for the purpose of constructing the EMAI architecture. These included the Goal Hierarchy, Event Graphs and Activity Digraphs.

Contribution 8: The way in which goals and concepts are represented in the agent's knowledge base also allows it to attribute emotional states to the concepts/elements that make up an event. Contemporary research attributes emotional state to a holistic view of events rather than assigning emotional reactions to the atomic components of an event, namely the action, object, time and context. The way in which an EMAI agent assigns these emotions to the elements is further used to extrapolate and predict the emotional states of events that the agent has not encountered before but contains elements from other events that it has experienced.

Contribution 9: The manner in which goals are structured in the EMAI architecture's Ontology also allows an EMAI agent to address the issue of dynamic goal interpretation via dynamic planning. EMAI agents can plan and re-plan the activities for satisfying their goals based on changing goals, environmental states and element appraisals that are associated with changing emotional states. As the emotional state of the agent changes, so does its goals. Furthermore, as the agent gathers experiences with elements in its environment and makes emotional assessments of these items, the agent's plans for satisfying its goals also change.

Contribution 10: In the same way that an EMAI agent can maintain a set of dynamic goals, an EMAI agent also maintains a dynamically changing experience against the elements of events that it has experienced. An EMAI agent is able to revise its experience toward the event elements at every encounter and can revise these assessments on a continuing basis determined by the outcome of the last encounter.

Contribution 11: Another restriction of previous emotional agent architectures that the EMAI architecture has overcome is the need to calculate emotional states through a series of constant production rules. For example, in the PETEEI pet-dog model, one rule is *if the agent is angry and bowl is taken away then bark.* In the PETEEI model the event, *bowl is taken away* followed by a *bark* implies the agent is *angry* or in a bad mood. Firstly, there is no mechanism to make this event a positive event resulting in positive emotions, it is programmed into the agent to have a negative effect on the agent. An EMAI agent allows its appraisal of events to change over time and with experiencing episodes of the event. Secondly, if the event *bowl is taken away* was to become a good event the production rule would no longer make sense. An EMAI agent does not use production rules, but instead maps the event's appraisals into the Affective Space in order to determine how it will affect the current state of the agent. In this way the agent's emotions are dynamic and it can use these changing affective appraisals to adapt to its environment.

Contribution 12: Another contribution made in this dissertation is the extension of the formalisation of the Theory of Reasoned Action. As the theory was developed for use in the psychological domain, the expressions that explain the relationship between beliefs, assessments and intention were too vague for implementation into a computation agent. Therefore, the expressions used in the theory were expanded and modified to best suit the EMAI architecture and used for single-dimensional affective decision making. Furthermore, the formalisation used in this dissertation for the Theory of Reasoned Action was further augmented and used to describe the nature of the six appraisal dimensions of Smith and Ellsworth's (1985) appraisal model and used to develop the EMAI agent's multi-dimensional affective decision making qualities.

Contribution 13: Finally, this dissertation contains the first attempt at a formalisation of a model of emotional intelligence that has been developed to describe the emergence of emotion in an artificial agent model. This model, as described in Chapters 8 and 9, was evaluated by a number of volunteers interacting with an experimental computerised character built using the EMAI architecture. While conclusive evidence does not exist to support this formalisation as being the true and correct way to implement an emotional decision making mechanism in an artificial agent, the preliminary results were supportive and encouraging.

11.4 Future Directions

During the course of the research presented in this dissertation, many issues and ideas for further work have arisen. This section examines three directions the research in this dissertation could take to improve on the EMAI architecture and concludes by discussing some further application areas.

11.4.1 Deeper Knowledge Representation of Event Element Relationships and Attitudes

As discussed in Chapter 5, values for attitude calculated using the Theory of Reasoned Action mean very little in isolation. For example, the attitude value for *buying a new red Honda sports car* may be 6, but when compared to nothing it is meaningless. Therefore, attitudes are calculated for related events such as *buying a new blue Ford ute* or *buying a secondhand white Mercedes Benz* and compared with the former event to put it in perspective.

However, the aforementioned events and their associated attitudes are calculated for the same action, in this example, *buying a car*. The Theory of Reasoned Action does not describe how the attitudes toward individual event elements, pertaining to one action, might relate to the same elements involved in a different action. For example, is the attitude towards the *red Honda sports car* in the event for *buying a car* the same, as it would be in the event for *driving a car*?

While the EMAI architecture allows for weightings to be applied to individual elements that make up an event it does not allow for the depth of relationships that exist between the individual elements. In Chapter 7 the Theory of Reasoned Action was used to calculate the attitude toward performing an event. In brief, the agent's attitude toward performing an event is determined by the agent's set of beliefs about the event elements and how the agent assesses each of these beliefs. Although the attitude calculation equation is used in the Theory of Reasoned Action to holistically assess an intention, according to Petty and

Cacioppo (Petty and Cacioppo 1996), it can also be used to assess attitudes towards elements involved in events such as people, objects and issues. However, it does not describe how the attitudes toward individual elements from one event might be combined with elements from other events to calculated new attitudes.

The EMAI architecture extends the Theory of Reasoned Action and applies it to finding assessment values for six appraisal dimensions. These values are used to determine the agent's current emotional state. While this has been evaluated as adequate for the task of representing the goals of a pet dog and combining event space elements to produce reasonable emotional states, it is unclear how more complex event spaces and associated elements would be handled.

This possible shortcoming of the EMAI architecture and limitation of the Theory of Reasoned Action introduces an objective for further research. It seems that in an event, the action and object of that action have the most influence over the attitude toward the event and thus the associated emotional state evoked by the event. Consider the pet dog example. Assume the pet dog has two objects in its environment: BALL and TOWEL, and the agent also has two actions to choose from, namely, CHEW and PLAY. These elements can be combined to create four event spaces, thus:

E_1 = {[CHEW],[BALL]}
E_2 = {[CHEW],[TOWEL]}
E_3 = {[PLAY],[BALL]}
E_4 = {[PLAY],[TOWEL]}

Assuming the agent, through experience, believes the action CHEW is bad and thus evokes a negative emotional state and the action PLAY is good and results in a positive emotional state, the agent will select the action PLAY over the action CHEW. If the agent believes the BALL evokes positive emotions and the TOWEL evokes negative emotions, the agent will choose, in this example, the event E_3 that represents *playing with the ball*. In this case, the event's action is disassociated from the object. Each object is combined with each action to produce an event space.

Now consider the actions for another agent; KISS and PLAY TENNIS and the objects BILL and TED. Combined these elements create the following four event spaces.

E_1 = {[KISS],[BILL]}
E_2 = {[KISS],[TED]}
E_3 = {[PLAY TENNIS],[BILL]}

E_4 = {[PLAY TENNIS],[TED]}

In this case if the agent believed the event E_1 that represents *kissing Bill* evoked positive emotions and the event E_4 (*playing tennis with Ted*) also evoked positive emotions, it would calculate that both the events, E_2 and E_3 (*kissing Ted* and *playing tennis with Bill*) would also evoke positive emotions. However, this may not be the case, Ted may be a terrible kisser and Bill an awful tennis player. Therefore, attitude associations that might work in the EMAI architecture for the goals and events of a simple pet dog would not work in the same way for more complex and disassociated events.

Here lies the need for a more complex event space definition within the agent's Ontology. Restrictions need to be placed on event element combinations. This could be achieved by creating a deeper knowledge representation of an event using the existing conceptual graph concepts. Restrictions could be placed on events that allow and disallow particular combinations of elements. For example, a restriction on which object (in this case, agent) can be combined with the action KISS may be that the object should be a good kisser. If the object does not meet the necessary criteria then the element combination is not allowed.

This type of knowledge representation is possible within the current constructs of the EMAI architecture but as yet has not been thoroughly explored. It may also be the case that these restrictions on event element combinations are still not adequate for producing reasonable emotional states in agents with complex goals, beliefs and events. An investigation would need to develop a formal relationship between actions, objects, time and context in event spaces. This research path needs to be explored further to determine if the current EMAI architecture is capable of achieving these reasonable emotional states for complex event spaces or if further enhancements need to be made.

Within the issue of deeper knowledge representation also lies the issue of scalability. While the objective of this dissertation was to discover the plausibility of modeling emotion in AI, scalability was not addressed. However, it is an important issue to consider and will be briefly discussed here.

As the knowledge base grows and the type and relation hierarchies become more complex, the topics of load handling and fault tolerance arise. Although combinatorial explosion is the main issue of scalability with respect to the knowledge base and the generation of events and plans, it allows the agent to fully explore its environment. As suggested in the Case Study 4 in section 7.8, Albert's decision making process demonstrates that humans do not need to consider all of the possibilities in their environment to make a decision. However, when an initial set of possibilities is considered unusable, another set is contemplated. As the EMAI architecture handles combinatorial explosion in a similar manner, the issue of scalability has not arisen. However, as scalability of small scale AI systems is a concern, further experimentation should be performed to discover the limitations of the EMAI architecture.

11.4.2 Domain Specific Event Appraisals and Emotional Theories

Another area that is worthy of research is that of the appraisal dimensions selected for constructing the Affective Space. The dimensions chosen for EMAI's initial Affective Space belonged to an existing appraisal model by Smith and Ellsworth (Smith and Ellsworth 1985). As the objective of the research herein was not to evaluate the appraisal dimensions but to evaluate the use of such a model in an intelligent agent architecture, different appraisal dimensions were not evaluated. It has become evident during this research that while the appraisal dimensions used for EMAI's Affective Space have been successful in representing emotional states with respect to events, there are many other dimensions that could be equally as effective. Furthermore, stimuli other than events can evoke emotional states and in these cases, an entirely different set of appraisal dimensions may be more appropriate.

The model produced by Smith and Ellsworth (1985) provides an excellent starting point for the creation of the Affective Space as they had developed the model by assessing a number of appraisal dimensions and identifying the discrete location of emotions along these dimensions through non-invasive human experimentation. The Smith and Ellsworth model demonstrates a strong relationship between a subject's interpretation of an event and the emotional reaction to it. They identified six orthogonal appraisal dimensions that show how their selected 15 emotional states are unique.

Besides the six appraisal dimensions that were included in the Smith and Ellsworth model used for the definition of EMAI's Affective Space (*pleasantness, effort, certainty, attention, responsibility and control*) psychologists have suggested many others. These include legitimacy (Roseman, Jose et al. 1990), compatibility, novelty and perceived obstacle (Scherer 1982) and danger or threat (Smith and Lazarus 1990).

Because the Smith and Ellsworth model was developed through experimental procedures and evaluated as a successful theory for relating their chosen appraisal dimensions with 15 separate emotional states, it formed an excellent basis for the Affective Space of the EMAI architecture. As the objective of this dissertation was to develop an emotionally intelligence artificial intelligence, and not expand the psychological theory of emotion, further research into other appropriate appraisal dimensions and other emotional states was not explored. This is not to suggest that this type of research would not be beneficial to the advancement of affective computing. In fact, the appraisals in the EMAI architecture are restricted to the emotional states evoked from events. However, emotional states can be evoked from a number of other cognitive appraisals.

Music, for example, can evoke a variety of emotional states in the listener (Juslin and Madison 1999). Tempo, timing, articulation, key, tone and instrument choice all influence the emotions experienced by the listener.

Colour, is another concept that can elicit an emotional response from the viewer (Jacobson and Bender 1996). Emotions experienced as the result of viewing a specific colour differ culturally. For example, black is associated with funerals and mourning in the western world and white and purple are associated with funerals and mourning in the eastern world. However, the research clearly shows that differing colours can evoke different emotional states.

This suggests that emotional experiences are not restricted to the cognitive appraisals of Smith and Ellsworth's model. Exactly what appraisal dimensions are appropriate in assessing concepts other than events is ill-defined in current literature. Initially a full analyses of the domain where the affective application is developed needs to be undertaken before further identification of more appropriate appraisal dimensions is made.

It could be the case that in some domains cognitive appraisal models of emotion theory are inadequate for the full realisation of a reasonable emotionally intelligent artificial being. For example, besides the theories of emotion presented in Chapter 2 that give the western world philosophy on the subject, there are also a number of eastern views on emotion. The traditional Chinese medicine of acupuncture defined 12 main Qi channels that run through the body (Rogers 2000). These channels related to major organ and intestinal parts of the body such as the stomach, lungs and large intestine. The lungs are said to relate to the emotion of *sadness*, the stomach corresponds to pensiveness and the heart *happiness*.

In the Chinese martial art of Tai Chi, emotions are related to the universal five elements; wood, fire, earth, water and metal. Each element corresponds to both a negative and positive emotional state. For wood, the positive emotion is *humour* and the negative emotion is *anger*.

Exactly how these views on emotional theories could be integrated into AI is unclear but also worth further investigation. It may be that some western philosophies are suited to a particular set of application areas and eastern philosophies to others.

11.4.3 Computerised Character Building

Having completed an evaluation of the EMAI architecture in the form of a simple pet dog and a comparison of its performance against a randomly programmed model, an issue has arisen that questions the depth of architecture needed to create a suspension of disbelief. During the evaluation, the random modelled pet dog performed in a manner that was acceptable to the evaluator. While the EMAI model did outperform the random model, the depth of the EMAI architecture may have implemented more mechanisms than necessary in order to *fool* the user.

Stern in (Hirsh 1998) identifies some key characteristics for the successful creation of a computerised character. He suggests they need to

- act like humans – to use natural language, reason and discuss events;
- have knowledge about their world and memories of personal history; and
- exhibit emotions and personalities that fuel their motivations and desires.

Stern also explains one important key point in the development of a life-like computerised character. All they need to do is make the user *believe* they are real. The characters do not have to be internally *alive*, just perceived as such. This approach is called *user-perception-based* development.

The EMAI architecture has not been designed in an attempt to create an agent that is operating as if programmed to be *alive*. The concept of the Affective Space has been based upon an existing psychological model of emotion in an attempt to create a more efficient, adaptable and believable agent. The experiment performed in Chapter 10, illustrates the successfulness with which the EMAI architecture was used to model a pet dog called Fido.

The most complex part of creating a computer character with the EMAI architecture is in the programming of the Ontology. This is a labourious task of gathering domain specific information, goals and associated plans to fill the agent's knowledge base. As a further research project it would be critical to define a generic Ontology that could have *plug in* type modules that would be of use to application developers who were *interest*ed in integrating the EMAI agent's into their software.

This research could lead to the development of a computer character generation engine. Similar to a games engine, this would allow novice programmers and applications developers easy access to building computerised gaming characters with the EMAI architecture without having a depth of knowledge that spans conceptual graph theory, cognitive appraisal theories of emotion and goal setting and planning. Rather an ideal front-end could be written that makes the internal workings of an EMAI agent invisible to the developer. This way creative designers could put more energy into the development of the personality and background of the character rather than be concerned with artificial intelligence concepts.

This type of application would require extensive research into the psychology of human-computer interaction, what the user believes they are perceiving when interacting with a computer generated character and the tools needed by games developers and interactive fiction writers.

11.4.4 Other Application Areas

A number of application areas for affective computing have been developed in recent years. The majority of these areas focus on the computer's ability to predict the emotional state of the user. Several projects undertaken by the MIT Media Laboratory focus on wearable computers and affective tutoring systems that rely on their ability to determine the user's emotional state from a series of physiological changes. The research into the EMAI agent has taken a different perspective. Rather than predicting the emotions of the user, the computer itself is generating emotions that affect its own behaviour. This type of behaviour would be beneficial in the creation of an intelligent tutoring system that not only predicts the user's emotional state but also could adapt to it and provide an emotional and personalised interface that could appear sympathetic to the user.

Another obvious application area for the EMAI architecture is in the development of computer game and interactive fiction characters. This was discussed in Section 11.4.3.

Over the years computer games players have seen a significant advancement in the quality of the sound, visual effects and virtual reality environments that have been created by computer games programmers. Stern, a leading computer games developer, in (Stern 1999) comments that, what computer gamers are really seeking is affective AI and in particular the creation of believable artificial life or A-life. He comments that by combining a direct interactive interface with highly expressive and personality rich characters the users are easily able to suspend their disbelief and form emotional relationships with the characters. The EMAI architecture achieves a partial step in the direction to achieving this in computer characters by modelling an emotion rich decision-making mechanism that gives a computer character a perceived personality and the ability to use emotions to adapt to its environment.

E-commerce systems are increasingly recognising the importance of giving additional value to users by providing personalised transactional experiences (Meech and Marsh 2000). Two specific factors that are apparent in the impact of personalised transactions in e-commerce are those of *trust* and *personality* and how these may be integrated into web based interfaces. Shoppers in face to face transactions with another human being usually encounter these traits. The principle behind a life-like intelligent agent to act as a host at a website is the same as for the creation of computer games characters. The aim is to create a suspension of disbelief for the user. The qualities necessary for this are personality, emotions, social relationships and the illusion of life (Mateas 1997). With the ability to create an artificial personality with emotions, the EMAI architecture could be used to develop a personalised face for an e-commerce business by providing a *face* for the company and a personal shopping guide.

A further application of the EMAI architecture that couples with the e-commerce interface is a field called *persuasive computing*. This field deals with the use of computers in marketing and advertising. It involves the use of computer systems intentionally designed to change a person's attitudes or behaviour in a predetermined way (Fogg 1999). The Theory of Reasoned Action (Fishbein and Ajzen 1975) has been used to study and calculate human attitudes regarding particular subjects and topics. This theory has further been used in the analysis of persuasion techniques through advertising and marketing (Petty and Cacioppo 1996). As the EMAI architecture has been built on the concepts of the Theory of Reasoned Action it makes it an ideal candidate for use in this field of application and research.

As can be seen from the aforementioned application areas, the EMAI architecture is ideal for integration into applications that require an intelligent system and an interface with personality and emotional capabilities. The deep knowledge representation of the agent's Ontology allows it to be programmed to meet the needs of any domain. Its emotional abilities allow it to create a personalised and non-threatening interface for computer systems and web interfaces and its attitude development mechanisms allow it to not only to meet the current needs of the user but to also be able to perceive future needs.

11.5 Conclusion

This dissertation has examined the concept of emotions from many viewpoints including the philosophical, the neurological and the psychological. Of all the opposing and complementary theories of emotion one point is always made; *emotions influence behaviour*. Although emotions were originally thought to impede rational thinking, new schools of thought have arisen that consider emotions a vital part of rational intelligent behaviour and necessary not only for survival, but also intellectual success. In recent years emphasis has been placed on the concept of emotional intelligence. This has lead many researchers in the area of artificial intelligence to examine the contribution that theories on emotion could have in the development of intelligent agents. Thus the domain of affective computing was borne.

The contributions in this dissertation endeavour to further enhance the domain of affective computing by providing a different perspective on the way in which emotions are modelled in an artificial agent. The continuous Affective Space in the EMAI agent is used to replace the static gauges that are used to represent emotions in EMAI architecture's predecessors. Initial evaluations of the EMAI architecture's ability to simulate a simple computer character have been promising and provided valuable and positive feedback. This data will be useful in determining future enhancements and further experimentation.

What the future holds for the field of affective computing is unclear. As it is very much in its infancy, researchers need to continue to examine and assess the elementary concepts of emotion generation. No one theory stands out from the rest as being the ideal. The complexities of human emotions may be too extreme to include them holistically within an artificial intelligence at this time. Only those segments of emotional behaviour that are advantageous to the goals of an artificial being should be considered. Further categorisation of emotional behaviours is necessary to identify domains where particular aspects of emotions are of advantage.

As more and more effort is put into small projects that contribute to the understanding of emotions, the field of affective computing comes one step closer to a model for general, emotional, artificial intelligence.

Appendix A

The OCC Cognitive Structure of Emotions

A number of theories exist that have been developed in order to specify how many and which appraisal criteria are minimally needed for differentiating emotional states. The Ortony, Clore and Collins (OCC) model presents 22 emotional states (*happy-for, gloating, resentment, pity, joy, distress, pride, shame, admiration, reproach, love, hate, hope, fear, satisfaction, fears-confirmed, relief, disappointment, gratification, gratitude, remorse and anger)* with respect to valenced reactions to events, agents and objects. The model categorises emotions based on the consequences of events, the actions of agents and aspects of objects. Within these categories, the model further distinguishes emotional states by determining whom the event affected or who performed the action. Figure C.1, taken from (Ortony et al. 1988), illustrates the categorisation of the 22 emotions.

Although the OCC model was never intended to be used for generating synthetic emotional states in artificial beings the authors did believe that it was important that machines be able to reason about emotions with respect to natural language processing, distributed problem solving and planning. Despite the authors' original intentions for the model during the 1990s, it became the default model of choice for synthesising emotions in artificial intelligences. Partial implementations can be seen in the architectures of *Tok* (Bates 1992) and *The Affective Reasoner* (Elliot 1992).

Furthermore, the OCC model only addresses the cognitive generation of emotions. It does not consider learning from emotional states, blending of emotional states or emotional reasoning or decision making. For this reason, artificial intelligence researchers are using this model as a starting point for building affective computers.

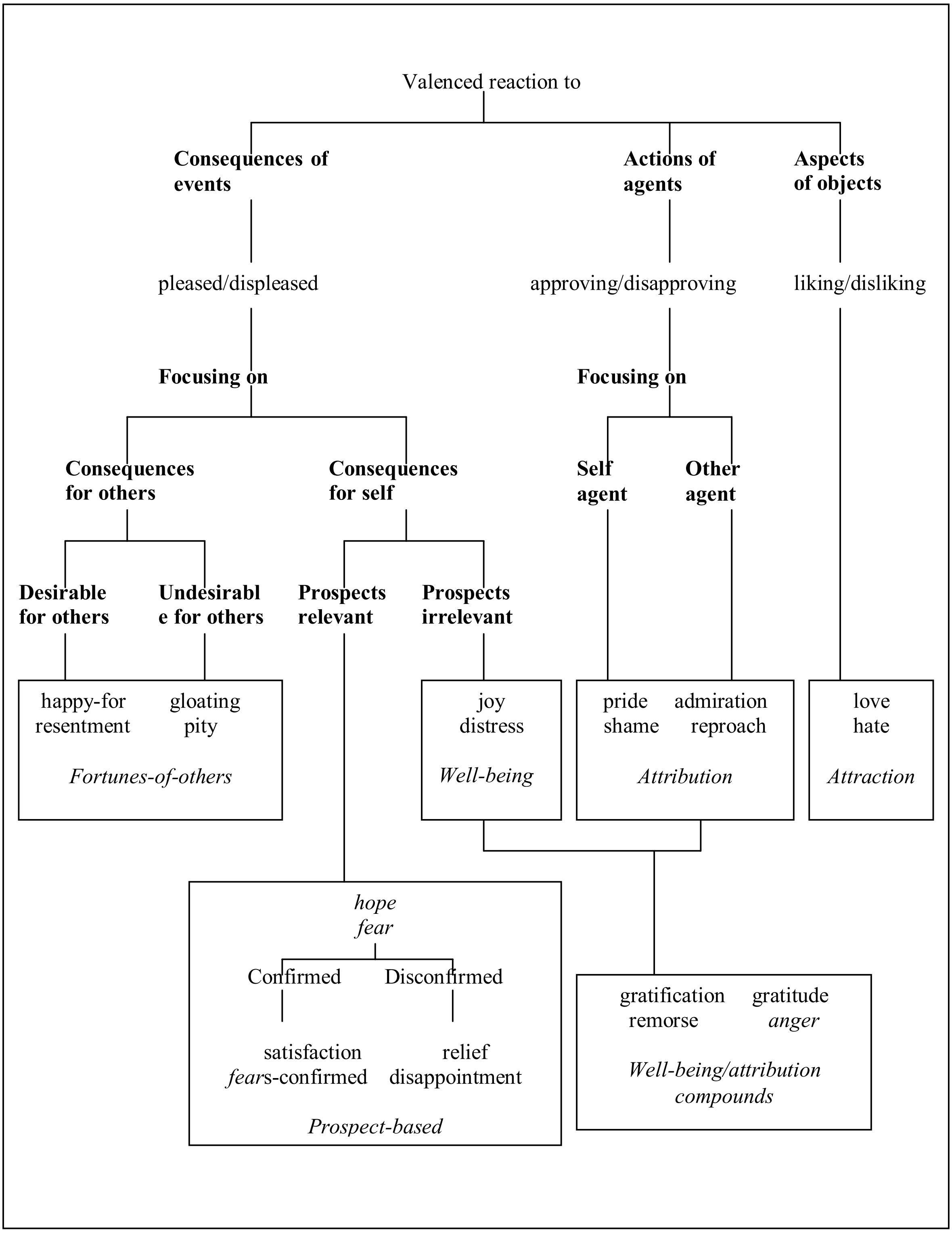


Figure A.1 The OCC Cognitive Structure of Emotions (Ortony, et al. 1988)

Appendix B

Emotional State Categorisation in The Affective Reasoner Architecture

Group	Specification	Name and Emotion Type
Well-Being	appraisal of a situation as an event	***joy***: pleased about an event **distress**: displeased about an event
Fortune-of-Others	presumed value of a situation as an event affecting another	***happy*-for**: pleased about an event desirable for another **gloating**: pleased about an event undesirable for another **resentment**: displeased about an event desirable for another **jealousy**: resentment over a desired mutually exclusive goal **envy**: resentment over a desired non-exclusive goal **sorry-for**: displeased about an event undesirable for another
Prospect-based	appraisal of a situation as a prospective event	***hope***: pleased about a prospective desirable event ***fear***: displeased about a prospective undesirable event
Confirmation	appraisal of a situation as confirming or disconfirming an expectation	**satisfaction**: pleased about a confirmed desirable event **relief**: pleased about a disconfirmed undesirable event ***fear*s-confirmed**: displeased about a confirmed undesirable event **disappointmen**t: displeased about a disconfirmed desirable event
Attribution	appraisal of a situation as an accountable act of some agent	***pride***: approving of one's own act admiration: approving of another's act *shame*: disapproving of one's own act reproach: disapproving of another's act
Attraction	appraisal of a situation as containing an attractive or unattractive object	**liking**: finding an object appealing **disliking**: finding an object unappealing
Well-being/Attribution	compound emotions	**gratitude**: admiration + *joy* ***anger***: reproach + distress **gratification**: *pride* + *joy* **remorse**: *shame* + distress
Attraction/Attribution	compound emotion extensions	**love**: admiration + liking **hate**: reproach + disliking

Table B.1 Emotional State Categorisation in The Affective Reasoner

Appendix C

Conceptual Graph Theory

C.1) Introduction

Conceptual Graphs (CG) are based on the existential graphs of Charles Sanders Peirce (Sowa 2001) and the semantic networks of artificial intelligence. They are formally defined by an abstract syntax with the purpose of expressing meaning in a model that is logical, human readable and computationally trackable (Sowa 1984). A CG has no meaning in isolation. It is part of a larger knowledge representation called a semantic network. A semantic network is a collection of all the relationships that concepts have with respect to one another, to procedures and motor mechanisms.

According to Sowa (Sowa 1984), a conceptual graph is created through the process of perception. Perception occurs in an individual when sensory input is received. This could be in the form of a visual, audible or other sensory stimulus. Through experience, a person builds up a stock of what are called *percepts* in long-term memory. Each of these percepts are a person's mental representation of previously received sensory inputs. For example, a person may have mental representations of objects such as *dog*, *cat* and *apple* and actions such as *sit*, *eat* or *sleep*. The conceptual graph itself resembles a completed jigsaw puzzle that has been constructed by joining smaller pieces together, each piece representing a single percept. On receiving sensory input, long-term memory is searched for percepts matching all or part of the input. These percepts are then assembled into a working model that best describes the input. When a person sees a cat sitting on a mat, percepts are gathered for *cat*, *sit* and *mat*. These percepts are then combined to form a complete image of the sensory input. A conceptual graph represents the logical associations and organisations of the percepts within the model created by the sensory input.

The process of perception is an individualistic process. For example, if two people were to hear that *the cat sat on the mat*, each person might develop a different image in their mind's eye based on their individual stock of percepts. One person may imagine a white cat sitting on a deep plush pile mat and the other person may imagine a black cat sitting on a doormat. The image in the mind's eye is analogous with the completed jigsaw puzzle where pieces representing the percepts cat, mat and sit have been connected and a conceptual graph could be created to diagramatically represent this image. What the image means to the individual cannot be expressed as a conceptual graph. Conceptual graphs must derive meaning in the semantic network through their links to context, linguistic rules, procedures, definitions, episodes and emotions. However, as this appendix is dedicated to the syntax of conceptual graph theory, the numerous links to the semantic network will not be discussed here.

A *conceptual graph* is constructed using a series of *concept nodes*, that represent the percepts, and *conceptual relation nodes* that specify the role that each percept plays in connecting the concept nodes.

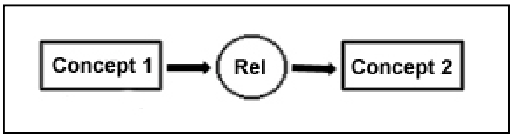


Figure C.1 A Simple Conceptual Graph Form

Conceptual Graphs have the general form as shown in Figure C.1. This can be written using a linear notation in the form,

$$[\mathrm{CONCEPT}_1] \texttt{ -> } (\mathrm{REL}) \texttt{ -> } [\mathrm{CONCEPT}_2]$$

This may be read as *a relation of a* $concept_1$ *is a* $concept_2$. The arrows in the conceptual graph indicate the direction of the reading. A reversal of the arrows in Figure 1 would result in the reading *a relation of a* $concept_2$ *is a* $concept_1$. The following is an example of a simple conceptual graph:

```
[CAT] -> (PART)->[TAIL]
```

This graph reads *a part of a cat is a tail.* Expressed like this the description sounds ungrammatical and in reading more complex conceptual graphs in this manner, it is. However, this is a useful aid in interpreting any conceptual graph. Once the user is familiar with conceptual graph structures, they could express the above CG, as *a tail is a part of a cat*.

Concept nodes represent elements such as entities, attributes, states and events while conceptual relation nodes connect the concepts. A concept has two parts, a *concept type* and a *referent*. The concept type is a label that associates the concept with a real world perception. Examples of a concept type are *entity, action,* and *time.* A concept type does not refer to any specific instance of an entity, attribute, state or event but instead acts as a predicate for which the semantic network to which the CG belongs, has a definition. It is the purpose of the referent to identify a particular instance of the concept type. The concept type *t* and referent *r* are displayed in linear form between square parenthesis with the concept type on the left and the referent on the right, thus:

`[t: r]`

Figure 2 integrates a number of concept types for example `[ENTITY].` In this instance the referent has been left blank to indicate an existential quantifier meaning the actual referent is a physical entity that exists somewhere in the semantic network. To define a particular instance of an exact entity type a referent could be included thus:

`[ENTITY: John]`

This would mean that `John` exists as an instance of an entity in the semantic network.

While a CG may be used to represent a general description of an event, it can be linked to episodes indicating when such an event took place and also emotional associations during that episode that indirectly confer emotional overtones on the types of concepts involved. As this dissertation concentrates primarily on the emotional assessment of events the following example focuses on the definition of an event using CG notation. In keeping with Ajzen and Fishbein's definition of a behavioural event (Petty and Cacioppo 1996) (see Chapter 6), an event can be represented as a CG as shown in Figure C.2.

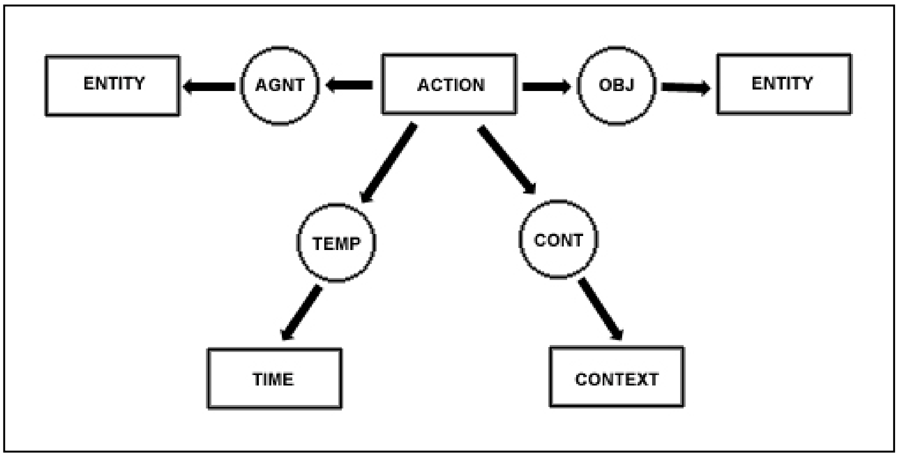


Figure C.2 A Conceptual Graph Representation of an Event

Graphical or character-based notation can be used to depict a conceptual graph that represents knowledge with its series of concept nodes and conceptual relation nodes (Sowa 1984, 2001). In what is called the *display form* of Figure C.2, rectangles represent concepts and conceptual relations are drawn as circles or ovals. The way in which the conceptual relations exist between the concepts is drawn as an arrow. In linear form, the boxes are abbreviated as square brackets and the circles as rounded parentheses.

The textual representation of a conceptual graph in linear form is less cumbersome and more easily parsed and dealt with from an analysis point of view and will be used from this point forward. Figure C.2 can be written as:

```
[ACTION] –
    (OBJ) -> [ENTITY]
    (AGNT) -> [ENTITY]
    (TEMP) -> [TIME]
    (CONT) -> [CONTEXT]
```

The hyphen on the first line indicates that the relations attached to `[ACTION]` are continued on subsequent lines. A longer version could have been written as follows (the English-like translations appear in brackets after the graphs):

`[ACTION] –> (OBJ) -> [ENTITY]`	(an entity is the object of an action)
`[ACTION] –> (AGNT) -> [ENTITY]`	(an entity is the agent of an action)
`[ACTION] –> (TEMP) -> [TIME]`	(a time is the temporal value of an action)
`[ACTION] –> (CONT) -> [CONTEXT]`	(a context is the context of the action)

This can also be more simply expressed in a condensed English-like sentence and read, *an entity is the object of an action on an entity at some time and in some context.* Therefore, the sentence *a cat entity is the agent of the action sitting whose object is a mat entity, whose temporal value was the time Tuesday and whose context was in the rain* could be expressed as:

```
[SAT] –
    (OBJ) -> [MAT]
    (AGNT) -> [CAT]
    (TEMP) -> [TUESDAY]
    (CONT) -> [RAIN]
```

It could also be condense further into a more sensible English-like sentence as *the cat was sitting on the mat on Tuesday in the rain.*

Notice that the labels for the conceptual relations are kept relatively short and usually abbreviated (three or four characters) to keep the notation brief. For future reference,

Table C.1 displays the conceptual relations to be used in this appendix along with their abbreviations.

Conceptual Relation	**Abbreviation**
object	OBJ
agent	AGNT
own	OWN
context	CONT
location	LOC
temporal	TEMP
past	PAST
instrument	INST
manner	MANR
between	BETW

Table C.1 Conceptual Relation Abbreviations

Like the concept type, the conceptual relation has two parts, the *relation type* and a nonnegative integer called its *valence.* The conceptual relation acts as a connector between concept types and specifies the role that exists between them. The relation type is a label that explains the nature of the relationship and the valence is the number of arcs that belong to the conceptual relation. Most of the common relations are dyadic (with two arcs) such as:

```
[ACTION] -> (OBJ) -> [ENTITY]
```

However, a conceptual relation may have any number of arcs for example,

```
[ACTION] -> (PAST)
```

depicts an action concept occurring in the past with `(PAST)` as a monadic conceptual relation and

```
[PERSON: Joe] <- (BETW) –
                      <- [HOUSE]
                      <- [CAR]
```

describing a person called *Joe* being between *a car* and *a house* where the conceptual relation `(BETW)` is triadic as the concept `[PERSON: Joe]` is one and the same thus giving the `(BETW)` conceptual relation three arcs.

In summary and as formally defined by the NCITS.T2 Committee on Information Interchange and Interpretation in the Conceptual Graph Standard of March 21, 2001 (Sowa 2001), A *conceptual graph g* is a bipartite graph, which consists of two kinds of nodes called *concepts* and *conceptual relations*.

- Every *arc a* of *g* is a pair (*r, c*) consisting of a conceptual relation *r* and a concept *c* in *g*. The arc *a* is said to *belong* to *r*; it is said *link r* to *c*; but it does not belong to *c*.
- A conceptual graph *g* may have concepts that are not linked to any conceptual relation; but every arc that belongs to any conceptual relation *r* in *g* must link *r* to exactly one concept *c* in *g*.
- Three kinds of conceptual graphs have distinguished names:
    1. The *blank* is an empty conceptual graph with no concepts, conceptual relations, or arcs.
    2. A *singleton* is a conceptual graph that consists of a single concept, but no conceptual relations or arcs.
    3. A *star* is a conceptual graph that consists of a single conceptual relation *r*, every arc that belongs to *r*, and every concept *c* that is linked by some arc (*r, c*) that belongs to *r*.

Both the concept and conceptual relation nodes are labeled with English words. These are percepts that are understood by the semantic network where the conceptual graphs have been created and also aid in the conversion of the graphical notation into English sentences. These labels are organised within the semantic network as hierarchies known as type hierarchies and relationship hierarchies. Both hierarchies are partially-ordered sets of labels that support the inheritance of properties from supertypes to subtypes. Supertypes within the hierarchy are more abstract and subtypes are more primitive. The concepts in the CG of Figure C.2 are quite abstract. Examples of these hierarchies are illustrated in Figure C.3 and C.4.

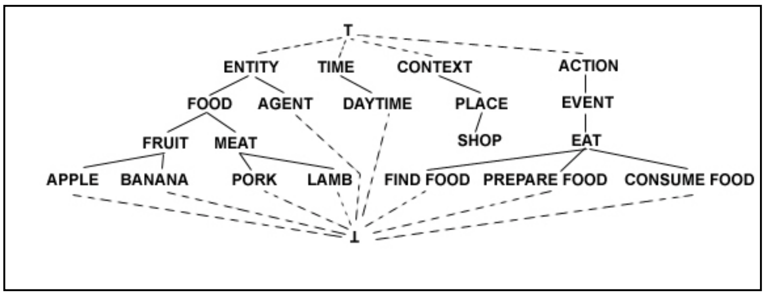


Figure C.3 An Example Type Hierarchy

Figure C.3 shows a simple Type Hierarchy. The type denoted by **T**, at the very top of the hierarchy, is the most abstract type and is known as the *universal* type. All other types in the hierarchy are subtypes of the universal type. As can been seen, moving down through the hierarchy leads to subtypes that have the properties of their parents. At the very bottom of the hierarchy is the type abstract denoted by ⊥.

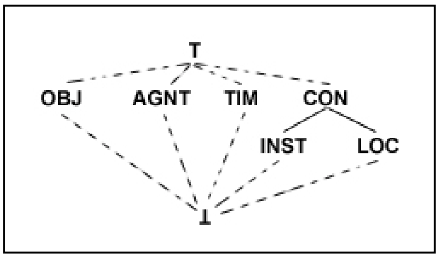


Figure C.4 An Example Relation Hierarchy

The hierarchy in Figure C.4, which shows a Relation Hierarchy, is structured in the same manner as the Type Hierarchy. There is a universal relation that acts as the common supertype to all other relations. In this example, the relation **CONT** (context) is a parent to the relations **LOC** (location) and **TEMP** (temporal). These relations could be used to describe the location of an entity or when a particular action occurred.

Concept types from the Type Hierarchy can be combined with conceptual relations in the Relation Hierarchy to form countless conceptual graphs. However, not all combinations of concept nodes and conceptual relation nodes will make sense. For example, in (Sowa 1984), Sowa gives the following example:

```
[SLEEP] -> (AGNT) -> [IDEA] -> (COLR) -> [GREEN]
```

which may be read *Some act of sleeping has an agent, which is an idea, which has a colour, green.* To eliminate such nonsensical conceptual graphs, Katz and Fodor (Katz and Fodor 1963) created a set of semantics that impose constraints on the combinations of concept and conceptual relation nodes that are allowed. These constraints allow for the creation of meaningful graphs that represent possible situations in the world. These graphs are called *canonical*.

C.2) Canonical Graph Formation

There are three processes by which a canonical conceptual graph may be formed. These are:

1. When a new real world situation is encountered and a graph is formed by means of perception, the graph formed by the new combination of concept nodes and conceptual relation nodes is deemed canonical.
2. Graphs that are introduced into the semantic network by a knowledge engineer are referred to in conceptual graph terms to resemble insight and are considered canonical.
3. New conceptual graphs can be formed through a series of operations called formation rules. Although there are a number of formation rules that can be applied to conceptual graphs there are only four considered to produce canonical graphs. These are *copy, restrict, join* and *simplify*.

As the first two processes listed above could be achieved in a variety of ways, they are not discussed here. However, the formation rules of item three are part and parcel of conceptual graph theory and therefore, warrant definition at this point.

As its name suggests, the *copy* rule creates a copy of a conceptual graph. If the knowledge base of an application includes a conceptual graph g_1, then by copying g_1 another conceptual graph called g_2 can be produced that is an exact replica of g_1. As g_1 existed in the knowledge base as a canonical graph, then g_2 must also be a canonical graph. It may seem pointless to have two conceptual graphs that are exactly the same in the knowledge base, however, if a conceptual graph is manipulated by another formation rule, then it becomes different and the original is lost. Therefore, it is pertinent to make a copy of a conceptual graph before applying other formation rules on it, if the original conceptual graph will be needed later.

The *restrict* rule can perform changes to concept types or referents or the concept nodes of a conceptual graph. The restrict rule allows the concept type to be constrained to another type that exists in the Type Hierarchy as a subtype of the concept type. For example, given the Type Hierarchy of Figure C.3 and the conceptual graph,

```
[ACTION] -> (AGNT) -> [ENTITY]
```

the concept `[ACTION]` could be restricted to the concept `[EAT]` forming the new canonical conceptual graph,

```
[EAT] -> (AGNT) -> [ENTITY]
```

Furthermore, the concept `[ENTITY]` could be restricted to the concept `[ANIMAL]` or even further to the concept `[DOG]` to create the graph:

```
[EAT] -> (AGNT) -> [DOG]
```

To prevent the concept [ENTITY] being restricted to a concept such as [ROCK] that would make a nonsensical graph such as:

```
[EAT] -> (AGNT) -> [ROCK]
```

a base set of canonical graphs can be included within a conceptual graph processor to place selectional restrictions on the concept types of restrict operations that can be performed. In the above example, given the canonical graph `[EAT]->(AGNT)->[ANIMAL]`, the concept `[ENTITY]` could never be restricted to a concept type that was not `[ANIMAL]` or a subtype thereof.

The restrict rule can also perform restrictions on a concept type by adding a referent when none exists. For example, the above graph could be restricted to a particular instance of the concept type `[DOG]` and create the graph:

```
[EAT] -> (AGNT) -> [DOG: Deefa]
```

The *join* rule takes two canonical conceptual graphs and joins them at the intersection of concept nodes. This operation is formed by identifying an identical concept in two graphs, deleting this concept from one of the graphs and then linking all arcs from the deleted concept to its identical concept in the other graph. For example, consider the following two graphs:

```
[EAT] -> (OBJ) -> [BONE]
```

and

```
[EAT] -> (AGNT) -> [DOG: Deefa]
```

By using the *join* rule and specifying the concept `[EAT]`, the following new conceptual graph can be formed:

```
[EAT] –
    (OBJ) -> [BONE]
    (AGNT) -> [DOG: Deefa]
```

This graph can be read *a dog called Deefa eating a bone*.

The final canonical formation rule is *simplify*. This rule acts on a conceptual graph to remove duplicate concept nodes if there are any present. This operation is usually performed after graphs have been joined to remove any redundant nodes.

For example, the graph:

```
[EAT] –
    (OBJ) -> [BONE]
    (AGNT) -> [DOG: Deefa]
```

is joined by the concept `[EAT]` to the graph:

```
[EAT] –
     (AGNT) -> [DOG: Deefa]
     (TEMP) -> [TUESDAY]
     (LOC) -> [BACKYARD]
```

to form the graph:

```
[EAT] –
     (AGNT) -> [DOG: Deefa]
     (TEMP) -> [TUESDAY]
     (LOC) -> [BACKYARD]
     (OBJ) -> [BONE]
     (AGNT) -> [DOG: Deefa]
```

As the resultant graph, contains a redundant link to the concept `[DOG: Deefa]` from the concept `[EAT]` via the conceptual relation `(AGNT)`, the above graph can be simplified to remove the redundant link and produce the graph:

```
[EAT] –
    (AGNT) -> [DOG: Deefa]
    (TEMP) -> [TUESDAY]
    (LOC) -> [BACKYARD]
    (OBJ) -> [BONE]
```

C.3) Generalisation and Specialisation

The canonical formation rules of join and restrict, described in the previous section, are also known as *specialisation rules*. This is due to the nature of the resultant conceptual graphs. Specialisation occurs when concepts in a graph become less generic by converging the concept type to one of its subtypes in the Type Hierarchy or when a specific referent is designated.

The operations that reverse the effects of specialisation are called generalisations. After a graph has been generalised the concept type labels become more generic taking on type labels of their supertypes from the Type Hierarchy. It should be noted that generalisation does not necessarily preserve the canon. For example the graph:

```
[DOG] <- (AGNT) <- [EAT] -> (OBJ) -> [BONE]
```

could be generalised on both the [DOG] and [BONE] concepts resulting in the graph:

```
[ENTITY] <- (AGNT) <- [EAT] -> (OBJ) -> [ENTITY]
```

This graph implies that *some entity eats another entity* and while this may be true in some circumstances, it is not the case for all combinations of `[ENTITY]` subtypes that could be specialised into this graph from the Type Hierarchy. For example, the graph:

```
[BONE] <- (AGNT) <- [EAT] -> (OBJ) -> [DOG]
```

or *some bone is eating a dog*, is a nonsensical specialisation. Therefore, by generalising the first graph that defined the constraints that only a dog can eat a bone, the knowledge base has lost its canonical nature. However, operations that result in generalisations of graphs are necessary from time to time because they form the basis for logic within the knowledge base and are necessary for adding new insights (see Section 2 of this appendix).

C.4) Type and Relation Expansion and Contraction

Type and Relation expansion in conceptual graph theory is somewhat similar to variable substitution and expansion that takes place in mathematical expressions. Given a definition for a type in the Type Hierarchy or a relation in the Relation Hierarchy, these conceptual graphs can be substituted into existing conceptual graphs where the type or relation appears. This will expand the conceptual graph and is called type or relation expansion depending on the concept being replaced. The opposite operation that replaces sections of conceptual graphs with one type or relation concept is called contraction.

Consider the following type and relation definitions:

```
type KISS(x) is
     [TOUCH: *x] –
           (AGNT) -> [PERSON] -> (PART) -> [LIPS]
           (MANR) -> [TENDER]
```

If a graph existed thus:

```
[KISS] –
   (AGNT) -> [PERSON: HARRY]
   (OBJ) -> [PERSON: MARY]
```

it could be expanded by replacing the concept type `KISS` with the type definition for `KISS` above. The resulting graph would look like this:

```
[TOUCH] –
   (AGNT) -> [PERSON: HARRY] -> (PART) -> [LIPS]
   (MANR) -> [TENDER]
   (OBJ) -> [PERSON: MARY]
```

Type contraction would return the above graph back to its former state. Relation contraction and expansion works in exactly the same way.

C.5) Other Graph Operations

Before concluding this appendix, two other operations that are used in manipulating conceptual graphs should be mentioned. These are the rules of *projection* and *maximal join*.

The projection rule allows one conceptual graph to be projected onto another conceptual graph resulting in a graph that consists of concept nodes and conceptual relation nodes that are present and directly connected in each graph. Basically, it is finding a graph that is a common subgraph of the two given graphs where specialisations are held. For example, given the graphs:

```
[ACTION] -> (AGNT) -> [ENTITY]
```

and

```
[DOG: Deefa] <- (AGNT) <- [EAT] -> (OBJ) -> [BONE]
```

by projecting the first graph onto the second graph, the following graph results:

```
[DOG: Deefa] <- (AGNT) <- [EAT]
```

Maximal join operates in a similar manner to the join rule, except that the graphs are joined on all maximally common projections. For example, given the graphs:

```
[EAT] –
    (AGNT) -> [DOG]
    (MANR) -> [FAST]
```

and

```
[EAT] –
    (AGNT) -> [DOG]
    (OBJ) -> [FOOD]
```

the resultant maximally joined graph would be:

```
[EAT] –
    (AGNT) -> [DOG]
    (OBJ) -> [FOOD]
    (MANR) -> [FAST]
```

where the graphs have been joined on both the `[EAT]` concept and the `[DOG]` concept in the same operation.

C.6) Conclusion

There have been many books, tutorials and conference papers written on conceptual graph theory. The nature of conceptual graphs is as broad as it is deep. This appendix has provided a brief introduction to the topic and included necessary syntactical and operational information necessary to the comprehension of the use of conceptual graphs throughout this dissertation.

Appendix D

Evaluation Questionnaire

Simulation of Fido.

Background: EMAI (Emotionally Motivated Artificial Intelligence) is the software agent that has been designed by Penny Baillie as the result of her research into emotional intelligence for her PhD. EMAI can be used to simulate any type of computer character. For this experience, EMAI is simulating a simple computerised pet called Fido. Fido is a simulation of a pet dog with simplistic behaviours. The simulation has not been designed to evaluate how well EMAI simulates a pet dog, but rather to gather your evaluation about the emotional intelligence produced by the agent. A pet dog was chosen as the computer character as most people are familiar with and have expectations about the behaviour of pet dogs.

Interacting with Fido

If you received the fido.zip file via email you will need to unzip the contents into a new directory. A directory called EMAI on your desktop will do. EMAI (Emotionally Motivated Artificial Intelligence) is the model that has been developed during the research for Penny's Ph.D. For this experiment/evaluation EMAI is being used to model a pet dog called Fido.

There are two executable files for you to evaluate; fidogreen.exe and fidoblue.exe. Each program implements a different type of model for producing emotions and behaviours in the pet dog Fido. The evaluation that you are about to undertake will assist in the further development of the EMAI software. For this reason your answers about each model need to be as honest as possible.

What you wo not be evaluating

Not having the resources, skills, staff and money of multimillion-dollar software companies, the EMAI software has a primitive interface, no sound and no sophisticated computer graphics so you will not be evaluating the software for these things.

The actual behaviours of a pet dog are quite complex. Although EMAI is simulating a pet dog, you will not be evaluating the accuracy or depth of the pet's range of actions.

What you will be evaluating

The purpose of this evaluation is to obtain your opinion about the emotional intelligence of the pet dog. Things that this will include are goal setting, motivation, attitude development, learning and emotional states.

How to Proceed

People's experiences with Fido can be many and varied. In order to gather data about similar experience with Fido the following exercises have been created for you to follow.

Read an exercise, run the EMAI software and perform the exercises with Fido then answer the questions. You will need to perform this routine twice; once with **Fidogreen** and once with **Fidoblue**. When finished with the exercises close the program. On closing you will find that a text file has been generated in the same directory as the programs, they will be called **fidogreen.txt** and **fidoblue.txt**. You can use these to review your simulation if you need to. These files will be overwritten each time you run the program.

When you are finished return your questionnaires to Penny. Exercises begin on page three.

EMA's User Interface

This document explains the interface for EMA's simulation of Fido, a computer character pet dog.

Motivational Urges

The motivational urges section of the interface consists of three urge gauges. These gauges represent 1) how hungry Fido is; 2) how sleepy Fido is, and 3) how much exercise Fido wants. These are controlled by the simulation and the user cannot adjust the values.

Environmental Display

black window where an image of Fido is displayed. The image and simple animations of Fido illustrate Fido's behaviours and emotions. Along with the image of Fido, the environmental display also shows which items you have placed in Fido's environment. These are displayed as small individual icons.

Simulation Narration

Tool Bar

to put items in and take items out of Fido's environment. When the user adds an item to the environment a corresponding icon will appear in the environmental display. The second part of the tool bar are two buttons that control the running of the simulation. The pause button allows the user to pause the simulation at anytime to review the narration and the start button starts the simulation running again.

Discipline Buttons

The final items on the EMAI interface are the large buttons with the sad and smiley faces . These buttons can be pressed by the user to tell Fido whether or not he is bad or good. The buttons can be pressed any number of times. For example if Fido is chewing the shoe, the user can click on the BAD DOG button several times to emphasize to Fido just **how**

Running the Simulation

X button in the top right hand corner. On closing the program the narrative from the simulation run will be written to a file called "fidogreen.txt" or "fidoblue.txt" in the same directory as the EMAI software.

Exercises

Introduction - IMPORTANT

To perform these exercises select one of the Fido simulations, either Fidogreen or Fidoblue. Perform ALL exercises and complete the questionnaire for ONE FIDO ONLY before repeating the exercises and questionnaire for the other Fido.

Each questionnaire should take you about 45mins to complete. If you do not have time to complete the questionnaires for both Fidos, try to finish one in a sitting then come back later and complete the other questionnaire. Please try and complete these questionnaires as quickly as possible. You will need to print out 2 copies of the questionnaire to complete if you do not already have a paper copy.

Exercise 1: Default Behaviour.

a) Run Fido.

b) Before beginning run the mouse over each part of the screen. Hint boxes will pop up that informs you what each part/button is for.

c) Click on the Start Button to Begin. Notice that Fido initially has low motivation for all of the motivational urge gauges (at the top). As the values on the motivational urge gauges go up, Fido will begin to do things.

d) For this exercise just watch what Fido does without interacting with him.

e) When Fido's *hungry* gauge reaches the extreme right, wait until "FIDO IS HUNGRY" has printed in the narrative box 3 times and press the pause button.

f) Use your experience in watching Fido and the narration in the text box to answer the questions about exercise 1 on the questionnaire.

g) Close Fido

Exercise 2: Simple Learning.

a) Run Fido.

b) For this exercise use the BAD DOG and GOOD DOG buttons ONLY.

c) By using the GOOD DOG button to praise Fido when he performs a good action and the BAD DOG button to reprimand Fido when he is bad try and teach Fido to prefer *playing with the ground* (rolling about) rather than other activities. Continue this until the *hungry* gauge reaches the extreme right, wait unit "FIDO IS HUNGRY" has printed twice. Press the pause button.

d) Use your experience in watching Fido and the narration in the text box to answer the questions about exercise 2 on the questionnaire. DO NOT CLOSE FIDO WHEN FINISHED.

e) After answering the exercise 2 questions, continue to Exercise 3 with the same Fido program open.

Exercise 3: Satisfying Motivational Urges

a) Continuing from the last exercise, press the start button to start the simulation going.

b) Click on the Food Button to give Fido some food.

c) Once Fido has finished eating the food, take notice of his emotional state and his next behaviour.

d) When Fido is *hungry* again, click on the Food Button again. Continue this until you have served up about 5 lots of dog food. NOTE: You cannot add more dog food until he has begun eating the lot before.

e) Click on pause and answer the questions for exercise 3.

f) Close Fido

Exercise 4: Complex Learning

a) Run Fido.

b) Add the ball and towel to the environment.

c) Teach Fido to that playing with the ball is GOOD and playing with the towel is BAD.

d) Do not forget to give Fido dog food when he appears *hungry*!!

e) When you think that Fido has been trained, attempt to reverse the Fido's training. See if you can make him think that the towel is GOOD and the ball is BAD.

f) When you think you have done this take the towel away. Observe Fido's behaviour.

g) After a while, add the towel back into the environment.

h) Click on pause and answer the questions for exercise 4.

i) Close Fido

Exercise 5: Complex Behaviour

a) Run Fido.

b) For this exercise it is your task to train Fido to like an item of your choice for playing with. Add all of Fido's items to the environment including food items and try to teach Fido to prefer playing with and eating certain objects. Get Fido to have several favourite items, so if you take one away he plays with the other.

c) Observe Fido's behaviour.

d) Click on pause and answer the questions for exercise 5.

e) Close Fido

Overall Impression of Fido

Complete the remainder of the Questionnaire. If you need to review your interaction with Fido, there will be a .txt file in the same directory as the program that has your last interaction recorded. (fidogreen.txt or fidoblue.txt), or you can run the Fido program again.

Repeat Exercises and Questionnaire NOW for the other Fido Simulation.

Questionnaire

These evaluations were made about which Fido? ❑ Green ❑Blue

How to fill out this questionnaire:

You will need to fill out TWO copies of this questionnaire. One for Fidogreen and one for Fidoblue.

The answers to most questions will provide you with a rating scale like this:

none of the time | | | | | | | | all of the time

Place an X in the appropriate box in the range to indicate your answer. e.g. for all of the time put:

none of the time | | | | | | | X| all of the time

for none of the time put:

none of the time | | | | X| | | | all of the time

for half of the time put:

none of the time |X | | | | | | | all of the time

For other answers, tick the yes or no box where indicated, or for an explanation, print a couple of brief sentences.

Exercise 1

a) How did the progression of Fido's behaviour's appear to you?

erratic |__|__|__|__|__|__|__|__| calculated

Did you feel that Fido's behaviour was driven by the values on the motivational urge gauges (the three gauges at the top of the window)?

❑ Yes ❑No

Explain with an example:

b) How would you rate Fido's changing emotional state taking into consideration his behaviour?

c) How would you rate Fido's changing behaviours taking into consideration his emotional state?

unreasonable |__|__|__|__|__|__|__|__| reasonable

Exercise 2

a) Did you find Fido easy to train? ❑ Yes ❑No

b) Do you think that Fido learned anything during the training? ❑ Yes ❑No

c) How would you rate Fido's emotional state with respect to your training techniques?

unreasonable |__|__|__|__|__|__|__|__| reasonable

d) Briefly explain your training techniques and Fido's emotional and behavioural responses.

Exercise 3

a) Did Fido eat only after he expressed he was *hungry*?

never |__|__|__|__|__|__|__|__| always

b) When Fido was *hungry* and there was food available did Fido eat the food before doing something else?

never |__|__|__|__|__|__|__|__| always

c) During the times when he was *hungry*, did Fido seem to retain knowledge of his previous training?

❑ Yes ❑No

d) How would you rate this type of behaviour?

unreasonable |__|__|__|__|__|__|__|__| reasonable

Exercise 4

a) How easy do you think that Fido is to train?

impossible |__|__|__|__|__|__|__|__| very easy

b) Overall, do you think that Fido prioritises his behaviours? ❑ Yes ❑No

c) How would you rate Fido's emotional states during the training?

unreasonable |__|__|__|__|__|__|__|__| reasonable

Exercise 5

a) How easy do you think that Fido is to train in multiple tasks?

impossible |__|__|__|__|__|__|__|__| very easy

b) Do you think that Fido was able to devise a list of favourite and unfavourite items?

❑ Yes ❑No

c) How would you rate Fido's emotional states during this exercise?

unreasonable |__|__|__|__|__|__|__| reasonable

d) Do you think that Fido learned from his emotional states? ❑ Yes ❑No

Explain:

An Overview of your Evaluation of Fido

Use your experience with Fido during the exercises to answer these questions.

Intelligence

There are many and varied definitions of intelligence. For the purpose of this experiment, intelligence is being evaluated from several perspectives.

a. Human behaviour is considered as goal orientated. This means that we set goals and then try to satisfy them. Often there are several actions that can be taken to achieve the same goal. If one attempt fails, we can fall back on another plan of action to satisfy our goals. Goal setting also means that if we do not WANT to do something then we wo not.

Do you think that Fido set's goals? ❑ Yes ❑No

Did Fido ever perform actions that were not related to his WANTs? ❑ Yes ❑No

Do you think that Fido can formulate ways in which to satisfy his goals?

❑ Yes ❑No

Do you think that Fido tries other plans of action when one fails? ❑ Yes ❑No

b. Attitude has been defined as a person's predisposition toward objects, actions and other elements. These predisposition are referred to as *liking* or *disliking*. Do you think that Fido shows the ability to form attitudes?

❑ Yes ❑No

c. Do you think that Fido uses his attitudes to make decisions? ❑ Yes ❑No

d. Emotional Intelligence has been described as a person's ability to reason and make decision's based on emotional experiences and intuition.

Do you think that Fido shows the ability to form emotions? ❑ Yes ❑No

If yes, Do you think that Fido's emotions influence his behaviour? ❑ Yes ❑No

Explain:

e. Another definition of intelligence is the ability to adapt to certain environmental changes. Do you think that Fido has this form of intelligence?

❑ Yes ❑No

f. Overall how would you rate Fido's intelligence?

unreasonable |__|__|__|__|__|__|__|__| reasonable

Explain:

Learning

a. How would you rate Fido?

very disobedient |__|__|__|__|__|__|__| very obedient

b. Do you think Fido learns about his environment to plan for his actions?

❑ Yes ❑No

c. Do you think Fido learns about good and bad actions? ❑ Yes ❑No

d. Do you think Fido learns about good and bad objects? ❑ Yes ❑No

e. Do you think Fido learns about his behaviours from his emotional reactions?

❑ Yes ❑No

f. Overall how would your rate Fido's learning using the criteria listed above? Explain your answer.

Behaviour

a. Overall, how would you rate the Fido's behaviour?

unpredictable |__|__|__|__|__|__|__| predictable

unreasonable |__|__|__|__|__|__|__| reasonable

Emotions

a. Do you think that Fido's emotional states were reasonable? ❑ Yes ❑No

Explain:

b. To what degree were you able to predict Fido's emotional response to your actions?

unpredictable |__|__|__|__|__|__|__| predictable

Personality

Explain in your own words what you think about Fido's personality. Rate your answer in terms of how the character conveyed this personality to you.

unpredictable |__|__|__|__|__|__|__| predictable

emotionally unstable |__|__|__|__|__|__|__| emotionally stable

stupid |__|__|__|__|__|__|__| intelligent

erratic |__|__|__|__|__|__|__| goal orientated

Briefly describe your feelings toward Fido:

Expectation

The following questions relate to your expectations of the Fido simulation.

a. Do you think that the actions that Fido performs throughout the simulation meet your expectations of a computerised character?

none of the time |__|__|__|__|__|__|__| all of the time

Explain:

b. Do you think the emotions produced by Fido seem reasonable?

none of the time |__|__|__|__|__|__|__|__| all of the time

Explain with an example:

c. During your interaction with Fido did you find that Fido produced actions that you thought were outrageous or out of the ordinary?

none of the time |__|__|__|__|__|__|__|__| all of the time

Give an example:

d. Did you find that at any time during the simulation that you experienced an emotional attachment to Fido?

❑ Yes ❑No

Explain:

Improvement

Suggest one improvement for the emotional and behavioural model used to simulate Fido and explain why it is needed.

Comments

Any other comments?

Debriefing Questions

Answer these questions AFTER you have completed all the exercises with BOTH Fidos.

In your opinion,

1. Which model possessed the best goal setting skills? ☐ Green ☐ Blue
2. Which model learned from your feedback the most? ☐ Green ☐ Blue
3. Which model seemed more intelligent? ☐ Green ☐ Blue

   Explain why?

4. How would you define the term 'emotionally intelligent'?
5. Which model seemed more emotionally intelligent? ☐ Green ☐ Blue
6. Overall what did you think of Fido Green?
7. Overall what did you think of Fido Blue?
8. Have you ever interacted with a program like Fido before? ☐ Yes ☐ No
9. If yes, what program? How does Fido compare to this?
10. Your final ratings of the Fidos compared to what you expect from a computerised pet dog.

	Fido Green	Fido Blue
Learning Skills	unreasonable |__|__|__|__|__|__|__| reasonable	unreasonable |__|__|__|__|__|__|__| reasonable
Behaviour	unreasonable |__|__|__|__|__|__|__| reasonable	unreasonable |__|__|__|__|__|__|__| reasonable
Emotional States	unreasonable |__|__|__|__|__|__|__| reasonable	unreasonable |__|__|__|__|__|__|__| reasonable
Goal Setting	unreasonable |__|__|__|__|__|__|__| reasonable	unreasonable |__|__|__|__|__|__|__| reasonable

Disclaimer:

I understand that in undertaking this survey I will remain totally anonymous. No personal details or my name will be published with this data. I also understand that I may discontinue this evaluation at anytime before or during the experiment. Data resulting from this survey will be collated with other surveys and statistically analysed and used in research papers and Penny Baillie's PhD. Any explanations that I give for certain answers could also be published however, I would still remain anonymous.

Signed____

Date:____/______/_______

Please return (just the questionnaire) by June 29th, 2001 to:

Penny Baillie, Department of Mathematics and Computing, University of Southern Qld. Toowoomba, Qld., 4350 Email: penny.baillie@usq.edu.au Fax: 4631 1775

References

Allen, J.F., Kautz, H., Pelavin, R. & Tenenberg, J. 1991, *Reasoning About Plans*, Morgan Kaufmann Publishers Inc., San Mateo.

Asimov, I. (1950) *I, Robot*, Gnome Press, New York.

Baeker, R. M., Grudin, J., Buxton, W. & Greenberg, S. (eds) 1995, *Readings in Human-Computer Interaction: Toward the Year 2000, 2nd Edition*, Morgan Kaufmann Publishers, San Francisco.

Baillie, P. & Lukose, D. 1999, 'To Believe, to Desire, to Intend: Motivating Autonomous Intelligent Agents with a Mind of their Own', in *Proceedings of Managing Enterprises 99, Newcastle, Australia*, 1999, The University of Newcastle, Newcastle, pp. 235-240.

Baillie, P., Lukose, D. & Toleman, M. 2000a, 'How to Give an Agent An Attitude', in *Proceedings of Agents in Simulation Workshop I, Passau, Germany,* 2000, University of Passau, Passau, pp. 7-12.

Baillie, P., Lukose, D. & Toleman, M. 2000b, 'Emotional Intelligence for Intuitive Agents: Agents with Attitude', in *Proceedings of PRICAI 2000, Melbourne, Australia,* 2000, Springer-Verlag, Germany, p. 818.

Baillie, P., Toleman, M. & Lukose, D. 2000c, 'Could Emotions be the Key to Real Artificial Intelligence', in *Proceedings of ISA'2000, Wollongong*, *Austrailia*, 2000, ICSC Academic Press, pp. 45-50.

Baillie, P., Toleman, M. & Lukose, D. 2000d, 'Emotional Intelligence for Intuitive Agents', in *Proceedings of AISAT 2000, Hobart, Australia*, 2000, University of Tasmania, Hobart, pp.134-139.

Baillie, P. & Lukose, D. 2001a, Navigating an Emotional Space for Intelligent Agents, *Technical Report SC-MC-0124*, *Department of Mathematics and Computing, University of Southern Queensland*, Toowoomba, Australia.

Baillie, P. & Lukose, D. 2001b, 'Urging Desire: Motivational Mechanisms for Intelligent Agents with Minds of Their Own', *Cybernetics and Systems,* Taylor and Francis, Philadelphia, vol. 30, no. 1, pp. 701-718.

Baillie, P. & Toleman, M. 2001, 'Creating an Emotional Space for Artificial Beings', in *Proceedings of Agents in Simulation Workshop II, Passau, Germany,* 2001, University of Passau, Passau, pp. 7-12.

Baillie, P., Lukose, D. & Toleman, M. 2001, Agents with Attitude. Technical Report, *Department of Mathematics and Computing, University of Southern Queensland*, Toowoomba, Australia.

Baillie, P. & Lukose, D. 2002, 'Multidimensional Affect Appraisals for Artificial Intelligences', in *Proceedings of ACE 2002, Vienna, Austria*, University of Vienna, Vienna.

Baillie, P. 2002, 'An Agent with a Passion for Decision Making', in *Proceedings of Agents in Simulation 2002, Passau, Germany*, University of Passau, Passau.

Baillie, P. Lukose, D & Toleman, M. 2002, 'Engineering Emotionally Intelligent Agents', in *Intelligent Agent Software Engineering,* eds. V. Plekhanova & S. Wermter, Idea Publishing Group, Hershey.

Balkenius, C. 1995, *Natural Intelligence in Artificial Creatures,* Lund University Cognitive Studies, Lund.

Bates, J., Loyall, A.B. & Reilly, W.S. 1992, 'An Architecture for Action, Emotion, and Social Behaviour', in *Proceedings of Artificial Social Systems: Fourth European Workshop on Modeling Autonomous Agents in a Multi-Agent World*, *Pittsburg*, 1994, Springer-Verlag, Germany, pp. 55-68.

Blumberg, B.M., P.M. Todd and Maes, P. 1996, 'No Bad Dogs: Ethological Lessons for Learning in Hamsterdam', in *From Animals to Animats 4: Proceedings of the Fourth International Conference on Simulation of Adaptive Behaviour*, *Cape Cod*, MIT Press/Bradford Books, Cambridge, pp. 295-304.

Canamero, D. 1997, 'Modeling Motivations and Emotions as a Basis for Intelligent Behaviour', in *Proceedings of the First International Conference on Autonomous Agents, New York*, 1997, ACM Press, New York, pp.148-155.

Clarke, A. C., and Kubrick, S. 1993, *2001: A Space Odyssey*, New American Trade Library, New York.

Cohen, P. R., Greenberg, M. L., Hart, D. M., and Howe, A. E. 1989, 'Trial by fire: Understanding the design requirements for agents in complex environments', *AI Magazine*, AAAI, Menlo Park, vol. 10, no. 3, pp. 32-48.

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

Elliot, C. 1992, The Affective Reasoner: A process model of emotions in a multi-agent system, Ph.D. Dissertation, Northwestern University.

El-Nasr, M. S. 1998, Modeling Emotion Dynamics in Intelligent Agents, M.Sc. Dissertation, American University in Cairo.

Fellous, J. 1999, 'The Neuromodulatory Basis of Emotion', *The Neuroscientist,* SAGE Science Press, Thousand Oaks, vol. 5, no. 5, pp. 283-294.

Fishbein, M. & Ajzen, I. 1975, *Belief, Attitude, Intention and Behaviour*, Addison-Wesley, London

Fogg, B. J. 1999, 'Persuasive Technologies: Introduction', *Communications of the ACM,* ACM Press, New York, vol. 42 no. 3, pp. 26-29.

Foster, R. & Smith, M. 2000, *Dog Behaviour: An Evolutionary View*, Foster & Smith, Inc. Bakersfield.

Furth, H. 1987, *Knowledge As Desire, An Essay on Freud and Piaget*, Columbia University Press, New York.

Galotti, K. M. 1989, 'Approaches to studying formal and everyday reasoning', *Psychological Bulletin*, American Psychological Association, Washington, vol.105, pp. 331-351.

Goleman, D. 1995, *Emotional Intelligence*. Bantam Books, New York.

Hanks, S., Pollack, M.E., & Cohen, P.R. 1993, 'Benchmarks, Test Beds, Controlled Experimentation, and the Design of Agent Architectures', *AI Magazine*, AAAI, Menlo Park, vol. 14, no 4, pp. 17-42.

Hanley, R. 1997, *The Metaphysics of Star Trek*, Harper Collins Publishers, Inc., New York.

Healey, J. 2000, Wearable and Automotive Systems for the Recognition of Affect from Physiology, *Technical Report*, MIT Media Laboratory, Cambridge.

Healey, J., Picard, R., & Dabek, F. 1998, 'A New Affect-Perceiving Interface and Its Application to Personalized Music Selection', in *Proceedings of 1998 Workshop on Perceptual User Interfaces*, San Francisco, CA, November 4-6.

Hirsh, H. 1998, 'Interactive Fiction', *IEEE Intelligent Systems*, Institute of Electrical and Electronics Engineers, Washington, vol. 13, no. 6, pp.12 - 21.

Hull, C. L. 1943, *Principles of behaviour*, Appleton-Century-Crofts, New York.

Jacobson, N. & Bender, W. 1996, 'Color as a Determined Communicator', *IBM Systems Journal* MIT, Cambridge, vol. 35, no. 3, pp. 526-538.

Jenkins, H. 1998, 'From Homer to the Holodeck: New Media and the Humanities', paper presented at the *Post Innocence: Narrative Textures and New Media Conference*, *Transforming Cultures Project*, University of Technology, Sydney, Australia, October 3, 1998.

Juslin, P. & Madison, G. 1999, 'The role of timing patterns in recognition of emotional expression from musical performance', *Music Perception*, The University of California Press, New Haven, vol.17, pp. 197-221.

Kaehms, B. 1999, 'Putting a (Sometimes) Pretty Face on the Web', *WebTechniques*, CMP Media, Issue: September 1999, http://www.newarchitectmag.com/archives/1999/09/newsnotes/.

Katz, J. & Fodor, J.A. 1963, 'The structure of semantic theory', *Language*, Linguistic Society of America, Washington, vol. 39, pp. 170-210.

Klein, J., Picard, R. & Riseberg, J. 1997, 'Support for Human Emotional Needs in Human-Computer Interaction', juried position paper presented at *HCI Research and Practice Agenda Based on Human Needs and Social Responsibility Workshop, CHI'97: Human Factors in Computer Systems conference*, Atlanta.

Koestler, A. 1967, *The Ghost in the Machine*, Penguin Books Ltd., London.

Kolb, B. & Whishaw, I.Q. 1990, *Fundamentals of Human Neuropsychology*, W.H. Freeman and Company, New York.

Kort, B., Reilly, R. and Picard, R. 2001, 'External Representation of Learning Process and Domain Knowledge: Affective State as a Determinate of its Structure and Function', in *Proceedings of AI-ED 2001*, San Antonio, Texas, pp. 64-69.

Lefton, L. A. 1994, *Psychology: Fifth Edition*, Allyn and Bacon, Boston.

Leuba and Lucas 1945, 'The effects of attitudes on descriptions of pictures', *Journal of Experimental Psychology, American Psychological Association*, Washington, vol. 35, pp. 517-524.

Lukose, D., 1993, 'Executable Conceptual Structures', in *Proceedings of ICCS'93*, Quebec City, Canada, pp. 223-237

Lukose, D. 1996, 'Modeling Extendible Mobile Agents', in *Proceedings of DAI'96*, Cairns, Queensland, Australia, pp. 32-47

Lukose, D., 1997, 'Complex Modeling Constructs in MODEL-ECS', in *Proceedings of the International Conference on Conceptual Structures (ICCS'97)*, Seattle, Washington, USA, August 4-8, 1997.

Luxo Jr. (short animated film) 1986, Pixar Animation Studios, Emeryville, California.

MacLean, P. 1958, 'Contrasting Functions of Limbic and Neocortical Systems of the Brain and their Relevance to the Psycho-physical Aspects of Medicine', *American Journal of Medicine*, Association of Professors of Medicine, San Francisco, vol. 24, no. 4, pp. 6711 - 6726.

Maes, P. 1995, 'Artificial Life Meets Entertainment: Life like Autonomous Agents', *Communications of the ACM*, ACM Press, New York, vol. 38, no. 11, pp. 108-114.

Magai, C. & Hunziker, J. 1998, 'To Bedlam and Part Way Back: Discrete Emotions Theory and Borderline Symptoms', *Emotions in Psychopathology: Theory and Research*, eds. W. F. J. Flack and J. D. Laird, Oxford University Press, Oxford.

Mateas, M. 1997, An Oz-Centric Review of Interactive Drama in Believable Agents, *Technical Report CMU-CS-97-156*, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA. June 1997.

Meech, J. & Marsh, S. 2000, Social Factors in E-commerce Personalization, in *Proceedings of the CHI 2000 Workshop Designing Interactive Systems for 1-to-1 E-commerce*, The Hague, NRC 43664.

Minsky, M. 1985, *The Society of Mind*, Simon and Schuster, New York.

Minsky, M 2000, *Future Models for Mind-Machines*, invited plenary paper presented at the Symposium on How to Design a Functioning Mind at the Artificial Intelligence and the Simulation of Behaviour 2000 Conference, Birmingham.

Munday, C., Cross, J., Daengdej, J., and Lukose, D., 1996. *CGKEE: Conceptual Graph Knowledge Engineering Environment User and System Manual*, Research Report No. 96-118, Department of Mathematics, Statistics, and Computing Science, University of New England, Armidale, 2351, N.S.W., Australia.

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

Okuda, M. & Okuda, D. 1997, *The Star Trek Encyclopedia: A Reference Guide to the Future*, Pocket Books, New York.

Padgham, L. & Taylor, G. 1997, 'A System for Modeling Agents having Emotion and Personality', *Lecture Notes in Artificial Intelligence*, Springer-Verlag. vol.12, no. 9, pp. 59-71.

Penrose, R. 1989, *The Emperor's New Mind*, Oxford University Press, New York.

Pert, C. & Synder, S.H. 1985, Neuropeptides and Their Receptors: A Psychosomatic Network, *Journal of Immunology*, The American Association of Immunologists, New York, vol.135, no.2, pp. 820-826.

Pert, C. B. 1997, *Molecules of Emotion*, Simon and Schuster, New York.

Petty, R.E. & Cacioppo, J.T. 1996, *Attitudes and Persuasion: Classic and Contemporary Approaches*, Westview Press, Boulder.

Picard, R. 1997, Does HAL Cry Digital Tears? Emotions and Computers, in *HAL's Legacy: 2001's Computer as Dream and Reality*, ed. D. G. Stork, MIT Press, London, pp. 279-304.

Picard, R. 1997, *Affective Computing*, The MIT Press, London.

Picard, R. 2000, 'Toward Computers that Recognize and Respond to User Emotion', *IBM Systems Journal*, MIT Press, Cambridge, vol.39, nos. 3&4. pp. 705-719.

Pisanich, G. & Prevost, M. 1996, 'Representing human characters in interactive games', in *Proceedings of 1996 Computer Game Developers Conference*, San Francisco, Miller-Freeman Inc, pp. 377-388.

Plutchik, R. 1980, *Emotion: A Psychoevolutionary Synthesis*, Harper and Row, New York.

Pollack, M. & Ringuette, M. 1990, 'Introducing the Tileworld: Experimentally evaluating agent architectures', in *Proceedings of the Eight National Conference on Artificial Intelligence*, MIT Press, pp.183-189.

Rao, A. S. & M. P. Georgeff, M.P. 1992, An Abstract Architecture for Rational Agents, in *Proceedings of Third International Conference on Principals of Knowledge Representation and Reasoning*, *San Mateo*, Morgan Kauffmann Publishers, San Mateo, pp. 439-449.

Rao, A. S. & Georgeff, M.P. 1995, BDI Agents: From Theory to Practice, *Technical Report*, Australian Artificial Intelligence Institute, Melbourne.

Reilly, W. S. N. 1996, Believable Social and Emotional Agents. Ph.D. Dissertation, Carnegie Mellon University.

Rogers, P. A. M. 2000, Acupuncture and Homeostasis of Body Adaptive Systems, *The Web Journal of Acupuncture*, http://users.med.auth.gr/~karanik/english/hels/helsfram.html.

Roseman, I. J., Jose, P.E. & Spindel, M.S. 1990, 'Appraisals of Emotion-Eliciting Events: Testing a Theory of Discrete Emotions', *Journal of Personality and Social Psychology*, American Psychologists Association, Washington, vol.59, no.5, pp.899-915.

Rosenfeld, A. 1997, ‘Eyes for Computers: How HAL Could See’, in *HAL's Legacy: 2001's Computer as Dream and Reality*, ed. D. G. Stork, The MIT Press, Cambridge, pp. 211-236.

Rothgeb, C. (ed.) 1973, *Abstracts of the Standard Edition of the Complete Psychological Works of Sigmund Freud*, International Universities Press, New York.

Russell S. & Norvig, P. 1995, *Artificial Intelligence: a modern approach*, Prentice-Hall Inc., Upper Saddle River.

Sansweet, S.J. & Zahn, T. 1998, *The Star Wars Encyclopedia*, Lucas Books.

Scherer, K. R. 1982, ‘Emotion as process: Function, origin and regulation’, *Journal of Experimental Psychology*, American Psychologists Association, Washington, vol. 29, pp. 497-510.

Schmidt, B. 2000, *The Modeling of Human Behaviour*, SCS Publication, San Diego.

Shaver, P. R., Schwartz, J., Kirson, D. & O’Connor, C. 1987, ‘Emotion Knowledge: Further exploration of a prototype approach’, *Journal of Personality and Social Psychology*, American Psychologists Association, Washington, vol. 52, pp.1061-1086.

Sloman, A. 2001, ‘Beyond Shallow Models of Emotion’, *Cognitive Processing,* Pabst Science Publishers, Lengerich, vol. 2, no. 1, pp. 177-198.

Smith, C. A. & Ellsworth, P.C. 1985, Attitudes and Social Cognition, in *Journal of Personality and Social Psychology,* American Psychologists Association, Washington, vol. 48, no. 4, pp. 813-838.

Smith, C. A. & Kirby, L.D. (eds.) 2000, *Consequences Require Antecedents: Towards a Process Model of Emotion Elicitation. Feeling and Thinking: The role of affect in social cognition*, Cambridge University Press, London.

Smith, C. A. & Lazarus, R.S. 1990, ‘Emotion and Adaption’, in *Handbook of Personality: Theory and Research,* ed. L. A. Pervin, Guilford, New York, pp. 609-637.

Smith, C. A. P. 1997, ‘A Causal Model of Individual Decision Making Under Time Pressure’, in *Proceedings of Americas Conference on Information Systems*, *New Orleans*, The Association for Information Systems, Tel-Aviv.

Sowa, J. F. 2001, *Knowledge Representation: Logical, Philosophical, and Computational Foundations*, Brooks Cole Publishing Co., Pacific Grove.

Sowa, J. F. 1984, *Conceptual Structures: Information Processing in Mind and Machine*, Addison-Wesley, Reading.

Stern, A. 1999, 'AI Beyond Computer Games', in *Proceedings of 1999 AAAI Spring Symposium, Artificial Intelligence and Computer Games*, Menlo Park, AAAI Press, pp.77-80.

Stern, A., Frank, A & Resner, B. 1998, 'Virtual Petz: A Hybrid Approach to Creating Autonomous, Lifelike Dogz and Catz', in *Proceedings of Second International Conference on Autonomous Agents*, ACM Press, pp.334-335.

Sternberg, R. J. 1985, *Beyond IQ*, Cambridge University Press, Cambridge.

Stork, D. G. 1997, Scientist on the Set: An Interview with Marvin Minsky, in *HAL's Legacy: 2001's Computer as Dream and Reality*, ed. D. G. Stork, The MIT Press, Cambridge, pp.15-32.

Tambe, M. 1997, 'Agent Architectures for Flexible, Practical Teamwork', in *Proceedings of the National Conference on Artificial Intelligence*, AAAI Press, pp.198-202.

Velasquez 1999, 'From Affect Programs to Higher Cognitive Emotions: An Emotion-Based Control Approach', in *Proceedings of Workshop on Emotion-Based Agent Architectures*, Seattle, USA, pp. 10-15.

Velasquez, J. D. 1998, 'Modeling Emotion-Based Decision-Making', in *Proceedings of the 1998 AAAI Fall Symposium Emotional and Intelligent: The Tangled Knot of Cognition,* Orlando, AAAI Press, Menlo Park, pp.164-169.

Velasquez, J. D. 1998, 'When Robots Weep: Emotional Memories and Decision-Making', in *Proceedings of the Fifteenth National Conference on Artificial Intelligence and Tenth Innovative Applications of Artificial Intelligence Conference, Madison, Wisconsin*, The MIT Press, Cambridge pp.70-75.

Wallace, S. & Laird, J. 1999, Toward a Methodology for AI Architecture Evaluate: Comparing Soar and CLIPS, in *Intelligent Agents VI*, eds. Jennings N. & Lesperance, Y., Springer, Germany, Lecture Notes in Computer Science vol. 1757, pp. 117-131.

Wells, S. and Nazarro, J. 1997, *Inside Blake's 7*, Boxtree, London.

Zajonc, R. B., Murchy, S.T. & Inglehardt, M. 1989, 'Feeling and Facial Efference: Implications of a novel observation', *Psychological Review*, American Psychologists Association, Washington, vol.96, pp. 395-416.