Affective Synthetic Characters

by

Song-Yee Yoon

Submitted to the Department of Brain and Cognitive Sciences
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2000

Author . . . .
Department of Brain and Cognitive Sciences
February 16th, 2000

Certified by . . . .
Gerald E. Schneider
Professor of Neuroscience
Thesis Supervisor

Certified by . . . .
Bruce M. Blumberg
Assistant Professor of Media Arts and Sciences
Thesis Co-Supervisor

Accepted by . . . .
Mriganka Sur
Head, Department of Brain and Cognitive Sciences

Affective Synthetic Characters

by

Song-Yee Yoon

Submitted to the Department of Brain and Cognitive Sciences
on February 16th, 2000, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

We have long wanted to create artifacts that resemble us: artifacts that not only look like us, but act, think and feel like us, with which we can have sympathy and empathy. In particular, the purpose of this thesis work is twofold: One is to understand animal behavior, the role of the motivation system and of affect, in terms of the theory embodied in the software architecture for modeling interactive synthetic characters. The other is a practical one: to build believable synthetic characters that we can interact with, utilizing our best understanding of animal behavior in creating characters that are perceived as sympathetic and empathetic to humans. We have done this with the help of a specific understanding of the roles of the motivation and affect systems of animals.

To accomplish these goals, I propose a creature kernel model which is largely based on the approach of Blumberg [7]. The creature kernel is modeled as a sum of four main parts, the perception, motivation, behavior and motor systems. Among these four components, despite the fact that the motivation system plays a crucial role in daily survival of creatures in nature, its functional importance has often been neglected in attempts to create *intelligent* artifacts because it has been thought of as the "opposite" of rationality. Thus, in the system proposed in this thesis, emphasis is placed on the roles of the motivation system and how it acts as the integrator of the four parts of the creature kernel, and enables a creature to exist as a functional whole.

The plausibility of the proposed system is demonstrated through two projects, for which the characters were built using the creature kernel. The first project is called **Sydney K9.0**, in which the main character is Sydney, a virtual dog. Human participants can train the dog to do certain tricks using various physical input devices: voice command, clicker sound, milkbone box and a training stick. Learning and training phenomena are observed as operant and classical conditioning, and it can be explained how each subsystem is functioning inside of the character's mind to implement that functionality. The other project is called **(void*): A cast of characters**, through which three distinctive characters – Earl, Elliot and Eddie – are introduced. In this project, a human participant can 'possess' one of these three characters using *buns-*

and-forks interface and control the possessed character's dance movement by wiggling the interface in various ways. Learning and adaptive change of attitude through the interaction as well as expression of different personalities and its effect on interaction are emphasized. How well the personalities were represented is explored through the results of a survey of a number of novice users of the system.

Thesis Supervisor: Gerald E. Schneider
Title: Professor of Neuroscience

Thesis Co-Supervisor: Bruce M. Blumberg

Title: Assistant Professor of Media Arts and Sciences

Acknowledgments

I thank everyone in the Synthetic Characters Group for friendship and for having been wonderful colleagues; Professor Gerald Schneider, Professor Bruce Blumberg, Professor Emilio Bizzi and Professor Sebastian Seung for being great teachers and the guidance that they provided to me during my time in graduate school; and all my family for their patience and love.

Contents

List of Figures

List of Tables

Part I

OVERTURE

People have long been fascinated with artifacts that resemble us. Fiction writers and engineers have described and tried to build artifacts that not only look like us, but act, think and, more recently, even to feel like us, and with which we can have sympathy and empathy.

The beginning of such efforts to build such artifacts can be traced back at least to 1938, when Thomas Ross made a machine which showed learning in a test of animal intelligence – finding the way out of a maze [137]. It was the first attempt to make a machine that would imitate a living creature in performance, as distinguished from appearance. Since then, there have been many researchers from various disciplines pursuing this kind of goal, some of whom have focused on parts of the system that are crucial for creating such artifacts and some of whom have dealt with the problem of building a whole system.

Building lifelike creatures satisfies two natural desires. One is the desire to understand ourselves, or living organisms in general, better. The process of implementation provides us the best way of finding out where in the system our lack of understanding or comprehension lies, or whether our knowledge really is correct. The other desire which we would like to satisfy by making lifelike artifacts is the desire to find another instantiation of ourselves for sharing the essence of our lives - and relieving solitude. Nowadays, such artifacts are labeled as toys, agents or robots to serve as friends, companions or servants, and are designed to function at various places in various contexts with the purpose of making our lives better.

This thesis describes work that extends such efforts and concentrates on the problem of building a whole system rather than focusing on a subsystem. The purpose of the work is to build a framework for constructing artificial creatures that can act and react with humans, in a believable manner.

(a) (b)

Figure 0-1: These are two snapshots of some of the software creatures for which I have implemented the behavior system. Throughout this thesis, I will discuss efforts to create lifelike characteristics of this kind of creature, characteristics which bring them to life – and enable them to interact with humans in a believable manner.

Chapter 1

An interesting parallel: history of work on artificial creatures, and evolution of real animals

The efforts toward building lifelike artifacts continued after Thomas Ross' compelling demonstration in 1938. His work led to various incarnations of lifelike artifacts, and that effort is still continuing in the forms of intelligent robots [16, 17], software creatures [2, 122, 9], and agents [96, 24].

Interestingly, parallels can be seen between implementation techniques and our understanding of how the brain of a living organism operates, and the development of such techniques bears some resemblence to the evolution of the brain.

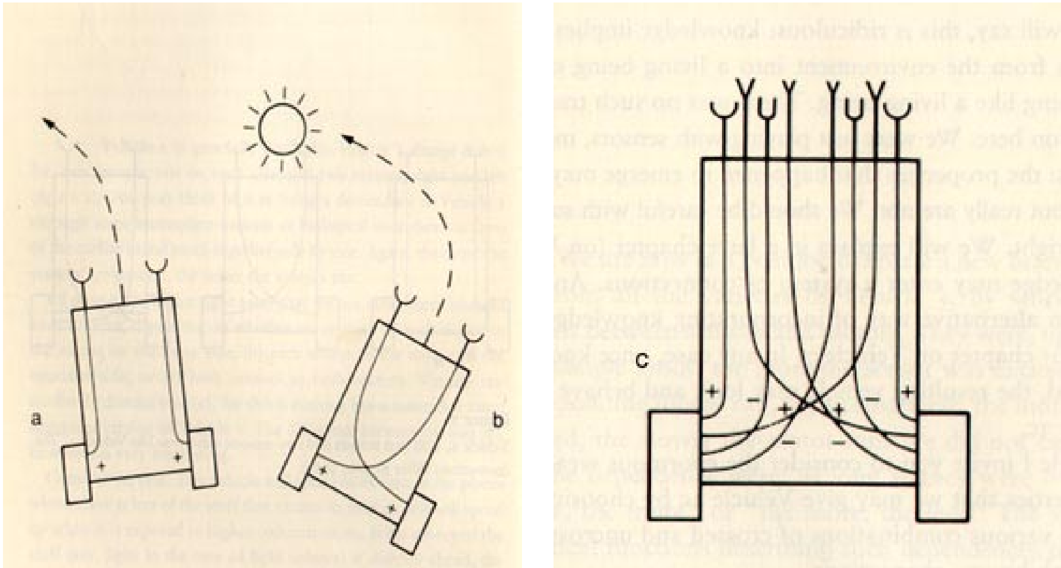
The simplest form of a creature can be described in a stimulus-response (SR) [68] framework. This framework stems from Descartes who invented the idea of the reflex arc in a living organism; Descartes' idea can be considered one of the first theories of how a living creature works. Straightforward S-R behavior patterns of animals that are optimized for particular environmental settings have been fixed by the genes, and can result in an animal being fooled by a *supernormal stimulus*, even if the response is not necessarily helpful for survival. This demonstrates the fact that at least some animal behavior patterns are tied to very specific stimulus configurations, and are elicited without analysis of other available information [126].

On the implementation side, Braitenberg suggested a number of vehicles based on a simple SR framework, which showed compellingly lifelike attributes [11]. Some of the first vehicles that appear in his book are not more than a couple of light sensors and wheels connected to the sensors. However, based on the way the wires are arranged, vehicles show either positive or negative phototactic behaviors, and people easily perceive them as vehicles "in love" with the light or which are "afraid" of the light, which are attributes that are usually ascribed only to living creatures. Braitenberg showed how interesting and life-like behaviors can arise from very simple SR connections.

Foreseeable problems arose as he attempted to take the obvious next step: the addition of more complex repertoires of behaviors while keeping the simple SR connection scheme. Connections necessarily got more and more complex, and coordination for invoking desired behaviors became more difficult (Figure 1-1). The evolution of the central nervous system probably has faced the same challange (Figure 1-2). The cerebellum and corpus striatum sit one level above the spinal cord and brainstem, which serve as the center of reflex arcs and action pattern generators, and the higher centers orchestrate them to generate desired behavior patterns [47, 43]. An obstacle for the Braitenberg vehicles' proceding one step further was the lack of such coordination centers, which are situated at one higher level or layer in the control hierarchy.

Efforts at finding a solution and creating mechanisms that deal with this coordination problem led to the implementation of central control mechanisms [142, 25] and then, to the coining, at a higher level of coordination, behavior engine and action selection mechanisms [15, 72, 8].

As the need for this coordination mechanism was realized by those who are interested



(a) (b)

Figure 1-1: (a) In his book, *Vehicles* [11], Braitenberg showed that proper wiring of sensors and servos preserving the simple SR framework is enough to evoke lifelike behaviors and emotional attributes such as love and hatred. (b) Provoking more complex patterns of behavior while keeping the simplest SR framework ends up complicating the circuit itself and it becomes harder to generate well coordinated behaviors of the creature (Coursey of the MIT Press, *Illustration, Maciek Albercht.*).

in constructing lifelike creatures, it was realized that such mechanisms have been conceptualized by those who closely observe real creatures – the ethologists [70, 68, 126]. In particular, Tinbergen proposed a hierarchical network to explain the way animals' behaviors are organized for meeting their survival needs, and Lorenz pursued the idea in more detail and explained various behaviors that he observed in his research in the framework of hierarchical organization with different operations that happen at each level.

These works of the ethologists further inspired those who are interested in constructing artificial lifelike creatures and led to interesting and promising results. Wilson started the new field of *animat* development [141], bringing together people who were interested in making intelligent systems with people interested in simulating the behavior of animals. Brooks's subsumption architecture led to success in constructing insect-like robots that operate reactively. Maes's action selection mecha-

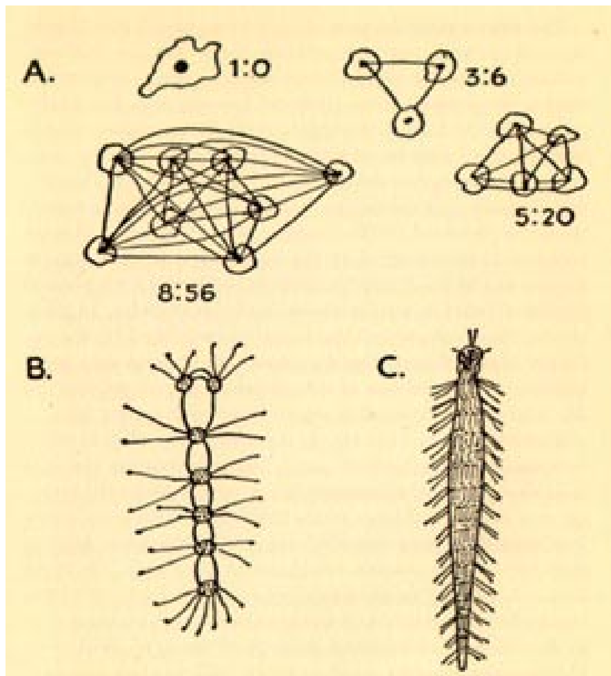


Figure 1-2: Figures from *The Living Brain*, W. Grey Walter [137]. This figure demonstrates the early evolution of brain as progress in the organization of nerve cells. (A) Jellyfish is a good example of a creature that probably has no gating mechanism other than a net of simple nerve cells. (B) Ganglia, or nerve knots, are really simple forms of brain that appear early as the mechanism for gating outputs from nerves; this coordination scheme is observed in insects. (C) As a slightly more advanced form, lancelots show a nerve cord with a budding brain.

nism has proven to be useful for designing sofware agents [72] through various implementations. Other implementations also proved that this approach is more robust and adaptive [138, 102, 8, 71] than traditional approaches and further proceeded to demonstrate its generalizability to coordinated behaviors among individuals [75, 123] in addition to coordination of behaviors within an individual. One can perceive these advances in the programming of robots and synthetic characters in terms of a biological analogy: the computer brains were evolving structures above the level of the spinal cord and brainstem [43, 22].

The next challange is to carry this creature with its expanded repertoire of behaviors into various novel environments. Just as in natural evolution, we are brought up against the notorious nature-versus-nurture controversy [26, 68, 45]. Some of the necsssary conditions for a species' survival are that it must have a structure that fits and performs well in the particular environment confronting it, and it must do this efficiently. It must have robust adaptive plans–plans which are efficient over the range of environments that it may encounter. One strategy for simplifying this problem would be maintaining a highly optimized structure for a certain environment. This strategy tends to sacrifice complex, plastic development for efficiency reasons. However, for a species to spread over a broad range of environments, a complex repertoir of behaviors is required or the behaviors must be able to change in order to remain adaptive in the face of such environmental variety [50].

Also, since organisms live in an ever-changing environment, they must be able to modify their adaptations in response to these changes. Thus, there are a variety of changes to which an organism must continually adapt during its lifetime. It must be able to institute short-term, reversible changes in response to temporary environmental changes. Finally, for its survival it is important for an animal to be able to learn from experience in such a way that it modifies its behavior adaptively.

The major mechanism for adding robustness to creatures is called *learning.* Learning depends on changes at different levels within the central nervous system. It may include simple sensitization or desensitization among neuronal connections as well as the use of strategic decisions in acquiring new plans.

Normally an animal's behavior is modified on the basis of individual experience so as to increase the animal's chances of survival. Such behavioral changes presuppose the presence of special phylogenetically developed genome-coded programs that prepare the individual for changes in its surroundings. Thus, the brain of an animal is programmed for adaptive changes of various kinds: it is programmed to learn.

Learning has directionality. It takes place to alter the behavior of the living organism to be performed more in one way than in other ways. This implies the need for a certain feedback mechanism for guiding this adaptive change and some kind of memory mechanism in addition, if the change is to be present over a certain period of time.

For learning, it is assumed that reward or punishment signals provide feedback necessary for learning. Positive reinforcement or reward follows a behavior that accomplishes something that increases the animal's chance of survival, and this, in turn, raises the likelihood of the animal's performing the behavior. By contrast, negative reinforcement or punishment follows a behavior that could decrease the animal's chance of survival, and this, in turn, reduces the likelihood of the animal's performing the behavior.

The evolutionary need for these components of learning mechanisms led to the appearance of *feel good* and *feel bad* circuitry associated with modules controlling approach and avoidance.

This circuitry was elaborated from visceral control mechanisms and from the chemosensory mechanisms which were so closely allied with feeding – the taste and olfactory systems – as well as from central mechanisms underlying pain and discomfort. The highly interconnected neuronal groups of this system have come to be called by neuroscientists the *limbic system* of the forebrain and midbrain [87, 14]. The activities of this collection of brain components are tightly connected to learning and memory.

The limbic system is seen as the center of motivation mechanisms which include drives and affect; it includes the neurons where activity can provide feedback with a directional bit – happiness and sadness, satisfaction and disappointment, each of which can serve as either reward or punishment. The result is a tendency to have the creature either pursue further the behavior which has brought the reward or suppress the behavior that originated the punishment, and grope for another way of doing things. From the learning perspective, drive and affect mechanisms are similar in the sense that both of them provide *signed* feedback signals to the creature. To reiterate, the appearance of memory and motivation mechanisms, which was possible through the evolution of limbic-system structures (including the hippocampus), made creatures become more robust and thus made it possible for them to survive in a broad range of environmental conditions.

Spier [112] has attempted to develop some of these ideas in computational modules. He developed a reactive learning framework based on the incentive learning paradigm. In his model, what he called motivation involves a minimal amount of explicit memory and, while still keeping the reactive system framework, he could demonstrate the ability to build an adaptive system which becomes updated through a well correlated, yet simple, *motivation* mechanism.

Change in quantitative values of the components that constitute the motivation system (i.e., change in neuronal activity states) not only serves as the basis for judging the plausibility of the behavior that triggered the change, but also may be expressed outwardly and communicated to other creatures in the form of *emotional expressions.*

One of Darwin's most significant theoretical statements on communication emerged from his consideration of animal emotions and the vehicles for their expression. In 1872, he set out to provide a general account of both acoustic and visual expressions of emotions in humans and nonhumans. As in all of his earlier discussions, he attempted to document and account for similarities and differences in expression across the animal kingdom by considering their design features and functions. The following quote captures the relevant details of Darwin's "bottom line."

> *No doubt as long as man and all other animals are reviewed as independent creations, an effectual stop is put to our natural desire to investigate as far as possible the causes of Expression. By this doctrine, anything and everything can be equally well explained; and it has proved pernicious with respect to Expression as to every other branch of natural history. With mankind some expressions, such as the bristling of the hair under the influence of extreme terror, or the uncovering of the teeth under that of furious rage, can hardly be understood, except on the belief that man once existed in a much lower and animal-like condition. The community of certain expressions in distinct though allied species, as in the movements of the same facial muscles during laughter by man and by various monkeys, is rendered somewhat more intelligible, if we believe in their descent from a common progenitor. He who admits on general grounds that the structure and habits of all animals have been gradually evolved, will look at the whole subject of Expression in a new and interesting light [28].*

Though this quote is restricted to visual expressions, Darwin adopted the same approach (descent with modification) to vocal expressions, and this included his explanation of function. Specifically, he argued that expressions were designed to convey information about the signaler's emotional or motivational state, with some signals reflecting an underlying ambiguity or conflict between different emotional states (his theory of "antithesis") such as fear and aggression.

Concerning communication, the grandfathers of ethology – Lorenz, Tinbergen, and von Frisch – largely accepted Darwin's treatment, especially the idea that signals were designed to communicate information about the signaler's motivational or emotional state. Thus Lorenz [67] explored the communicative exchanges between mothers and their young in a variety of avian species, focusing in particular on the signals and cues used during imprinting. For Lorenz, signals used during imprinting would insure attachment of offspring to the mother, and maternal response to offspring in distress.

Communications among organisms serve at least two different roles. One is coordination of behaviors among organisms by conveying behavioral intentions or necessary information for the good of the group. For example, Tinbergen [127, 128] looked at some of the key releasers during aggressive interactions between stickleback fish, in addition to exploring the variety of displays used by gulls during competitive interactions and courtship. Von Frisch [136] provided in-depth analyses of the honeybee's communication system, concentrating in particular on signals used to convey information about the location and quality of food. In all of these studies, very ritualized patterns were observed, each of which apparently has very clear meanings to the other animals of the same species.

With all these roles [1] being played, which also can be shown through various postural and behavioral configurations, the motivation system plays an important role in creating the lifelike impression made by synthetic creatures.

The communicative aspect of motivation caught the attention of those who were interested in building lifelike artifacts. Disney's magical creatures are probably the best example demonstrating what emotional attributes can add to a creature. Thomas et al [124] reflect that the addition of (exaggerated) emotional response to their characters was the key to bringing the illusion of life to the screen. Characters that smile or frown help the human beings watching or interacting with the characters believe that the characters *do* feel happy or sad, and are trying to convey the internal state of their feelings. People believe they have emotions, and thus have life.

[1] I discuss the roles of the motivation system in greater detail in Chapter 6

There has been a series of efforts for making artifacts with their own motivations and emotions in the robotics and software-agents domains as well [134, 37]. Breazeal [12] built a robot called Kismet that can express nine emotions: anger, surprise, fear, happiness, calm, interest, tired, disgust and sadness. In a social situation - interaction with a human user - Kismet displays one of those nine emotional states through its facial expressions. Duration and 'intensity' of certain types of interactions that the designer had in mind are the main factors that affect its drive states, which in turn, are used for selecting one of the nine emotional states as the primary emotion of that moment to be expressed through its face. Though Kismet expresses emotion which reflects the nature of its interaction with the human user, the system is designed within the framework of a kind of reflex model such that Kismet does not learn to try out a different strategy even if a certain situation causes it to undergo pain. While Breazeal's robot design is focused on a feed-forward operation of motivation, i.e., drives and emotional states influence behavior selection and facial expression, Velasquez's robot, Yuppy, was designed more with a feed-backward operation of emotion. Previous emotional experiences are fed back to the behavior system and influence future action selection strategies in the same or similar situation [134]. Within a behavior engine similar to that of Breazeal's, his robot keeps forming emotional memory, which affects behavioral attitude when it reencounters an object with associated emotional memory. However, since Yuppy does not have any generalization capability from the past experience, it does not know how to deal with objects or situations whose features - whether they cause pleasure or pain to the robot - were not pre-specified by the designer, and thus it does not show emotional response to a novel object or situation.

Bates led the OZ project [2, 76], which paid a lot of attention to emotional aspects of synthetic actors, called Woggles. Individual Woggles had specific habits and interests or baseline emotional states, which are shown as different personalities for different Woggles. However, designed with application to interactive cinema in mind, which is different from the goal of making an adaptive animat, emotion, as well as the rest of the components of a motivation system, functions like a set of rules for determining each Woggle's specific ways of reacting to events or objects.

Including the works mentioned above, the stream of research efforts for constructing artifacts with emotion led to many interesting demonstrations [96, 135]. However, those implementations have limitations primarily because emotion is modeled as superficial phenomena that make creatures *look* live, rather than as an important functional module that is crucial for proper functioning of the creatures, with emotional expression as one of the phenomena observable as the result of its operation. Instead of being carefully designed by programmers with specific aspects of emotional phenomena in mind, it should have been implemented as an emergent phenomenon that comes from the underlying motivation system that works as a part of the functional modules that are crucial for the operation of the creature. This is in accordance with a well-defined difference between the magical and the scientific imitation of life. The former copies external appearances; the latter is concerned with performance and behavior [137]. This criterion can be applied to the subparts of any system that claims to mimic life, and thus it is applicable to the motivation system, too.

Blumberg [6]'s work made an important advances beyond the older philosophy. He implemented an action selection method, and incorporated learning based on feedback from the drive system. The work demonstrated a way of understanding and implementing classical and operant conditioning within the behavior architecture. However, this implementation did not fully explore the affect (emotional feeling) part of the motivation system for providing reinforcers for learning, and thus was restricted to certain types of behavioral adaptation.

Chapter 2

The scope of this thesis

This thesis work aims at building artificial creatures that are like us and that can interact with us. I hope this attempt not only provides us with greater ability to build creatures that are fun to interact with but also contributes to our better understanding of living organisms, more specifically the mechanisms that underlie their operation.

The thesis builds mainly on work done by the Synthetic Characters Group, MIT Media Laboratory. The type of artificial creatures we create are called synthetic characters, and they can be described as 3D virtual creatures that are intelligent enough to do and express the right things in a particular situation or scenario [62]. The general research goal of our group is to create life-like synthetic characters that can interact with human participants in real time.

2.1 Problems

This goal of having artificial creatures intearact with human participants in real time imposes a number of challenges, which include both design issues and performance issues. Design issues include the structuring of a character's software brain so that it is intelligent enough to meet the demands of the interaction setting, and performance issues concerned with optimizing such software, using rapid graphics rendering routines and minimizing other hardware communication delays involved in implementing such scenarios well enough so that human participants do not lose interest in interacting with the characters. Among those various concerns, the following two are issues of special significance in my work.

Intentionality

Human participants can comfortably relate to interactive characters only when they feel that they can understand what is going on in the characters' minds. That is, they must be able to infer a character's beliefs and desires through its observed actions, and the quality of those actions [30]. Fundamentally then, a character's actions must be driven by and expressive of its beliefs and desires. This problem can be viewed as the need for communication between characters and human participants.

Adaptability

Real time interaction with human participants brings a strong measure of unpredictability to the virtual world. Not every situation can be predicted at the character design stage. So, it is very difficult to make characters behave in an intelligent manner solely based on the designer's comprehensive thought at the development stage. Furthermore, failures to show at least a very primitive level of intelligent behavior damages the life-like impression made by the characters. For example, it is hard to feel sympathy for a character that keeps approaching a participant who has been punishing the character at every instance of interaction, rather than avoiding such a participant. In other words, adaptability is crucial for it to "survive" as a character that can interact with humans in a compelling manner, where the details of interactions are not pre-scripted.

2.2 Proposed Approach

As discussed in Chapter 1, biological creatures are good sources that we can learn lessons from to solve problems that these interactive synthetic characters ought to be able to solve. Living in specific environments, biological creatures maintain both the hardware and software structures that have been evolved as viable solutions to the problem of survival. So, the approach this thesis is taking is to look for the solutions that those creatures have adopted and develop them as a mathematical and structural framework that can be easily incorporated as a design principle for building synthetic characters.

2.2.1 The motivation system

To look intentional, characters should know how to convey their intention. They need to be understandable characters, and thus, communicative characters. And, as I discussed in Chapter I, a good part of communication comes from facial, gestural and behavioral expressions of emotions, which is part of the function of the motivation system; this is exclusive of more explicit forms of communication such as language.

This external communication medium as a role of the motivation system has been witnessed in not only humans but also other creatures in nature. For example, monkeys may communicate their emotional state to others by making an open-mouth threat to indicate the extent to which they are willing to compete for resources, and this may influence the behavior of other animals. This aspect of the motivation system is well emphasized by Darwin [28], and has been studied more recently in humans by Ekman [35, 36]. He reviews evidence that humans can place facial expressions into the categories of happiness, sadness, fear, anger, surprise and disgust, and that this categorization may operate similarly in different cultures. He also describes how the facial muscles produce different expressions. Further investigations into the degree of cross-cultural universality of facial expressions, its development in infancy, and its role in social behavior are described by Izard [56] and Eibl-Eibelfeldt [33].

On the other hand, we have also discussed the role that the motivation system plays in providing reinforcement which is crucial for learning, which, in turn, is a crucial component for making creatures robust and adaptive. This brings us back to the need for the motivation system not just as a superficial addition to creatures, but with its proper implementations based on functional connections with other parts of the system.

Thus, the goal of this thesis work is to build a framework for creating interactive synthetic characters that incorporates the motivation system in such a way that it functions as an internal and external communication medium and as reinforcement for learning, which are crucial aspects of building intentional and adaptive synthetic characters, and, finally, a basis for shaping personalities and imposing diversity on characters, all as emergent phenomena.

2.2.2 Groundwork

There is no fuzzier terminology than in the areas of motivation and emotion. They are common words that we always use in everyday conversations, but there are no clear definitions of those words that everyone agrees with. So, the first effort should make clear the boundaries of those concepts that I will try to implement in the system. Functional definitions need to be made to explain necessary concepts.

Blumberg [7] developed a behavior engine based on the theory of hierarchical structure of behavioral organization [126, 68], which is widely accepted by ethologists. Efficiency and plausibility of the framework has been demonstrated by various interactive installations such as ALIVE [71] and Swamped! [9]. In my work, I extend that model to utilize a probabilistic framework, designing creatures based on a more explicit information flow model, and I have enabled various kinds of animal-like learning to make creatures truly adaptive. Most of the codes for this new model have been written in Java. The developed creature kernel model has been implemented in the form of various characters in various settings.

Thus, what I demonstrate through this thesis is a computational framework based on the motivation model, which can be connected to other parts of a creature's kernel, and implementations of synthetic characters in various interaction settings.

2.3 Development

The motivation system cannot be properly implemented by merely adding *ad hoc* adjustments of the look or reaction of the creature for each situation, or as a set of rules which tells the creature the way it *should* feel when a certain event happens. Instead, only a systematic and functional implementation of it can actually aid a designer developing a situated animat creature, and enable the creature to be truly adaptive and believable in the situation that it lives in.

2.3.1 A rough sketch

In this thesis work, the motivation system is modeled as a sum of the drive system and the affect system. The drive system is composed of the drives due to the internal state of the creature and the drives induced by its sensory inputs. The affect system is composed of affects associated both of these types of drives, and affects associated with appetitive and consummatory behaviors and with social signals. The number of basic components of drives and affect is specified to meet creature specific characteristics. The behavior system, which also is a part of the creature kernel, affects the motivation system both at its input and output ends. Performing consumatory behaviors satisfies associated drives, which, in turn, influence the affect system. Appetitive behaviors, on the other hand, influence the affect system directly; for example, just approaching an apple makes a raccoon happy even if it has not eaten it quite yet and thus has not lowered its hunger drive. On the output side, the change in affect is shown as various facial expressions and appropriate gestures.

2.4 Demonstration

Two projects were worked on to provide the necessary environments for implementing the motivation model: **Sydney K9.0** and **(void*): A cast of characters**.

Sydney K9.0 is a project that focuses on the motivation system's role as the reinforcement for learning. Sydney is the name of the dog in the project, and the setting provided is an interaction between a human participant and Sydney through speech and other auxiliary interfaces. Various types of animal learning, especially those discussed in Lorenz [68] are explored using the creature kernel framework described in this thesis.

(void*): A cast of characters is a project that focuses more on the communication medium aspect of the motivation system. Three characters with distinguishable personalities are introduced in this project. The project touches not only the issue of interaction between characters and human participants but interactions among characters themselves. Personality is the main factor that shapes the characters' reactions to participants' inputs and, in turn, their attitudes toward the interaction itself.

2.5 Evaluation

Evaluation of how well the proposed method works follows the description of the (void*) project.

Adaptive characters should not fail at dealing with novel situations that are generated by inputs from human participants and through interactions with other characters who have their own intentions. While continuing the interaction with the characters, human participants should feel comfortable thinking they understand what is going on and what characters are trying to convey, instead of feeling obstructed by an opaqueness of the characters' minds.

So, centering on the perceived believability of characters' behavior and apparent ease of communication from the human participants' viewpoint, survey questions were answered by volunteers and the results were analyzed to evaluate the plausibility of the proposed approach and, thus, the implemented system and algorithm.

To summarize, I have aimed at achieving the following three goals by incorporating the motivation system into synthetic characters through this thesis work:

1. Realization of synthetic characters that look more believable than heretofore, in other words, characters that give more lifelike impressions through their emotional expressions and reactions to events and human participants' inputs.

2. Design and implementation of synthetic characters that are adaptive even in situations that were not anticipated at the designing stage.

3. Provision of a framework which is easy for designers to use for implementing believable synthetic characters. Without a proper motivation system framework, emotional responses of characters have to be designed and added by hand for every little situation. This previous approach makes the job very difficult for the designers or programmers, because as the world becomes more complex, the number of possible situations grows exponentially, every one of which has to be carefully considered to evoke believable responses from characters. This older approach is not what happens in nature. I have implemented in the software of synthetic characters a theory of how real animals work.

Chapter 3

Glossary

Terminologies used in this thesis follow the conventions of the Synthetic Characters Group, MIT Media Laboratory, where all the implementation work described here was done. Below I describe the meanings of the key terminologies that appear frequently in descriptions of the system, to avoid confusion that might arise for readers outside of the group.

Virtual world refers to an environment we create as a software program and display in 3D graphics. It allows constant external inputs, which add unpredictability to the state of the world. All the projects described in this thesis display the world on a flat, vertical screen. However, it does not need to be restricted to a flat screen, and any other media which our creatures are allowed to inhabit meet the working definition of virtual world. Virtual world is a subset of **world** which refers to both virtual world and physical world. Unlike virtual world, **physical world** is filled with tangible objects and inhabited by robots, human beings and other animals. **Creatures** are inhabitants of a world, which include both real animals and software creatures. The common denominator of those two categories of creatures is that they both have perception, motivation, behavior and motor systems and live in an unpredictable environment. **Characters**, or synthetic characters, are creatures designed according to the behavior-based creature design framework [7]. They might be embodied physically or graphically so that human participants can perceive them or they might not

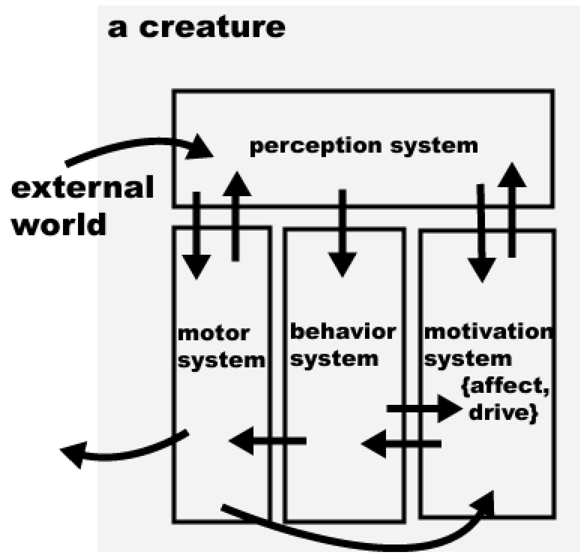


Figure 3-1: This figure is a schematic diagram of a situated creature, which is composed of the four main parts: perception, motivation, behavior and motor systems. Arrows represent information flows among those systems. Well coordinated communication among those four is required for a creature to successfully function in a dynamic world. See text for details.

be embodied when it is not necessary, as in the cases of music and camera creatures. Characters have distinct names which are used in their world, and personalities that can be perceived by other characters in the world and by human participants who are interacting with or watching them.

Creature Kernel is the integration of the four main elements – perception, motivation, behavior and motor systems – which, we think, are crucial for the operation of a creature in an unpredictable world. A schematic diagram of the creature kernel and how each part is connected to form a situated creature is shown in Figure 3-1. Introducing each concept briefly, the **perception system** refers to a system of sensors that are used to extract the state of the world. For animals in a physical world, the perception system includes visual, auditory and tactile sensors, whereas for virtual creatures, it refers to a set of visual sensors, which return the visual features of other objects in the same virtual world, smell sensors, which return the olfactory features of objects that they can smell, etc. The **motivation system** is the sum of drives and affect. Depending on personalities and what kind of species they are, characters have different sets of drives with different parameters. Affect refers to the state of feelings, often referred to as emotion, mood, etc. The **behavior system** is the control center of a creature's behavior. By putting together inputs from various parts of the creature kernel, action selection is made, and a command to the motor system is sent out that follows the decision made by the action selection. The **motor system** controls all movement of the physical body such as limbs and joints in the case of real animals, or the mechanical parts for robots and the graphical representation on the screen for virtual creatures.

Emotional expression is revealing of affect through the creatures' motor system, i.e., facial expressions, various gestures, etc. It enables and aids the communication between creatures and human participants, as well as that among creatures themselves. Human participants are users of our installation, who affect synthetic characters or the state of the virtual world through various input devices.

Part II

DEVELOPMENT OF THE SYSTEM

This thesis work is aiming at creating a whole graphical creature through which different parts of the underlying system – perception, motivation, behavior and motor skills and their realizations through the graphical 'body' – all interact in a plausible manner. In particular this work is directed at building creatures that are adaptive to an unpredictable world, and thus who are able to interact with human participants, whose exact input never can be predicted in advance. We call such creatures interactive synthetic creatures.

In Part II, I describe how the concept and framework were developed and implemented for supporting creation of such creatures.

Figure 3-2: K. F. Chicken, one of two main actors of SWAMPED! [9] in its virtual world.

Chapter 4

About building synthetic characters

The work of implementing a synthetic character is premised upon a broad understanding of a synthetic character as a situated creature in a dynamic and unpredicable world, as well as upon a deep understanding at the level of detailed operations.

A synthetic character is conceptualized as a situated creature in a world, and its existence and operation have their meanings not in isolation but through the character's action and reaction to the world, which includes not only various objects but also other creatures in the same world. Figure 4-1 summarizes this concept of a synthetic characater as a situated creature.

In its most abstract form, this idea of a situated system is rooted in the theory of classical control systems. The fundamental notion in classical control theory is that of a dynamic system. Such a system consists of a controller and an environment. Both the controller and the environment to be controlled are viewed as deterministic finite automata. The input of the controller is a signal which is the output of the environment; on the other hand, the input to the environment is an action provided by the controller [29]. Likewise, the operation of a synthetic character can be viewed as a stream of interactions between the creature and the environment. As the

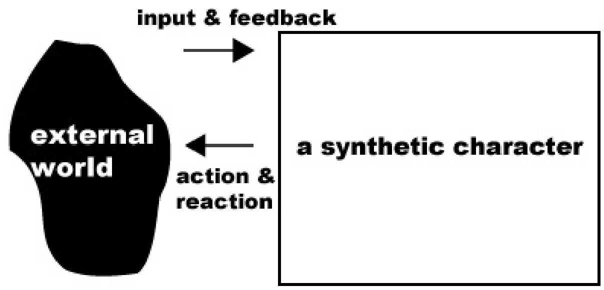


Figure 4-1: A synthetic character is modeled as a situated creature in the world and its operation is a series of actions and reactions to the dynamic and unpredictable world that surrounds the character.

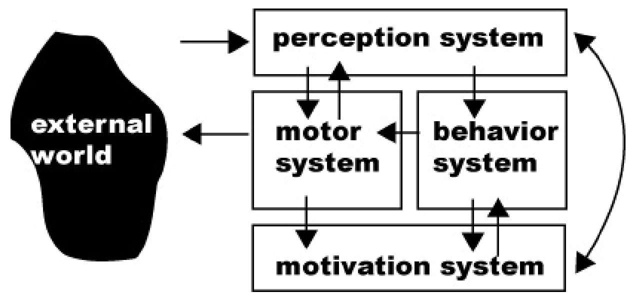
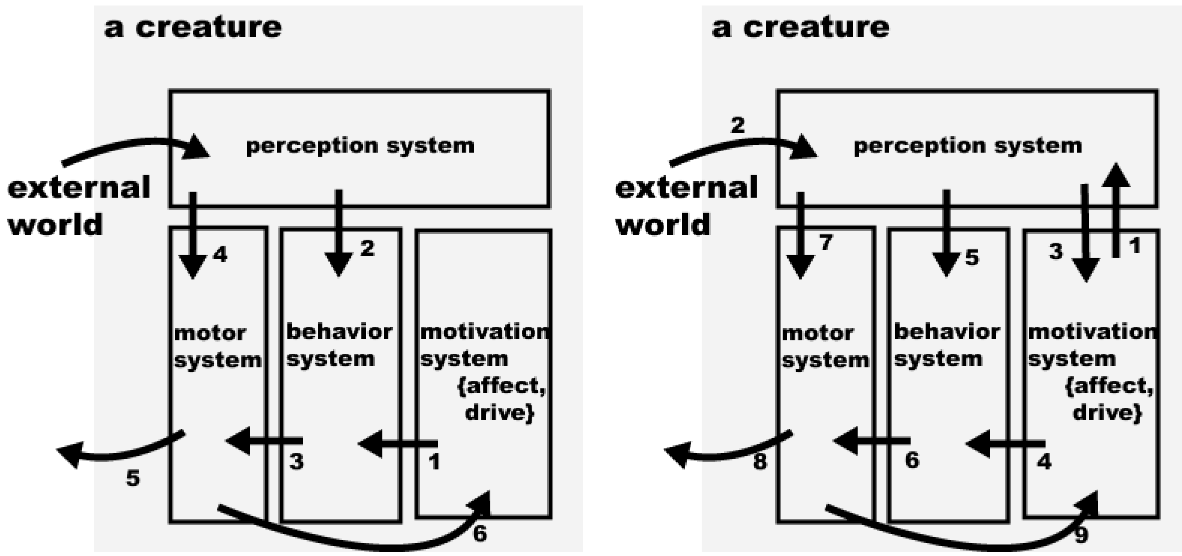


Figure 4-2: A synthetic character is modeled as a sum of four main components – perception, motivation, behavior and motor systems. Well-coordinated functioning of those four is essential for a character to survive in an unpredictable world.

task of a controller is to achieve a reaction in the environment, a situated creature seeks ways of managing its life to satisfy its goals and thereby survive, overcoming the challenges provided by the unpredictable surrounding world. Thus, the system function represents a set of correct actions to be executed for each possible perception.

Getting to the system-level description of a character that should behave in this manner, each character can be understood as a composition of four main components – perception, motivation, behavior and motor systems [49], which form the kernel of a creature. To function in an unpredictable world, these four main components of a creature should coordinate well as a whole, as shown in Figure 4-2.

Given this structure, a creature's actions and reactions or, more generally, its behaviors, arise either internally or externally. Internally aroused behaviors include be-

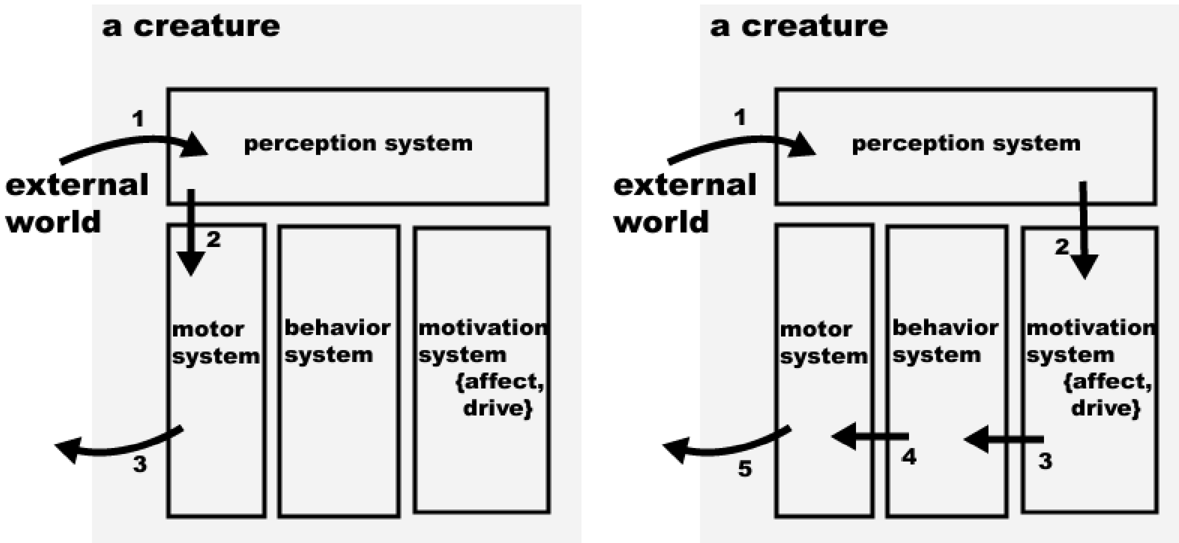


a b

Figure 4-3: These two figures show examples of information flow within a creature when it performs internally originated behavior. **a:** Need of certain behavior arises from motivation system and is signaled to the behavior system (arrow 1); combining with sensory input about the state of the external world (arrow 2), the behavior system selects a certain motor subsystem to be executed to satisfy the motivation system (arrow 3). The motor system also gets input from the perception system (arrow 4) and executes a motor command that it has been told to do (arrow 5). This execution is fed back to the motivation system (arrow 6) and serves as one of the bases for deciding whether the same cycle should be repeated or another drive should be taken care of. **b:** This is another example of information flow which explains the case when internally originated behavior is performed. Motivational input affects the perception system (arrow 1) which has the creature focus on certain objects (arrow 2) with which the creature can satisfy its motivational need. The finding of such an object then motivates the creature (arrow 3), for example, to *approach* the object, and the rest of the scenario is similar to that of **a**.

haviors for satisfying certain drives – for example, arousal of the curiosity drive makes a creature explore novel places, and arousal of maternal or paternal drives leads to behaviors such as child care by parents. These motivational inputs are fed into the creature's behavior system to activate relevant behavior modules, which in turn send execution commands to appropriate parts of the motor system as shown in Figure 4-3.

External stimuli may trigger behaviors as well. Perceptual input might go directly into the motor system to trigger certain reflex responses, or it may cause the motivation system to increase certain drives as shown in Figure 4-4. Except for purely reflex cases, the motivation system functions as a medium, i.e., a common currency,



a b

Figure 4-4: These are examples of information flow in the case of externally originated behaviors. **a:** Perceptual input (arrow 1)can be directly fed into motor system (arrow 2) to trigger a reflex like motor skill (arrow 3). **b:** Or the external state of the world (arrow 1) - such as the presence of an appetite raising object, which will affect hunger drive, or the presence of a predator, which will raise the drive to flee - is reported to the motivation system (arrow 2), which then influences the behavior system (arrow 3) to select an appropriate behavior to tell the motor system (arrow 4) to take an action (arrow 5).

that facilitates communication among the other systems [23, 10]. These motivational inputs are fed into the creature's behavior system to activate relevant parts, which in turn send execution commands to appropriate parts of the motor sytem [80] as shown in Figure 4-4. The motivation system talks to the behavior system to provide bias in action selection. Appropriate behavior is selected based on the creature's current needs combined with other factors. Signals from the motivation system bias the signals that are sent to the motor system, modulating the manner of execution of the creature's motor skill. For example, even if the creature is performing a walk skill, it could walk happily or sadly depending on its current mood. In turn, the result of the action selection in the behavior system and thus execution of the associated motor skills, lowers or raises various drives, or affects emotional states positively or negatively, depending on how they are connected to each other. The perception system may directly talk to the motivation system as well. Finding an apple might suddenly raise the creature's hunger drive. On the other hand, a creature would be more likely

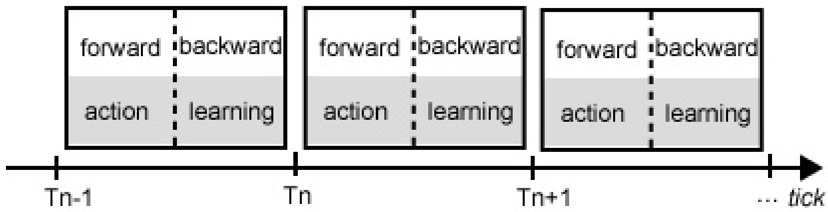


Figure 4-5: A character's operation is modeled and analyzed as alternating action and learning phases at every tick. These phases are referred to as forward and backward operations, respectively. Forward process includes evaluation of the state of the world and the character itself and making the right behavioral decision, and backward process refers to feedback from the result of the action, and updating of the character's world model, or *learning.*

to find foods when it is hungry than when it is not.

4.1 Forward and backward operations

As a character that has the ability to *learn*, the operation of a character can be expressed as a repetition of two different modes of operations – forward operation and backward operation.

The process I just described in the previous section summarizes the forward process of a situated character. It consists of determining the current state of the environment by evaluating the current input to the character, and then determining a suitable action strategy given the state of the environment. After performing the chosen strategy, given the output or response that resulted due to the character's action, the goodness or plausibility of the strategy, and thus the plausibility of the *character* as an estimator, is evaluated. Update of the evaluator follows, to increase the probability of its working better or using a better strategy the next time. This post-action system update process, i.e. learning, is referred to as backward operation of characters. This way of seeing a character's operation is summarized in Figure 4-5.

To summarize, no matter whether we think about characters for synthesis or analysis, we must include both the character itself and the environment, both how the character is structured and how it operates, and both how its forward operation is performed and backward operation is realized. No matter how local the focus of work on a synthetic character may be, all these action and reaction relations should be kept in mind. This thesis work is based on this framework and mindset.

Chapter 5

Creature Kernel

As discussed in Chapter 4, the creature kernel is composed of four subsystems – the perception, motivation, behavior and motor systems. In this chapter, I explain in more detail how each of these four systems are similar and different from each other and how each is implemented.

All four systems are modeled as a network of basis units. For example, the perception system is modeled as a hierarchically connected network of sensor units. Each unit represents a species-specific sensor such as a vision sensor, smell sensor, etc. and some units are structured for extracting compound information such as distance to a certain object. The number of levels in each network and the way units are connected are decided both by the designer and by learning processes, i.e., both by phylogeny and ontogeny of the creature. Units in a network interact with each other to excite or inhibit other units' activation in a manner similar to what was implemented by Blumberg [7]. This influence is modulated by the strength of the connection, which ranges between 0 and 1, as well as by its own activation level.

5.1 The behavior system

The behavior system is composed of behavior units. The output value of each behavior unit is a function of the prior output value of itself, output values of its parent

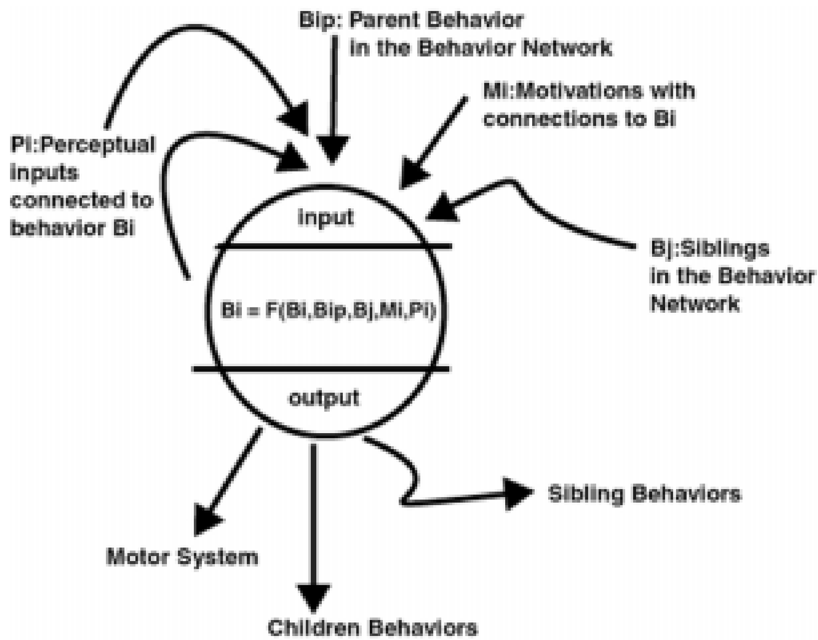


Figure 5-1: One behavior unit. Arrows represent connections to other units in the behavior system, or other systems in the creature kernel. Output value of this unit is a function of the output value of itself in the near past as well as inputs through those connections.

behavior unit and sibling behavior units (*via* mutual inhibition), and inputs from other subsystems i.e., the motivation and the perception systems. Figure 5-1 is a schematic diagram of one behavior unit where arrows represent connections to other units within a behavior network. Units that inherit output values of the units that belong to higher level of the network are called *children* units. Units that influence *children* units are called, in turn, *parent* units.

5.1.1 Behavior network

Behavior units are organized in a semi-hierarchical manner to form a behavior network. Figure 5-2 shows a portion of a possible behavior network as proposed by Tinbergen [126]. Arrows between behavior units are either excitatory or inhibitory connections. For simplicity, Figure 5-2 only shows input streams to behavior units from other parts of the behavior network, omitting influences from other subsystems of the creature kernel.

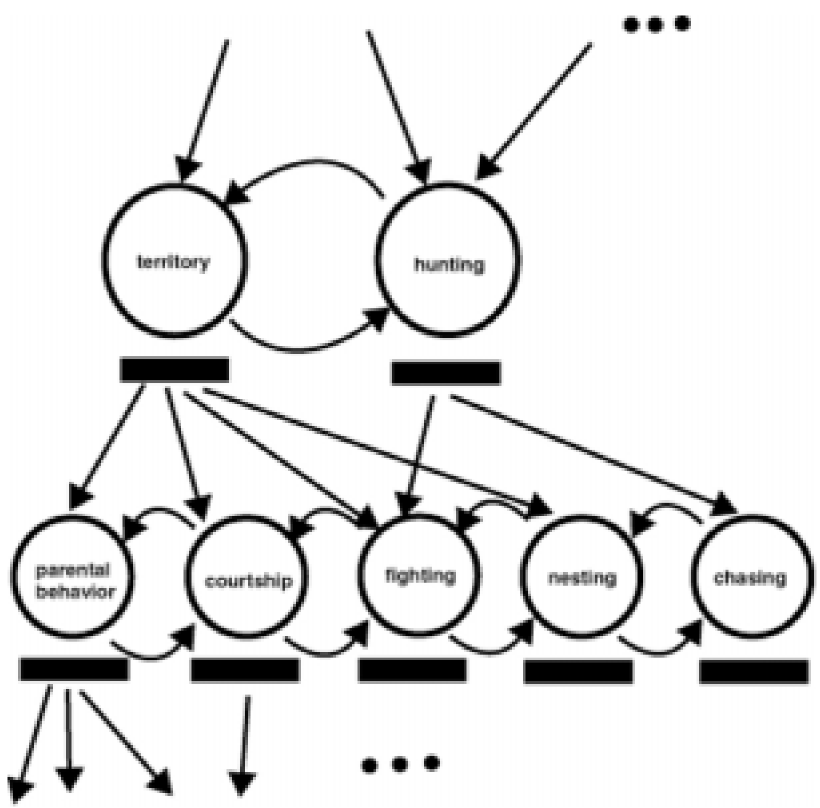


Figure 5-2: Hierarchical organization of behavior units as proposed by Tinbergen [126]. Vertical arrows represent influences from a parent behavior to its children behaviors and horizontal arrows represent mutual inhibitory connections among sibling behaviors.

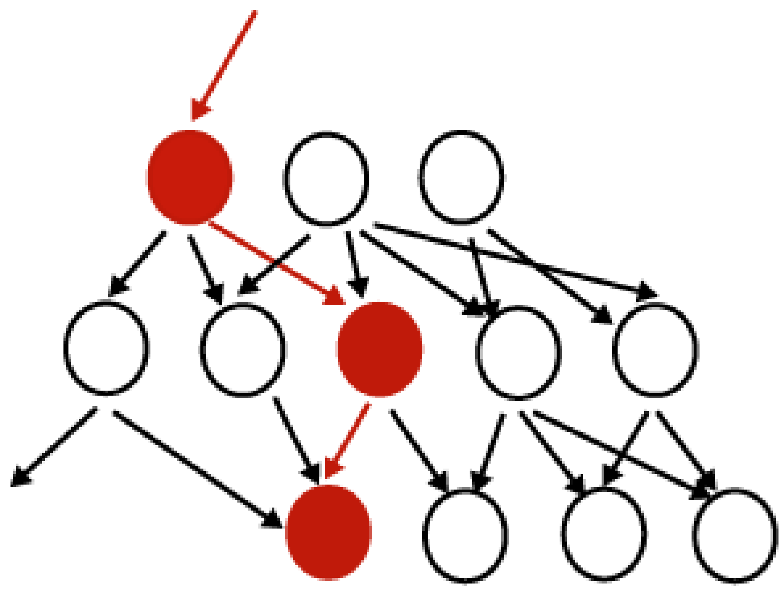


Figure 5-3: This figure shows how a behavior sequence may be executed within a behavior network. Filled circles and arrows show a possible activation of a behavior sequence in a behavior network.

5.1.2 Execution of a behavior sequence

A behavior sequence is executed through a chain of activation of behavior units, shown as connected by directional arrows in the figures. From the node where the sequence started, behavior units activate in an order following the direction of the arrows until the sequence terminates – i.e., for example, when it sends out an appropriate motor signal for bodily execution. One example sequence of execution is shown in Figure 5-3. Both the weight of the connections and how the behavior units are connected are important. It is assumed that every creature has a species-specific *a priori* structure of its behavior network and associated connection weights that it was born with. Both the presence and absence of connections and weights associated with the connections are flexible so that they can change over time for the creature to become more adaptive to the specific environment, and so it can learn new skills or strategies for solving problems it might face. Or the structure of the behavior network could just reflect personal habits that a particular individual has developed through its lifetime. I propose that these modifications in the behavior network are the result of learning; the way this learning has been implemented is discussed in Chapter 7.

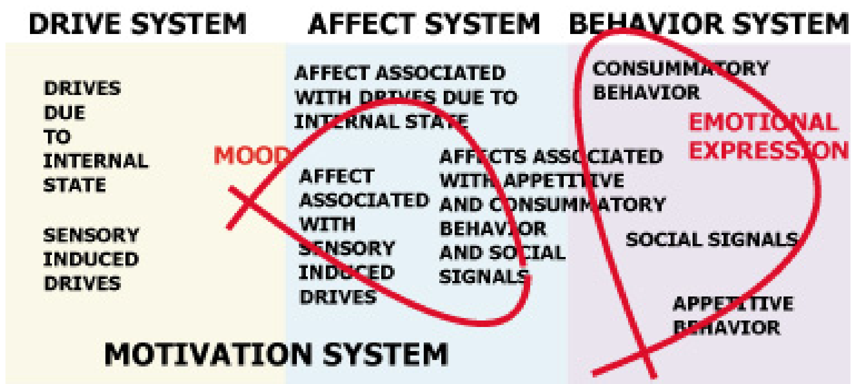


Figure 5-4: Components of the motivation system. The motivation system is composed of the drive system and the affect system, and each consists of corresponding subparts. Parts of the behavior system which have intimate connection to the motivation system are shown in this figure as well.

5.2 The motivation system

The motivation system is composed of two parts: the drive system and the affect system. The drive system includes those depending on internal states and sensory induced drives. It is organized as a semi-hierarchical network like the behavior system, and the connections that connect basis units are also flexible, i.e. the connection weights may be modified and updated as learning of the synthetic character proceeds. Figure 5-4 summarizes the components that all together form the motivation system. I include part of the behavior system in the figure, since there is a tight coupling between the behavior system and the motivation system which, among other things, is responsible for *expressing* the state of the motivation system outwardly through gestures and facial expressions.

5.2.1 The drive system

Consummatory behaviors and appetitive behaviors [68] have intimate back and forth connections with the drive system in particular, and the creature's affective state is expressed through its various emotional expressions which are built in (innate) or are shaped to fit cultural conventions. Every synthetic character is *born* with species specific drives that are connected in the *a priori* way, and over time, the connections between the drive units are modified as the character interacts with the world and gains experience. The connections and their strengths reflect the individual's specific interests and desires that are perceived as personality, and as characteristics which contribute to the individual's behavioral strategies and habits.

5.2.2 The affect system

Components of the affect system include affect associated with drives due to the creature's internal states, affect associated with appetitive behaviors, and affect associated with sensory-induced drives. Top level nodes in the motivation network represent so called *mood* which provides underlying bias to the character's behavior.

Approaching affect

> *When I use a word, Humpty Dumpty said in a rather scornful tone, 'It means just what I choose it to mean - neither more nor less.' 'The question is,' said Alice, 'whether you can make words mean so many different things.' 'The question is,' said Humpty Dumpty, 'which is to be master - that's all.' Lewis Carroll: Through the Looking Glass*

One of the challenges that we are faced with when we start modeling and implementing the affect system is that we need to be clear on what we mean by *affect.*

In antiquity, Chinese doctors believed that humans experience four basic affective states – anger, happiness, sorrow, and fear, which were thought to arise from activities of the liver, heart, lungs and kidneys. The ancient Greeks also viewed human nature as being constructed of four basic elements – fire, air, earth and water – and the classic temperaments - choleric, sanguine, melancholic and phlegmatic – are reminiscent of the psychobehavioral control systems that appear to be emerging from brain research as distinguishable emotive circuits [89].

Most modern taxonomies of human emotions, in line with earlier perspectives [31], posit the existence of a fairly small set of basic passions. For instance, Tomkins [131, 130] argued that joy, fear, rage, surprise, disgust, anguish, interest, and shame are the basic human emotions. Izard [54, 53] concurred, but also added contempt. Plutchik's [92, 93] list is similar to that of Tomkins's, but instead of shame, acceptance was put forward. Although all of these are distinguishable affects, and most can be recognized by facial countenance [35, 53, 55], it is doubtful that all should be considered to be of equal importance (i.e., reflecting robust, cross-species types of brain organization).

Despite all this effort, there has never been any clear agreement as to what the word means. Amongst philosophers, affect has almost always played an inferior role, often as an antagonist to logic and reason. Along with this general demeaning of affect in philosophy comes either a wholesale neglect or at least a certain distortion in the analysis of emotions [111, 69]. The lack of acceptable theory of emotion does not stem from a paucity of candidates [117]. Many books have been written presenting this or that theory of emotion. However, most have a selective view of observational and experimental evidence, often omitting whole areas of research that another theorist would deem pertinent. Both this disparity of viewpoint between theorists and the logical confusion which can be found within some theories are traceable to variations in the meaning of the word emotion [82].

One could argue that such verbal responses merely reflect cultural conditioning, and that all humans readily recognize the outward reflections of various emotions in others [35, 55]. At present, the more fertile starting assumption is that the verbal responses reflect a pre-linguistic, affective heritage, of distinct brain mechanisms passed down to us from our evolutionary past. These emotions appear to be represented coherently at a fundamental neural level. Although one could debate endlessly on how such systems should be verbally labeled, the essential issue is that concrete brain mechanisms for several emotions do exist, and scientific understanding of emotions can only be achieved through a direct study of these systems. Thus, although it may be impossible to convince an insistent skeptic of the proposition that the brain contains a separate 'panic system' [89], the essential issue is that brain circuits that elaborate separation distress do exist.

Therefore, the approach taken in this thesis is, instead of attempting to start from an inductive definition of affect and emotion, which has proven to be very difficult to work out, working with a deductive definition of affect which can be prescribed collectively through examples. In particular, the aspects of affect that are considered to be pertinent in terms of operation of the creature kernel will be presented in terms of their functionality (see Chapter 6), and the modeling and implementation of affect are designed to cover all those aspects important for functioning as a part of the creature kernel.

Three-axis model of affect

In addition to keeping the previously claimed notion of developing a model of the affect system that can support the functional need for the proper operation of the creature kernel, the design constraints I adopted include flexibility and communicatability.

- Flexibility: By being flexible, I mean the model's ability to cover a wide range of affective space, including that of a primitive animal that does not have a very complicated motivation system, and a more developed creature that possesses intricate and complex motivations and thus affects.

- Communicatability: A fundamental underlying requirement of this creature kernel framework is the ability of all four subsystems' to work together to make the creature a functioning whole. This should be supported by the subsystems' ability to communicate with each other.

Like all other subsystems in the creature kernel, the affect system is organized as a semihierarchical network structure. The basis affect units are connected via directional arrows, whose values represent levels of influence. The top level nodes of this network are three basis units – valence, stance and intensity – whose values vary along their respective axes [101, 12].

Putting the components of affect on axes make sense because the main descriptors of those components are signed values. A creature can experience both positive and negative states of one kind of affect (for example, happiness and unhappiness), which can be clearly represented as values along an axis that spans negative and positive values. Each state of one particular affect can be represented as a point on this axis, i.e. a creature can feel happy a lot or a little, or very unhappy or only a little unhappy, depending on the situation. This property of affect is essential for later development and especially for integration with learning which needs the guidance of signed quantities where values can be compared.

Intuitively, ths most basic attribute of affect is represented along the *feeling good – feeling bad* axis. A creature might feel good in anticipation of something good coming as a result of the behavior, or the behavior itself may be pleasing and rewarding, or the creature might feel bad due to the opposite reasons. This attribute of feeling good or bad varies along one of the three main axes and is called 'valence.' Stance corresponds to affective situations which lead to either approach or withdrawal responses in Schneirla's sense [73]. This also assumes positive and negative values, but is more concerned with a pronome [83] that the creature is reacting to or interacting with [1]. Stance is affect toward an object, another creature or situation that the creature is experiencing. The third, and the last, axis represents intensity. The need of a new dimension comes from temporal differences (for example, to express different rates of change) as well as qualitative differences in the expression types. In the case of human beings where six primary emotions can be considered as the primitives of affect [35], this axis aligns well with the affective attribute called 'surprise.' Surprise is often reported in conjunction with happiness, anger and distress and this observa-

[1] The difference between valence and stance can be viewed as similar to the comparison between intransitive verbs and transitive verbs. The basic either positive or negative affect that a creature experiences when the behavior can be described using an intransitive verb, would be closer to valence. Otherwise, it would be more proper to use stance, within the framework I am developing here.

tion supports the notion of separating this process from others [116]. When surprise occurs in close proximity to other aspects of affective arousal, surprise normally precedes others. As shown in the case of surprise and curiosity, affective response along the intensity axis is not really related to the attainment of a goal or bodily feedback and it precedes other evaluation-related affective responses, whereas good/bad and approach/withdrawal axes are more related to the maintenance of valued goals or perception of the relative position of self to a goal to be reached. Surprise, curiosity, and interest (or disbelief, which evokes a reaction similar to that of novelty – a conflict between expectation and perception) are indicative of the formation of new representations, whereas happiness, anger, sadness and fear are indicative of those emotions that occur when an event is being assessed with respect to the maintenance of valued goals.

In terms of the relationship with learning and behavior, while stance and valence axes contribute to a character's learning by providing evaluative quantities, i.e. serve as guidance for future behavior and belief, the intensity components contribute through absolute values as well as by determining the focus of attention. This issue is discussed in more detail in Chapter 7.

This affect system network, which starts with three basis nodes - valence, stance and intensity, can be expanded to arbitrary depth to represent complex affective states, or can be kept near the basis units for representing states in a simpler affect system. For example, Russell [101] succeeded at mapping out the six primary emotional states [35] on the plane spanned by these three basis axes. Within the proposed framework, this result can be implemented as a network of two layers with three nodes on the first layer and six nodes on the second layer that are connected to the first layer via arrows from the first layer with various weight combinations.

The motivation system is at the center of forward and backward operations

As explained in Chapter 4, the operation of a character is viewed as a chain of two phases of operations, forward and backward, which include evaluation of the world and making strategic decisions and performing actions, and of getting feedback and updating its world models or beliefs, respectively.

In particular, with the motivation system in mind, the operation of a character is viewed as follows.

During the forward operation phase, after evaluation of the current world or event, a behavioral decision which will bring the most positive valence or stance response as the result of the action, is made. The values of the affect system modulate the way the chosen behavior is performed. For example, even if the character made the same decision to 'walk', it may walk happily or sadly based on the current state of the affect system or anticipated outcome.

During the backward operation phase, feedback from the motivation system determines the goodness of the chosen action. In particular, the realization that the immediate goal of an action has not been attained or that the goal is not going to be attained in the near future, or that more than what has been expected will happen, violates expectation, i.e., there is a realization that something unusual or unexpected has occurred. Attention is then focused on two different dimensions. First, an assessment is made as to whether or not the event was encoded properly, and second an appraisal is made about the relative certainty that a particular goal has or has not been attained [115]. This is fed back to the creature as affective component values, which updates (adjusts) the decision making system so that the creature can predict or plan better next time in a similar situation. Positive valence and stance values support the recent decision whereas negative outcomes discourage repetition of the same strategic choice in the future. This lesson is reflected in the weight updates, or *learning* in the behavior and other creature kernel networks.

5.3 The perception system

The perception system is organized as a network of basis perceptual units. The kinds of perception units that a character possesses vary according to the type of species that the character belongs to. For example, the virtually omniscient camera creature has smell sensors that can detect almost all the creatures in the same virtual world if they have names. The raccoon in the Swamped! world [9] was dependent on visual sensors when it needed to detect K. F. Chicken, so it would not be able to realize the existence of Chicken if it was behind the wall or some other kind of obstruction. See the paper by Kline and Blumberg [62] for details on the sensors' software side implementations.

Like those in the behavior system, the connection weights in the perception system network are also subject to modification and update during the backward operation phase of the character. Certain connections within the network are strengthened to reflect the character's personal preference and interest, or the connection between a certain perception unit and a behavior unit is tightened to emphasize the contextual importance of the sensor output when the character performs that behavior.

5.4 The motor system

The underlying structure of the motor system is the same as that of other subsystems and is subject to same type of operation and thus learning processes. However, current implementation of the motor system maintains a more-or-less flat hierarchy and individual nodes correspond to the leaf nodes, or terminal branches of the behavior network. This flat motor system functions as the mediator between the behavior system and the graphics output layer. It receives a signal that indicates which graphics skill should be activated at the next tick and sends it to the graphics layer and has it play the appropriate animation sequence.

Chapter 6

The motivation system

Chapter 5 describes the creature kernel framework which is composed of four subsystems. Among those four subsystems, this chapter focuses on the motivation system. I list the roles of the motivation system that are crucial for proper functioning of a character and present the mathematical framework to show how it is actually implemented within the creature kernel framework to achieve the required functions.

No single function can explain the functional importance of the motivation system and there have been a number of claims and studies about what the functions of the motivation system are. The usual approach to the functional studies of drives and affects assumes them to be important chiefly as driving "forces"–giving direction and intensity to behavior – or as internally arising stimulation. In addition, one of the major functions of affective responses, which has been repeatedly discovered, is that they act as internal reinforcing mechanisms, and thus contribute to learned motivation of either a positive or negative nature. In another paper, Scott [105] pointed out that an internal emotional response may also magnify and prolong the results of external stimulation.

Among the functions of the motivation system which this thesis is particularly aimed at understanding and implementing in software are the following three: to serve as the internal and external communication medium, to provide the reinforcement signal for learning and to serve as the basis for shaping personality, thus to bring diversity to the society of creatures and synthetic characters.

6.1 Acts as the communication medium

The first function of the motivation system is to serve as a communication channel. A collection of *objects* that together exist as a pool or society of such objects, requires some means of communication if they need to coordinate in some way with each other. These *objects* could be a group of people in a society or a bunch of cells that together form a tissue.

In the case of synthetic characters, two kinds of communication mechanisms are needed for full functionality – internal and external communications. Internally, the creature kernel of a character is composed of four subsystems: the perception, motivation, behavior and motor systems. Since the coordination of those four subcomponents is essential for the survival of the character, communication among the four systems is required.

On the other hand, externally, as an *object* that interacts with other objects, i.e. other characters in the same environment or society, each character needs a means of communicating with other characters including human participants in character-human interaction settings. We human beings face a similar challange as a member of a society where more than one person resides. We have developed various communication channels such as language and gesture. In the case of animals that do not have language, bodily expressions of emotions and intentions including facial expressions are major ways of achieving direct communications [28].

As in animals, the motivation system serves as the means of both types of communications in the case of synthetic characters that do not have language.

6.1.1 Internal communications

For internal communication, the motivation system functions as the common currency that enables the exchange of necessary information among the four subsystems. As explained in Chapter 4, behaviors of a character originate either internally or externally and in both cases, the driving force is spread over multiple subsystems of the creature kernel to instigate and guide that needed behavior. The motivation system provides the common currency which enables this propagation of information over the boundaries of those subsystems.

6.1.2 External communications

Needs, intentions and affective states are the crucial contents that need to be communicated among individuals of a society. Animals have developed various mechanisms for expressing affective states such as gestures and facial expressions or affectively modulated execution of behaviors. For example, a dog walks happily or sadly based on its affective state. For species where language is not available for the means of communication this tight connection between the motor system and the motivation system provides the needed means of communication. Special effort is needed to keep the true emotional state from appearing on one's face.

A monkey, for example, may communicate its affective state to others by making an open-mouth threat to indicate the extent to which it is willing to compete for resources, and this may influence the behavior of other animals. This aspect of affect was emphasized by Darwin [28], and has been studied more recently by Ekman [35, 36]. He reviews evidence that humans can categorize facial expressions into the categories happy, sad, fearful, angry, surprised and disgusted, and that this categorization may operate similarly in different cultures. He also describes how the facial muscles produce different expressions. Further investigations of the degree of cross-cultural universality of facial expression, its development in infancy, and its role in social behavior are described by Izard [56], Fridhund [39] and Eibl-Eibesfeldt [33]. As described below, there are neural systems in the human forebrain – in the amygdala and overlying temporal cortical visual areas – which are specialized for the face-related aspects of this processing.

Not only advertising a creature's affective state at that moment, emotional expressions also serve as the warning or the encouraging signals which accelerate group learning [34, 28]. This mechanism is thought to play a very important role in making a group more adaptive.

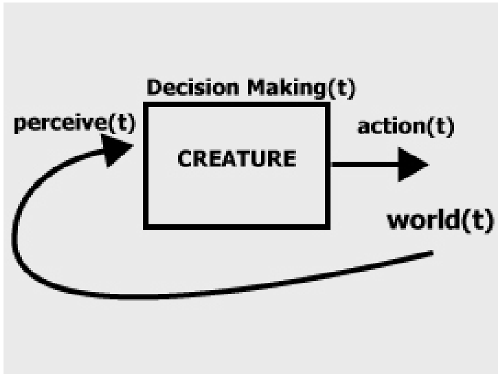
6.2 Provides the reinforcement signal for learning

By providing the reinforcement signal for learning, the motivation system plays an essential role in making a creature be adaptive to an unpredictable environment. Guided by its innate affects and drives, a creature behaves in a way that satisfies its needs. Drives, and Affect with its associated 'affective tags', bias a creature to behave in an adaptive manner by providing *values* to measure the level of goodness of its behavior [112] at a given instance.

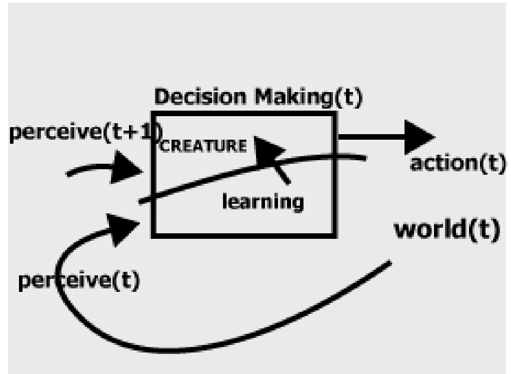
Two-factor learning theory [51] has been nicely generalized to include affect as reward or punishment for animal learning by Young [143]. In his paper, Young points out that affective processes are positive or negative in sign. Inferring its behavior as the cause or source of the affective process, a creature uses that internal affective response as a factor that sustains or terminates the aroused behavior pattern. The motivation system includes both the affect system and the drive system. Success (Failure) at performing consumatory behaviors, which would lead to reduction (increase) in certain drives or satisfaction (unsatisfaction) in certain motivations, or experience of happiness or joy (sadness or fear) function as positive(negative) reinforcers for a creature's learning, which in turn cause the creature to be more(less) likely to trigger the behavior pattern just aroused [85].

A distinguishing characteristic of experiences that involves affect is an effort to assimilate some type of new information into current knowledge schemes [74]. Stein and Levine [114] concluded from an experiment that people constantly monitor their environment in an effort to maintain preferred states. To succeed at this task, pattern-matching procedures are used to analyze and compare incoming data to what is already known. They concluded from their experiment that when new information is detected in the input, a mismatch occurs, causing an interruption in the current thinking process. Attention then shifts to the novel or discrepant information. With the attentional shifts comes arousal of the autonomic nervous system and a focus on the implications the new information has for the maintenance of valued goals. Thus, affective experience is almost always associated with attending to and making sense out of new information. This conclusion draws upon a very narrow definition of affect different from that used in this thesis. The usage of the term here does not always assume that an affective process is related to a novel event. But it is true that even the thesis definition of affective arousal includes cases where mismatch occurs between the perceived state of the world or self and the expected state, and the affective arousal that occurs forces the creature to learn from this new information.

Each subsytem operates in the form of alternations between two modes: forward mode and backward mode. Forward operation refers to the process of each node's getting inputs from various other nodes within and outside of the system. This results in each part's operations, leading to outputs which influence (1) the character's deciding to take certain actions, and (2) at the sane time, the updating of each system to the next state. External and internal reactions to past actions are fed back to the creature and this becomes the bias for updating parameters, and thereby influences the structures of the networks which form the creature's kernel. The necessary signal comes mainly through the motivation system, which decides the direction of updates. (See Figure 6-1.) This latter process is called backward operation and is viewed as *learning*.



(a)



(b)

Figure 6-1: These figures show operation of a situated character. **a:** Without learning ability, a character's operation is a series of perception → decision making (such as action selection) → action. Not all the information flows are shown here for simplicity. See Figure **??** for a complete diagram. **b:** With learning ability, the result of previous action affects the character's decision making system, i.e., updates its internal model of the world. Next time, when it faces exactly the same situation, it may react differently or it may react in the same way with more confidence. The arrow in the box represents the result of learning's updating the creature's internal model and **perception(t+1)** represents its getting new input after updating the internal model as the result of learning.

Consequently, learning almost always occurs during an affective episode. In an attempt to understand the nature of changing conditions, people and animals revise and update their beliefs about the conditions necessary for maintaining their goals. This explains the role of the motivation system as a reinforcement signal for learning.

6.2.1 Implementation

Within the computational structure, the motivation system sends out necessary signals for learning to subsystems of the creature kernel during the backward operation phase of the character. In the mathematical framework, each subsystem network is interpreted as a graphical model [59] where each node has output values with its connection weights representing influential strength. System output is determined as a function of connections, associated weights, state of the external world and the state of the creature itself [68, 126, 7].

In this section, I describe both forward and backward operations with emphasis on the backward operation that corresponds to learning. I discuss different kinds of learning and the way they are implemented in this proposed framework.

There are three different kinds of learning that are necessary for the proper operation of the creature kernel. They are designated organizational learning, concept learning and affective tag formation.

Organizational Learning

Organizational learning refers to the updates in the networks. I will explain it using the behavior network as an example. The behavior network is updated through modification of the way nodes are connected to each other and the weights associated with each node. The existence of a directed arrow from behavior A to behavior B indicates that the activation of behavior A could lead to the activation of behavior B. In other words, the strategy of performing behavior B after behavior A is one of the available strategies of the creature, who is the owner of the system. So this part of the organizational learning which involves reformation of the network is also called *strategy learning.* Often, behavior B is one of the children nodes of behavior A, and the weight associated with the link corresponds to the level of *preference* given to behavior B compared to other children behaviors of behavior A. The initial value of this weight reflects the innate tendency of the creature. Good or bad experiences such as the encountering of dangerous objects while performing that behavior or *consummation* of certain drives would decrease or increase the weight, respectively. In effect, the result is manifested as the change in behavioral tendency or habit of the creature through its behavior, and thus this type of update is called *preference learning.*

Behavior Groups The behavior network is organized into groups of mutually exclusive behaviors called "Behavior Groups." Activation of the Behavior Group is always preceeded by the activation of its parent behavior [7]. The activation level of each behavior is determined as a function of the creature's internal state, releasing stimuli, level of interest and inhibitory gain combined with the level of preference given to each behavior from its parent behavior. This preference is proportional to the expected "reward" from performing each behavior. This reward comes through the motivation system, which, in turn, is represented by stance and valence values. Assume that a behavior P has a child Behavior Group, which has N children behaviors as its entity. Then the activation level of $i-th$ child behavior at time t can be represented as

$$V_{it} = PR_i \cdot [LI_{it} \cdot Combine\{(\Sigma_k RM_{kt}), (\Sigma_n IV_{nt})\}] \tag{6.1}$$

, where LI_{it} is the level of interest for behavior i at time t, V_{it} is the value of behavior i at time t, RM_{kt} is the value of releasing mechanism k at time t, IV_{nt} is the value of internal variable n at time t [7]. Here $PR(i)$ represents the preference level toward the behavior i from the parent behavior p. This activation level is combined with the inhibitions from other sibling behaviors (i.e., behaviors that belong to the same Behavior Group), so only one of the behaviors in the Behavior Group ends up being active at the end of the update tick. This can be written as Eqn 6.2,

$$V_{i,t+(n\times\Delta)} = V_{i,t+((n-1)\times\Delta)} - (\Sigma_{m\neq i} N_{mi,t+(n\times\Delta)} \cdot V_{m,t+(n\times\Delta)}) \tag{6.2}$$

and, here, N_{mi} is the inhibitory gain from behavior m toward behavior i.

Bayesian Inference and Learning The level of preference from a parent behavior to one of its children behaviors is proportional to the expected reward from that particular child behavior. Expected reward for behavior i, $E(R_i)$ can be written as

$$E(R_i) = w_{vi}E(v_i) + w_{si}E(s_i) \tag{6.3}$$

where $E(v_i)$ and $E(s_i)$ are expected valence and stance value for behavior i, respectively, and w_{vi} and w_{si} are mixing weights , which are subject to vary from individual to individual creature, and from behavior to behavior as well. For simplicity, we calculate the expected values for valence and stance separately [1] and combine them together later.

This preference level from the parent behavior to a certain child behavior is updated after each child behavior's activation following the activation of the parent behavior. Let $EV_i(t)$ represent expected valence of child behavior i at time t; $EV_i(t+1)$ is updated as

$$EV_i(t+1) = EV_i(t) \cdot \frac{N+\alpha}{N+\alpha+1} + \frac{I \cdot v_i}{N+\alpha+1} \tag{6.4}$$

after the child behavior i is activated. Here, I is the *arousal* level of the affect and v_i is the actual valence that the creature just experienced as the result of the activation. α is the number of accumulated activations of the behavior group which this child behavior i belongs to, and N is a constant related to the confidence of the initial EV_i value. Let behavior j be another child behavior that belongs to the same behavior group; the expected valence is updated as:

$$EV_j(t+1) = EV_j(t) \cdot \frac{N+\alpha}{N+\alpha+1}. \tag{6.5}$$

Accordingly, the preference value for behavior i is calculated as the normalized value of $Clamp(0, 1, EV_i)$. Where $Clamp(min, max, v)$ has the following property.

$$\begin{aligned} Clamp(min, max, v) &= min \quad if\ v \leq min \\ &= max \quad if\ v \geq max \\ &= v, \quad otherwise. \end{aligned}$$

And we take the normalized value to make the sum of the preferences values of behaviors in the same behavior group equal to *one*. See Appendix A for the derivation

[1] Under this treatment, valence and stance are considered as independent entity in the defined space [101].

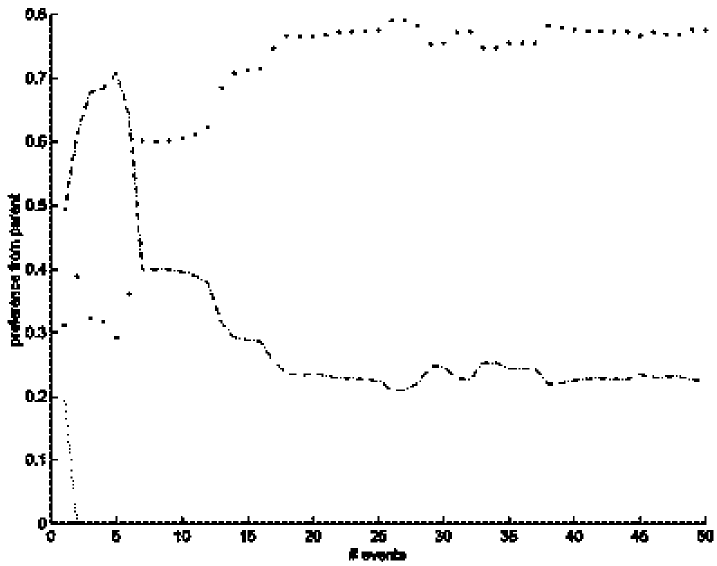
of the learning rule.

In Figure 6-2, I show an example of preference update traces. Here the number of behaviors in the behavior group is 3 and the figure shows the update in preference values from the behavior to each child behavior through 50 instances of the activation of their parent behavior. In this simple simulation run, it is assumed that stance value does not contribute to the reward. See Appendix C for actual matlab code for this simulation.

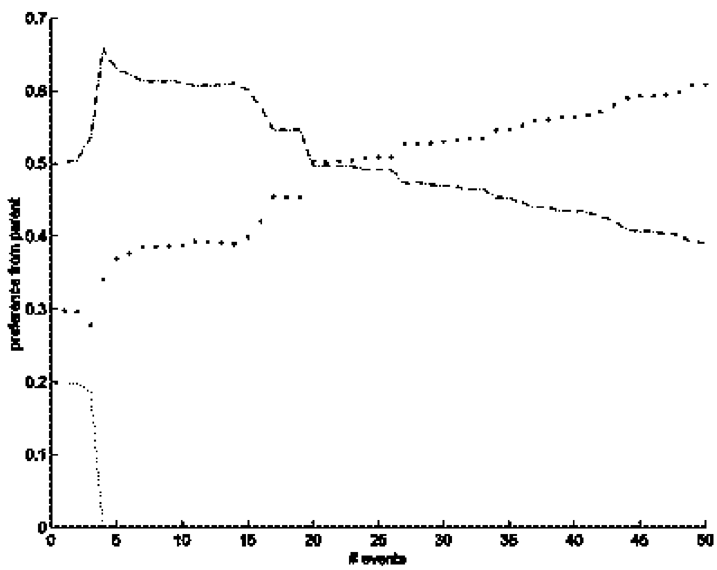
Concept Learning

By concept, we mean each character's attitude toward or belief about certain objects or events, which influences his behavior when he encounters such objects or events. For example, if there is a character who has a concept that tigers are scary, he will behave cautiously and will try to run away from it when he meets a tiger, whereas if there is another character who has a concept that animals are fun to play with, the best guess that he can make when he meets a tiger for the first time is that it is going to be fun to play with, and he will try to approach the tiger.

We assume that each character is *born* with some built-in concepts. Another assumption is that there is a set of features that each character cares about when he judges whether a certain concept is right or wrong with respect to a given object or event. For example, assume that a character uses size and brightness of objects as reliable features for judging the nature of the objects (e.g., scariness). This is graphically represented in Figure 6-3. Gray area is where he believes size and brightness features of scary animals are located. So this shows that initially his belief is that all animals are scary no matter how small it is or how bright the animal's color is. Given this, at the first instance of meeting an animal (e=1), he will behave under an assumption that the animal will be scary, following MDL (Minimal Description Length) principle. Assume that he met a white cat and behaved very cautiously. After figuring out that the white cat is fun and not scary, he will narrow his con-



(a)



(b)

Figure 6-2: Above two figures show the progressive change in the preferences in a hypothetical case of one parent behavior affecting a child behavior group which consists of three children behaviors. In this example, it is assumed for simplicity that stance component does not affect the learning. Each graph is plotted as line segments, dots and line, or dots, representing child behavior *one*, *two* and *three*. In this case, N is 10, the initial preferences for each behavior were 0.2, 0.3 and 0.5. Activation of each child behavior results in a stochastic process of reward with mean -4.15, 0.8 and -0.1, respectively, with variances being all 1 for **(a)** and 2 for **(b)**. Both figures show that as the event proceeds, the preference for behavior *two* increases, approaching 1.0, and the preferences for the other two behaviors converge to zero, but more slowly when the data have higher variance.

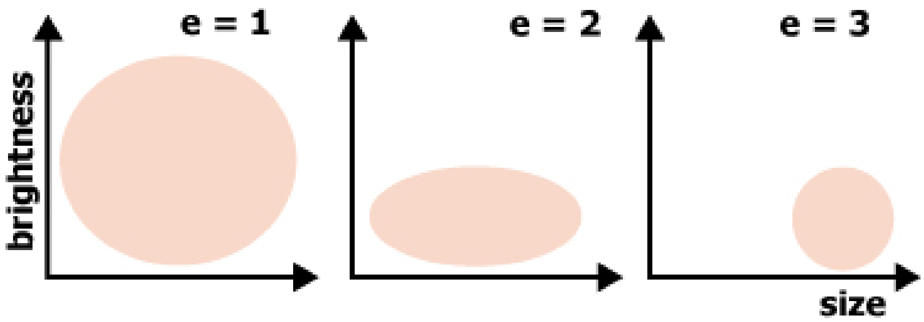


Figure 6-3: Assume a character whose initial belief (left side) is that all animals are scary no matter how small they are or how bright their fur colors are. As he interacts with animals that break this current belief, he updates his concept to reflect his own experience (righthand graphs). This is called concept learning and implemented using Bayesian belief update and MDL principle. See text for details.

cept down and start thinking that only dark animals are scary. Then he meets a little black mouse (e=2) with an assumption that the mouse will be scary because it is a dark animal. Then if he again figures out that the mouse is not scary, he will update his concept again to include only dark and large animals in the category of scary animals and apply this assumption when he meets an animal next time (e=3).

This process of belief update is called concept learning, and implemented as part of the creature kernel learning system as follows. Each object is represented as a vector whose dimension equals the number of features that the character cares about. And each concept is coded as a set of vectors that belong to the concept. So, assume that for a concept C, the character has example vectors $\overline{x_j}$ that all belong to a concept C, and each $\overline{x_j}$ is a vector whose dimension equals the number of features that the character cares about. When he encounters a new object y, he thinks that y belongs to the concept C with probability as shown in Equation 6.6.

$$P(\overline{y} \in C | \{\overline{x_j}\}) = \frac{1}{(1 + \frac{d}{R})^{n-1}} \tag{6.6}$$

Here n is the dimension of the vector, and R is the farthest distance between two example data that belong to the set $\overline{x_j}$ along each dimension. And d represents the shortest distance between y and a datum that belongs to the example set along the feature axis. This measure is multiplied for every dimension to calculate the probability of y belonging to the same concept that the set $\overline{x_j}$ represents. So, if all features of y fall in the area bounded by the example features, the creature will behave with the strong belief that y belongs to the concept C, whereas if d is large for every dimension, his belief in y's belonging to the concept C will be low. After interaction with y, if it was discovered that y belongs to the concept C even if $P(\overline{y} \in C|\{\overline{x_i}\})$ was not 1, the character will add this new datum y to the concept C representing set, and update the concept. As examples are added and R for a certain feature axis gets larger, that axis becomes less and less useful in describing the concept [121]. This learning rule is also used for representing and updating beliefs in certain events, situations or behaviors.

In Figure 6-4, I show a simple example of how a concept is built as experience accumulates. Dots represent data points gained through experience, which belong to the concept that the creature is currently learning. Here, it is assumed that the value along the x axis is the only feature that is relevant for determining the concept. Dotted line represent the concept distribution at the initial stage of learning. Without enough experience, the creature safely assumes a rather broad distribution. As more examples are gathered, the creature is now able to clearly distinguish the data points that belong to the concept and those that do not (solid line).

Affective Tag Formation

One of the primary roles of the motivation system is to offer a very efficient way of making quick decisions, through a mechanism that was called "Somatic Marker" by Damasio [27]. Here we introduce a more general concept, *Affective Tag*, in order to avoid the unproven hypothesis that such a mechanism is always associated with a peripheral state of the body. Affective tags provide bias to action selection in a form of emotional memory. Even when there is no other strong cue to prefer one way versus the other, an affective tag can still function as a bias to make the character feel neg-

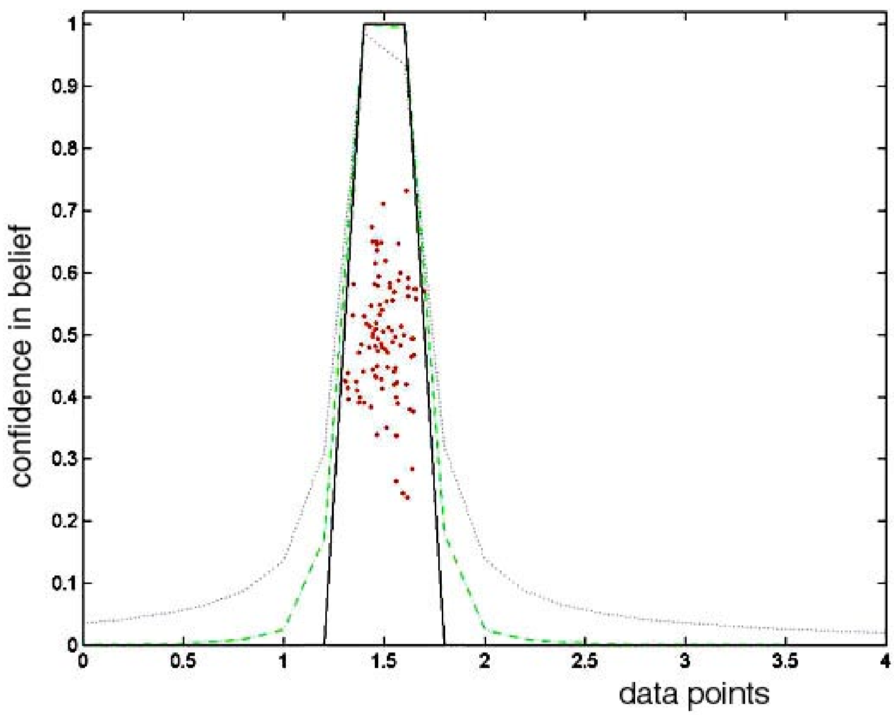


Figure 6-4: This graph demonstrates the change in a concept as a creature's relevant experience accumulates. Here, it is assumed that the value along the x axis is the only property that is relevant in determining the concept. The dotted line represents the concept at an early stage and the solid line represents the concept after ten iterations of exposure to the concept and learning. It shows that as experience accumulates, false positives are effectively suppressed and the confidence in the region where the actual data come from approaches one. See Appendix D for the actual matlab code which generated this hypothetical simulation.

ative stance to a red umbrella even though he has never seen it before, just because he had a bad accident involving a red car. Or it influences the character to decide to make an appointment to be on Wednesday as opposed to Thursday even if there isn't any clear advantage of having the meeting on Wednesday, because his joyful experience on last Wednesday biases him to have a positive attitude toward Wednesday.

In our system, this functionality is implemented as Hebbian connections [48] between the units in the motivation system and objects of interest; the connection weight $W_{i,j}$ is updated using the learning rule that can be written as $\delta W_{i,j} \propto O_i \cdot M_j$. This value is recalculated whenever an object O_i is involved in behavior and got motivational feedback M_j as the result of the behavior. So, for example, when a certain behavior selection is made, the motivation unit provides bias to behaviors differently based on objects of interest that each behavior is likely to engage. Though the influence is small compared to the influence of major drives, positive feedback resulting from the mutual inhibition mechanism biases behaviors with stronger affective tag inputs.

6.3 Basis for Shaping Personalities

Although the organizational structures of all of the four main parts of a creature contribute to shaping its personality, the motivation system plays the most fundamental role for making a creature distinguishable from other individuals. In different individuals, preferences are set differently for different drives and associated affects (desires). From mere preferences for certain colors, tastes or shapes to the relative hierarchy of internal motivational structure and also through details in the way each emotional state is expressed outwardly, various parts of the motivation system contribute to forming a unique personality of a creature. The motivational system nevertheless keeps it as adaptive and functional as any other individual of the same species with a different personality. The motivation system thus allows diversity among creatures, and it brings a real as well as a virtual creature to life by biasing its behavior and by being exhibited through emotional expressions [124].

Chapter 7

Learning and memory

Learning is an evolutionary solution that many of the creatures around us have found and adopted for the apparent advantages it guarantees for survival. The modifiability can be regarded as a species-characteristic adaptation to changes in the environment that are to be expected, but the directions of which are not predictable. This ability to learn is central for making characters adaptive and robust, which is a major theme of this thesis work. In this chapter, I explain various types of learning that are subject to implementation, and discuss how they are implemented and the necessary structures and assumptions that are necessary to support them.

7.1 Learning

When we use the word 'learning', it may refer to more than one type of behavioral adaptation. Here I briefly review the basic concepts and general theories of learning.

Rescorla [97] states that experiments on learning and memory should be based on examination of the organism's experience and behavior at two separate times. At the first time (t1), the organism is exposed to a particular experience – a sensory stimulus or a particular opportunity to learn. At a later time (t2), the investigator assesses the organism to determine whether the t1 experience has modified its behavior. The aim is to determine whether a particular t1 experience produces an outcome at t2 that would be absent without the t1 experience. This in turn implies the two step process of learning observed in nature, which is also the process that underlies its implementation. In the creature kernel framework discussed so far, this two-fragments view of learning is reworded as follows. First, the results of previous action or strategic choice are fed back to the creature after being sampled from the external world, *via* the creature itself through the creature's perception system, then second, modifications of appropriate parts of the creature kernel in an appropriate manner are decided and made.

The most prevalent way of categorizing types of learning is based on the second of these two fragments. It categorizes learning into three main types and one of the following three experimental paradigms is used to probe each of these three learning types.

1. Without any other event or constraint, a single stimulus (S1) is presented to the organism. The result may be habituation, dishabituation, or sensitization, which are all types of nonassociative learning.

2. One stimulus (S1) is presented in relation with another stimulus (S2). This paradigm (called Pavlovian, or classical conditioning) allows us to study how the organism learns about the structure of its environment and the relations among stimuli in its world.

3. A single stimulus(S1) is delivered in such a way as to reinforce a certain behavior. Called instrumental, or operant conditioning, this paradigm allows us to study how an organism learns about the impact of its own actions on the world.

Above experimental paradigms correspond to the three basic paradigms of learning and what each experimental paradigm studies are called nonassociative learning, classical conditioning and operant conditioning, respectively.

Nonassociative learning involves only a single stimulus at t1. Three kinds of nonassociative learning are habituation, dishabituation, and sensitization. Habituation is a decrease in response to a stimulus as the stimulus is repeated, whereas an increase in response amplitude over the baseline level is called dishabituation. Even a response that has not been habituated may increase in amplitude after a strong stimulus. This effect is known as sensitization. The response is greater than its baseline level because of prior stimulation.

Learning that involves relations between events – for example, between two or more stimuli, between a stimulus and a response, or between a response and its consequences – is called associative learning, one form of which is **classical conditioning** where an initially neutral stimulus comes to predict an event. In instrumental conditioning (also called **operant conditioning**), an association is formed between the animal's behavior and it consequences [125]. While these different types of learning occur, different brain regions may process different dimensions or attributes of a learning or memory situation: spatial, temporal, sensory dimensions, responses and emotional aspects [61, 79]. This broad classification scheme is widely accepted by contemporary learning theorists.

7.1.1 An ethologist's view of learning

A set of learning abilities necessary (and probably sufficient) for making characters robust and adaptive, are found in animals that are capable of surviving in a dynamic and unpredictable natural world. Within these broad boundaries, Lorenz classifies types of learning that are observed in animals as described below. I have used the Lorenz treatment of learning because it is more comprehensive than most, encompassing a greater variety of phenomena.

In animals, through adaptive modifications of behavior, the process of learning selects from many possibilities, contained in a pool of action primitives and possible combinations of them, the one that seems to fit current circumstances best. Environmental influences furnish the information about which possibility this is. Following

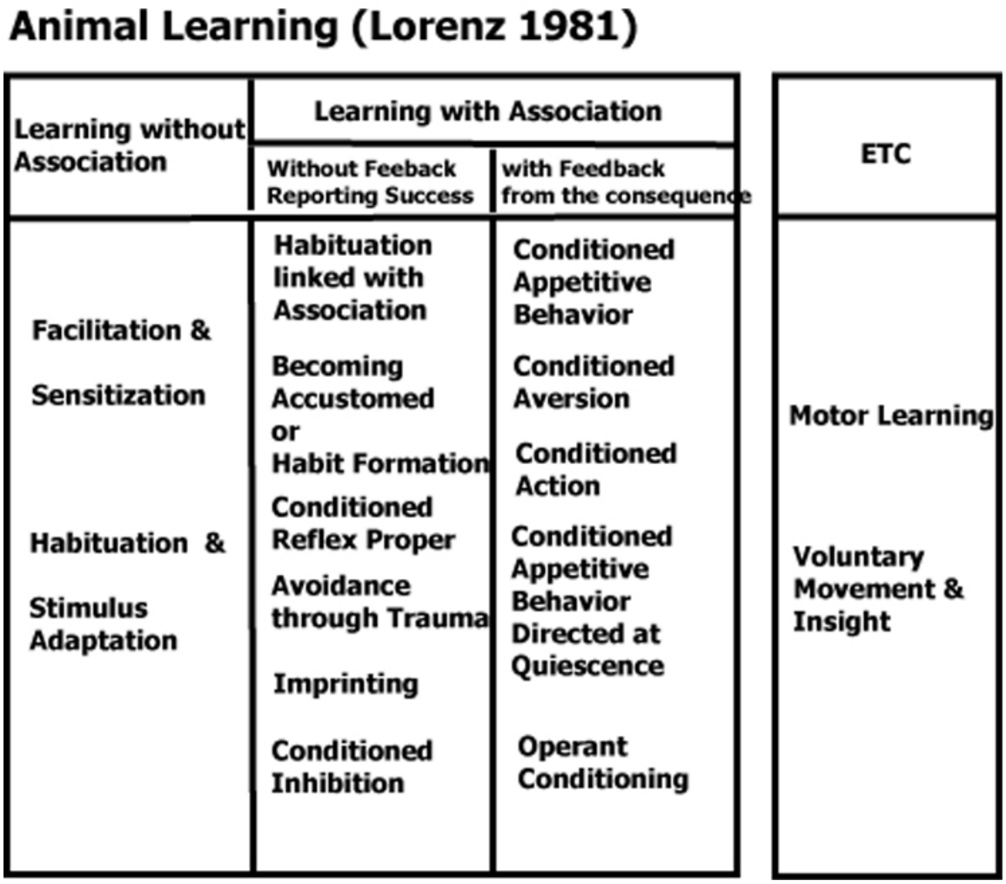


Figure 7-1: This diagram shows how Lorenz classifies types of learning that are observed in animals [68]. Roughly, learning types are categorized as nonassociative learning, associative learning and other types of learning. See text for details.

Lorenz's view [68], animal learning is classified into three main categories: nonassociative learning, associative learning and other types of learning. The various types of learning and his adopted classification scheme are shown in Figure 7-1, and a brief description of each is as follows[1].

Nonassociative learning

Nonassociative learning is the most primitive type of learning. It can be categorized into two kinds of learning: faciliation/sensitization and habituation/stimulus

[1]Actually, this list gets fairly long because there are a number of distinct learning types that he stresses. I list them with descriptions, which directly follow Lorenz's explanations in his book [68], because the way he uses some terminologies is different from more common usage of those words among learning theorists. Readers who are already familiar with Lorenz's usage of terminologies may skip this section.

adaptation.

Facilitation and sensitization Performance of some behaviors is smoother and faster after the behavior pattern is performed a few times repetitively. This phenomenon is called facilitation and, Lorenz specifically points out that this is different from maturation, even though the effect might be similar in growing organisms.

Similarly, the threshold of response to key stimuli get lowered as certain sensory input, and the response to it, happens repetitively. This process takes place in the sensory sector of the CNS and is called sensitization.

Habituation or stimulus adaptation Animals have the ability to deal with a constantly reoccurring stimulus situation. If there are specifically unchanging aspects of the environment, they adapt to these and still respond to more informative environmental cues adequately. As a result, the animals show releasing responses only when the stimuli are novel, and when they are detected constantly and repeatedly the animals no longer exhibit the response.

Associative learning

Learning through association can be viewed as a connection forming process. Here association is defined broadly as connections between *contents of awareness*. Associations are produced when two events happen at once or several times in the same sequence and within short intervals of time, where the interval varies based on species and the type of events. Again, associative learning can be divided into two categories - associative learning without feedback and associative learning with feedback.

Associative learning without feedback reporting success

- **Habituation linked with association.** Habituation to a certain innate releasing stimulus may be specific to a more complex stimulus configuration repeatedly associated with the releasing stimulus. Any change in the associated stimuli can result in a renewed response to the releaser.

- **Becoming accustomed or habit formation.** The process of becoming accustomed is a kind of habit formation in which an originally effective combination of key stimuli loses its releasing function unless associated with the perception of a more complex configuration. The original key stimuli are still essential, but are effective after the accustomization process only when they occur within the context of the associated more complex configuration.

- **Conditioned reflex proper or conditioning with stimulus selection.** 'Conditioned response' refers to the designation of the result of simple association. This association connects an originally ineffective stimulus with a reaction that can be regarded as a 'reflex,' insofar as its function does not depend on the changes of internal readiness. Skinner's operant conditioning experiment, where an intended behavioral response is learned as a response to an originally ineffective stimulus through conditioning of type S [110], and Pavlov's classical conditioning [97] experiment belong to this category.

- **Avoidance through trauma.** Some acquired avoidance responses are similar to a conditioned reflex, or are a special case of this. Here, the fact that the response that the originally indifferent, but newly associated stimulus releases is acquired very rapidly and is particularly long-lasting or even irreversible, as in very strong escape reactions [33], puts this in a special category.

- **Imprinting.** Imprinting is an irreversible early and strong learning phenomenon. At a particularly early stage of an organism's life, imprinting proceeds as programmed without any feedback from the effect of an action. As the result of this process, an innate response becomes linked to its biologically adequate stimulus [68].

- **Conditioned inhibition.** By repeatedly punishing a behavior right before it is initiated, inhibition of that behavior can be established. In this case, the association is made between the punishment and the intention to initiate the behavior. This is distinguished from a conditioned reflex in the sense that conditioned inhibition is associated not with a reflex nor with a passive readiness to react, but with a central nervous impulse.

Associative learning with feedback from the consequence The prerequisite for learning by success or failure is a "feedback mechanism," i.e., a regulation cycle. Von Frisch showed that a honey bee which has flown repeatedly to a blue flower and found no nectar but has found nectar when visiting a yellow flower will thereafter fly to yellow flowers only, even if the yellow color did not have a releasing effect originally [136]. This is an example where feedback affects the modification of behavior. In this case, animals gather information concerning the stimulus configuration in which a certain motor pattern brings a reinforcing reward to the animal. Through this process, the animal comes to learn the selection of a stimulus situation rather than the selection of an action.

- **Conditioned appetitive behavior.** When rewarding or punishing experiences that come after performance of a certain behavior are fed back to the precedent behavior, this feedback information affects the likelihood of any appetitive behavior that is directed toward the action. Conditioned appetitive behavior is distinguished from the conditioned reflex in the following sense. Whereas the conditioned reflex constantly results in releasing the same motor pattern when the associated stimulus configuration is presented, conditioned appetitive behavior can result in even other kinds of appetitive behavior that are phyletically programmed even if they did not occur during the original learning process.

- **Conditioned aversion.** In the same way that conditioned appetence lies functionally in parallel with the conditioned reflex, conditioned aversion can be seen as the functional counterpart of conditioned avoidance. The difference is that the direct connection between the conditioned stimulus and the response is affected in the former case, but not in the latter. In the latter case, the adaptive information is derived from the feedback, which affects the nervous pathways that lead to the aversion in the same way as the conditioned appetence facilitates even appetitive that did not occur during the learning process.

- **Conditioned action.** When a behavior pattern is rewarded right after an animal performs it, the preference to perform that behavior increases. The reward may come from an altogether different behavior system, and the next time the motivation is awakened the animal prefers to perform that previously rewarded behavior. It can be seen as the behavior being registered as a means of satisfying the motivational need.

- **Conditioned appetitive behavior directed at quiescence.** The motivation to get rid of a primary annoyer is such a strong one that the deliverance from a strongly disturbing stimulus acts as a very strong reward. The result is that the animal learns very quickly the corresponding appetitive behaviors, and does not easily forget. This explains why alcohol or drug addiction is so strong.

- **Operant conditioning.** Operant conditioning here is confined to learning processes in which not the stimulus situation but the behavior pattern is selected among the repertoire of behavior patterns that the animal is concerned with. Operant conditioning can be viewed as the process of learning the skill that maximizes reward within that environment. In nature, operant conditioning is very rare and mainly occurs under the influence of disturbing stimulus situations in an attempt to get rid of the annoying situations, i.e., within the context of appetence directed at quiescence.

Et cetera

Other than those described so far, there are additional learning types found in animals. These are not subject to implementation as far as this thesis work is concerned, but I will include a brief description of what each means for the completeness of the list.

Motor learning. Acquisition of a new motor pattern happens through formation of a chain of conditioned actions. For example, teaching a dancing sequence to a pigeon is done through this process by linking one directional turn after another.

Voluntary movement and insight. The evolutionary development of voluntary movement and the differentiation of mechanisms exploiting instant information, especially information about spatial arrangements, are closely connected. Insight is the result of these mechanisms. In animals, the ability to adapt movements to their physical limitations and the ability to gather goal-related spatial information are closely connected, so that the separation of sensory and motor functions is not always immediately apparent.

7.1.2 Integration with contemporary learning theory

Lorenz's classification scheme described so far is based on his close observations of real animal behaviors and emphasizes some learning types that are particularly strong. Figure 7-2 show how his classification scheme can be mapped onto the more common learning theorists' framework. In the figure, terminologies in the solid boxes are ones used by contemporary learning theorists, and those in the open boxes are the terminologies used by Lorenz to explain animal behavior.

Contemporary learning theorists classify learning into two main categories; they are nonassociative learning and associative learning. This categorization agrees with Lorenz. He refers to each as **learning without association** and **learning with association**, respectively, and with these he starts his further categorization. Associative learning is subdivided into classical conditioning and operant conditioning. These correspond to learning **without feedback reporting success** and learning **with feedback from the consequences** in Lorenz's terminology.

Nonassociative learning is subdivided into habituation, dishabituation and sensitization by modern learning theorists. Lorenz categorizes it slightly differently. He also explains what is called sensitization using his terminology, **facilitation or sensitization**, but he uses one item, **habituation or stimulus adaptation** to include

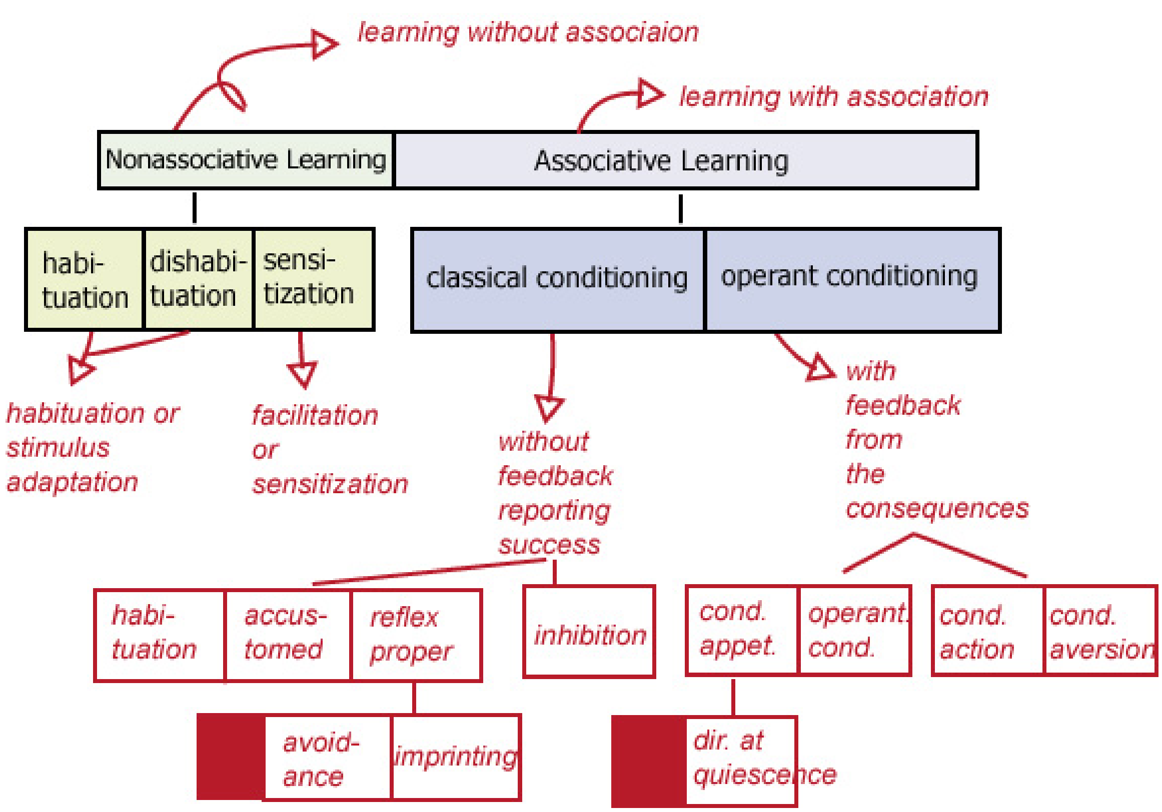


Figure 7-2: Lorenz categorized different types of learning based on his close observation of animal behavior. His categorization can be mapped onto more general categorization scheme as shown in this figure. See text for details.

both habituation and dishabituation. He views habituation and dishabituation as essentially the same process only with opposite effects on the behavior.

Comparisons for associative learning get a bit more complicated. Lorenz uses the terminologies, learning **without feedback reporting success** and learning **with feedback from the consequences**, to refer to classical conditioning and operant conditioning, respectively. But then he subdivides each to distinguish more specific cases.

He mentions six different subcategories to explain specific cases of classical conditioning type learning. Among them, **habituation linked with association**, **becoming accustomed or habit formation**, **conditioned reflex proper** and **conditioned inhibition** are distinguishable learning types based on where the actual learning takes place or where the necessary signals for learning come from. **Avoidance through trauma** and **imprinting** are phenomena that may be viewed as special cases of **reflex proper**. Both are not unlearnable and they strongly alter behaviors. But in terms of the mechanisms of learning, both can equally be seen as subcategories of **conditioned reflex proper**.

Five subcategories are presented as representatives of operant conditioning type learning. Again, among them, **conditioned appetitive behavior**, **true operant conditioning**, **conditioned action** and **conditioned aversion** are distinctive in terms of the related learning algorithms. He mentions **conditioned appetitive behavior directed at quiescence** as the fifth category of operant conditioning type learning, but it is a special case of conditioned appetitive behavior whose strength is distinctively large.

Here, the terminology *operant conditioning* is a bit confusing because it refers to two different kinds of learning. By modern learning theorists, operant conditioning is used more broadly for phenomena where an individual creature learns about the impact of its own action and how it affects the world. Lorenz prefers to use the term in a more narrow sense and he refers to learning that involves searching through an available repertoire of actions until the creature finds the one with a wanted impact, forgetting previously tried actions and sticking to the found action that brings the desired impact.

These different learning types are discussed again in section 7.3, locating where each happens and specifying the components involved in each learning type within the creature kernel framework.

7.2 Memory

The terms "learning" and "memory" are so often paired that it sometimes seems as if one necessarily implies the other. We cannot be sure that learning has occurred unless a memory can be elicited later. Like learning, memory is a collective term that refers to a number of distinguishable kinds. Here I introduce some basic concepts of memory, and then explain what are the necessary memory types and how they fit in the creature framework for implementation of the necessary learning types.

7.2.1 Classification by type

The basic distinction is betwen declarative memory and nondeclarative memory. Declarative memory is what we usually think of as memory: facts and information acquired through learning. It is memory we are aware of accessing. And nondeclarative memory is shown by performance rather than by conscious recollection. It is sometimes called procedural memory. Declarative memory deals with *what* and nondeclarative memory deals with *how* [40].

Each is also a collection of slightly different kinds of memory. Endel Tulving [133] defines semantic memory and episodic memory as two main categories of declarative memory. Semantic memory is generalized memory, such as knowing the meaning of

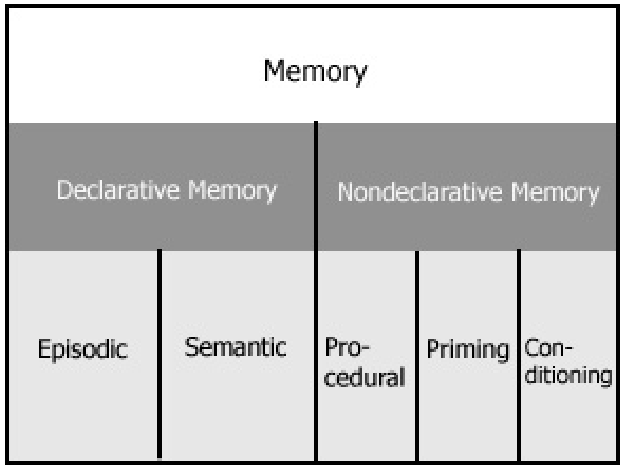


Figure 7-3: Memory is a collective term, which usually refers to different kinds of memories. This figure shows a way of classifying different types of memories in terms of inclusion relations [99]. See text for details of this classification scheme and explanation of each memory type.

a word without knowing where or when one learned that word. On the other hand, episodic memory is autobiographical memory that pertains to a person's particular history.

Procedural memory is the most representative of nondeclarative memory, whose subcategories are procedural memory, priming and conditioning. It deals with the problem of *how*. Skill acquisition is a good example of the formation of procedural memory. Priming, or repetition priming, is a change in the processing of a stimulus, usually a word or a picture, as a result of prior exposure to the same stimulus or related stimuli. Conditioning is a context-dependent memory type. It stores the information about a particular response's being appropriate to a particular stimulus in one setting but not in another. More formally, the conditioned stimulus(CS) or unconditioned stimulus(US) becomes associated with the context in which it occurs. A CS evokes a conditioned response(CR) in one context but evokes no response (or not as strong a response) in another. Figure 7-3 summarizes this categorization scheme based on memory types.

7.2.2 Classification by duration

Another way of classifying different memory types is classification by their durations. Some memories last longer than others, and this is what I mean by different "durations" of memories. The multiple trace hypothesis well classifies different types of memory by duration [81] and it can be summarized as follows.

Iconic memory is the briefest memory. An example would be transient impressions of a scene that are illuminated only for an instant. These brief memories are thought to reflect the continuation of sensory neural activity, the so-called sensory buffers.

Somewhat longer than iconic memory is the **short term memory** (STM). Storage of a telephone number that was never used before, which lasts until the call is made if nothing distracts, is one example of this type of memory. However, the usage of this terminology, STM, is not consistent among the investigators from different fields. Cognitive psychologists, who first used the term, found that if subjects are not allowed to rehearse, STM lasts only about 30 seconds [18, 91]. Many biologists define STM as memory that is not permanent but that lasts for minutes or hours, even up to a day.

Some memories last beyond the short-term memories but fall short of long-term memories. Remembering the parking spot that does not last more than a day is an example of such memories and they are called intermediate term memories. Intermediate memories are characterized as memories that outlast STM but that are far from being permanent [81, 98].

Beyond ITM are memories that last for weeks, months, and years; these are called long-term memories(LTMs). Because many memories that last for days or weeks do, however, become weaker and may even fade out completely with time, some investigators use the term permanent memory to designate memories that appear to continue without decline for the rest of the life of an organism, or at least as long as the organism remains in good health.

Given this background, next I present learning and memory types that are relevant or useful for the synthetic characters and how those features are implemented within the proposed creature kernel framework.

7.3 Learning for synthetic characters

Basically, within the creature kernel framework, the results of learning affect formation of associated weights of the links between basis units. For example, habits or available skills are coded as links that connect behavior basis units in the behavior network, and the comparative strengths of such connections reflect how much the creature is familiar or accustomed to those action sequences.

Whereas nonassociative learning deals with relations between the perception system and the behavior system, associative learning includes the motivation system as well.

Facilitation and sensitization are categorized in the same group in the sense that both result in strengthening of certain connections. Facilitation refers to strengthening, or increasing the weight of a connection within the behavior system, which can be implemented using the organizational learning algorithm. Sensitization affects connections between the perception system and the behavior system, which is implemented as a combination of concept learning and organizational learning. Sensitization relatively strengthens certain links between perception units and behavior units, which effectively shifts the focus of attention related to that behavior. On the other hand, habituation and stimulus adaptation weaken connections. Habituation also affects the link between the perception system basis unit and the behavior system basis unit and it is implemented as a combination of concept learning and organizational learning within the creature kernel framework.

The area where the associative learning types happen extends over the motivation system in addition to the perception and the behavior systems. **Habituation linked with association** is realized through having a previously effective stimulus be part of a new stimulus configuration. This corresponds to concept learning in the perception system, and organizational learning at the connections that connect the perception system and the behavior system which then form a new context as part of the creature's behavior repertoire. **Becoming accustomed, or habit formation,** is implemented as a concept learning process in the perceptual system linked to an existing behavior repertoire. The refined stimulus reflects the learning and experience in the situation or environment in which the character is located. Lorenz uses the term *habit formation* for this because it involves becoming accustomed. It should not be confused with habit formation which is the result of facilitation of a behavior sequence through preference learning within the creature kernel learning framework. **Conditioned reflex proper** has the broadest meaning within this category. It results in a new formation of connections and strengthening of those connections, i.e. organizational learning, between certain characteristics of the perception system and/or the motivation system with the behavior system. At the execution of the conditioned reflex, the associated formation (perceptual input or motivational state) leads to that behavior. Special cases of this type of learning are **avoidance through trauma** and **imprinting**, which are so strong that they are not unlearned and a single experience (i.e. learning example) is often enough to form the link. **Conditioned inhibition** is implemented as the addition of an effectively negative link to an existing context.

Associative learning with feedback from the consequence is different from the learning types within the category of **associative learning without feedback from the consequence** in terms of the learning processes involved. While the latter assumes that the interconnected target nodes involved are already known to the creature, the former requires the creature's active searching for such target nodes, within the behavior system.

Conditioned appetitive behavior is a process of adding more behavior nodes to an existing chain of actions that lead to a final consummatory behavior in the context of satisfaction of a specific motivation. Appropriate appetitive behaviors are found through various trials of the character and these are placed between the relevant motivation nodes and the associated consummatory behavior nodes. This learning results in adding connections between the motivation system and the behavior system as well as connections within the behavior system. **Conditioned appetitive behavior directed at quiescence** is a special case of **conditioned appetitive behavior** and it is dealt with as a special case because the motivation to get rid of the primary annoyer is very strong in animals, and this strong motivation sometimes challenges the animals to go through some unusual behaviors. This probably reflects the fact that the strong motivation makes the searching for the appropriate appetitive behaviors very intense and this does not allow the animals to give up the attempt to find one. Similarly, **operant conditioning** in Lorenz's sense consists of finding out the proper behavior nodes that allow or lead to the relevant consummatory behaviors and thus satisfy the current motivational need. It results in adding connections between the motivation system modules and the behavior system modules. Both **conditioned appetitive behavior** and **operant conditioning** are implemented by organizational learning.

Conditioned action and **conditioned aversion** are viewed as changing the individual behavior associated probabilities. Both are implemented as a process of affective tag formation. The former associates positive affect with the behavior, i.e., it influences the character to execute the behavior more, whereas the latter associates the behavior with negative affect, and this results in having the character decrease its preference for performing the behavior.

7.4 Memory for synthetic characters

By type. Declarative memory assumes symbolic representations of events or objects that are subject to memorization. For human beings, language is the main means of presenting the content of the memory to other individuals. Some way of representing or pointing to the contents are needed for internal communication purposes such as comparisons for checking matches. At least for now, we have not given language processing ability to characters. This part of memory and learning has not been fully explored yet, but initial attempts are integrated and explained in Chapter 9.

The contents of synthetic characters' nondeclarative memory are reflected through their later affective responses and behaviors. The ability to do certain tasks or preferences for those tasks manifest the results of their procedural learning and thus reveals procedural memory.

Within the proposed creature kernel framework, available skills or performable behaviors reflect the existence of connections to and between behavior nodes, with associated weight values. In other words, the contents of the nondeclarative memory of synthetic characters can be seen as the existence and relative strength of the associative weights. In particular, procedural memory is located as the weight connections between behavior nodes within the behavior network. These are shown as performable skils and habits that are manifested through the possible repertoire of behaviors. On the other hand, priming and conditioning memories are inter-subsystem associative weights, i.e. weights that connect nodes in the perception subsystem and the behavior subsystem or the motivation subsystem and the behavior subsystem, that all together form contexts in which certain behaviors are performed, preferred or prohibited.

By duration. Since the lifespan of synthetic characters is different from that of real animals or human beings, it is hard to categorize the type of their memories using the same time measure.

One approach would be inferring which memory type that each memory belongs to based on analogy. For example, skills or habits acquired in animals usually last for more than weeks, which can be regarded as a kind of long term memory. From this perspective, we can view the procedural learning which is the main type of learning for synthetic characters as a process of forming long term memory (LTM).

Or, real time measures may be adopted for judging memory types and this can be adjusted as the designer wishes. One example would be to have the characters develop some built-in skills that last over multiple sessions, while other learning lasts only while the character-human interaction lasts. If the newly learned skills may last over multiple sessions so that the acquired knowledge endures as long as it is not unlearned, the interaction can be viewed as a process of forming long term memory, whereas if the knowledge is all gone whenever a new interaction session starts, it can be viewed as short term memory or intermediate memory depending on how long each interaction lasts. But all this is adjustable within the creature kernel framework and subject to modification depending on application purposes.

Chapter 8

Personality

Distinctive personality is another factor that makes a character come to look alive, and look intentional [124]. Especially when there are more than one character that face very similar problems, it is obvious that exactly the same look of those characters or the same behaviors of the characters would make the illusion of life quickly fade away, even if they can all adequately deal with the dynamics of their world in an intelligent way, and convey their affective states and thoughts in well understandable ways. Intention is not duplicatable.

Thus, strong or unique personalities give a better illusion of life to characters [124], and the challenge lies not just in giving them personalities but having each character consistently show its given personality in various situations. Characters should *stay in charater* to associate themselves with certain personalities which gives them uniqueness, and thus lifelike impressions.

On the other hand, it is impossible to impose distinct personalities on autonomous characters by the designer's hand-crafting appropriate responses for all the possible situations. Like implementation of the affect system [1], considering the exponential nature of the number of situations to think about when such an approach is taken and

[1] discussed in Chapter 6

the difficulty of maintaining consistency in personality with that method, a principled way of designing personality should be adopted instead.

8.1 The Big Five theory of personality

The variations among people come both from observable attributes as well as from dynamic internal 'personality processes.' In general, people's personalities are perceived through various attributes such as their temporary conditions, social roles and evaluations by others, their appearance and physical characteristics, temperament, character traits and abilities [103]. Apart from physical contributions, it has been proposed that the internal personality process comes from five broad orthogonal factors that have been called Big Five [41, 78]. The five factors are extroversion (factor 1), agreeableness (factor 2), conscientiousness (factor 3), emotional stability (factor 4) and intellect-imagination (factor 5).

8.2 Personality for synthetic characters

The approach proposed here for imposing distinctive personalities upon characters is not by having designers hand craft various responses for every situation, but by adjusting a countable number of parameters in the creature kernel, and letting these apply to every situation that the character faces. The personality of an individual synthetic character that is designed by adjusting the creature kernel parameters, can be conveyed through the way it behaves, the way it seems to feel in certain situations, how it adapts to the environment, etc., in addition to what it looks like. Variations in all such channels are combined together to attribute a certain personality type to the character.

Except for factor 5, the attributes of the big five personality factors can be roughly ascribed to different parts and mechanisms in the creature kernel, which helps the character programmers in deciding where to work to deliver the intended personality [2].

In particular, the **extroversion** factor is represented by promptness of responses to environmental and external inputs and level of interactions with other characters and human participants. **Agreeableness** corresponds to the readiness to change one's habits and beliefs based on inputs that come from other characters and the environment. This factor also affects the learning rate within the proposed creature kernel framework. **Conscientiousness** can also be shown through the way behaviors are performed. Sticking to a safe strategy with a bit of delay in response and learning is interpreted as due to a high level of conscientiousness. **Emotional stability** is the extent to which, and how easily, the affect system is influenced by behavioral consequences and external inputs. In addition to this process oriented approach, the visual appearance and the way each animation is designed also affect the way each character's personality is perceived. More detailed implementation examples and discussions are found in Chapter 11.

[2] Factor five involves cognitive processes, corresponding mostly to the neocortex level of the brain, which is little provided for by the creature kernel proposed in this thesis.

Part III

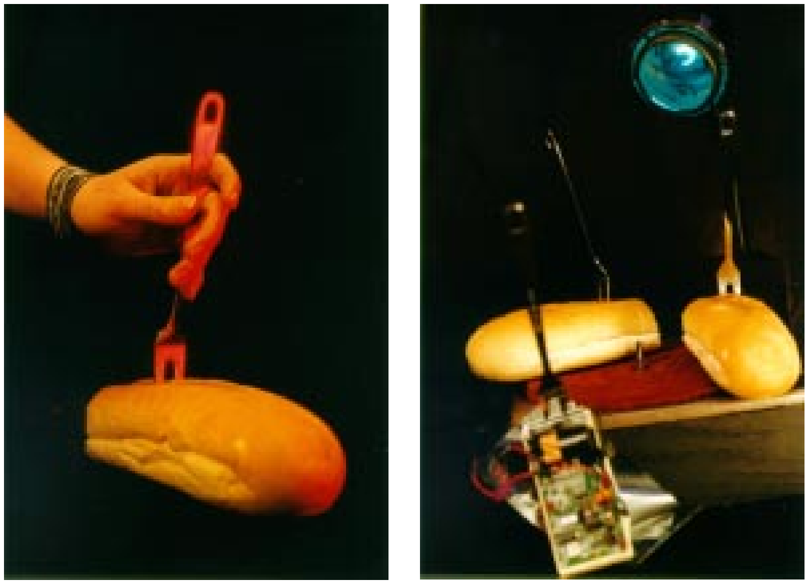
IMPLEMENTATIONS

This part of the dissertation explains two projects for which the characters have been designed using the creature kernel framework described in parts I and II. These two projects implement different virtual worlds and different human-character interaction settings.

The first one is **Sydney K9.0**, which implements a world inhabited by three graphical synthetic characters, Sydney the dog, Nabee the butterfly, and Fridge the refrigerator who essentially is the graphical instantiation of the human participant. Human participants can influence the world through speech signals delivered through a microphone, and a training stick which directs Sydney's attention. Like typical pet dogs, Sydney is a dog with its own desires and drives such as curiosity, hunger and fatigue. The person who is interacting with the dog can teach it some tricks, rewarding Sydney with yummy virtual food which satisfies Sydney's hunger drive. This interaction raises various interesting questions for the model such as how motivation and learning interact, and the roles of perception, attention and interference from other drives.

The second project is **(void*): A cast of characters**, which implements a dining hall setting inhabited by three distinctive characters – Earl the trucker, Elliot the Salesman and the fast dude, Eddie. Here, a human participant plays the role of the spirit of the night and can *possess* one of the characters by holding the *buns-and-forks* interface implements above the corresponding plate on the table. Moving the interface at the person's will makes the possessed character dance in the corresponding way. This dancing experience influences the character and make it feel in a certain way, and affects the attitude toward possession, and thus toward the human participant.

The characters in both projects implement the creature kernel framework described in this thesis and demonstrate both the plausibility and the limitations of the approach. I describe both projects in detail in the subsequent chapters and discuss findings learned through the actual implementations. In particular, the discussion of **Sydney K9.0** centers on the various issues related to learning and memory, whereas our consideration of **(void*): A cast of characters** leads to a discussion concerning the creation of characters with diverse personalities and how this influences the actual implementation as well as how the personalities influence the other characters and the human participants.



(a) (b)

Figure 8-1: (a) A hand holding the *buns-and-forks* interface. (b) The early implementation of **(void*): A cast of characters** was controlled by a record player and a pair of buns-and-forks interfaces. These are early versions of interface hardwares for the (void*) project, and I put these pictures here for historical interest. This project implements a dining hall setting where human participants can interact with characters' dance using a pair of buns-and-forks interfaces. See text for details.

Chapter 9

Sydney K9.0

Sydney K9.0 implements a virtual world where a dog named Sydney (See Figure 9-1.), which is designed using the framework of the developed creature kernel, lives. As

Figure 9-1: As the main character of the project, **Sydney K9.0**, Sydney shows off his various drives and affects such as hunger, curiosity, fatigue as well as surprise and satisfaction.

a dog, Sydney has his own desires and beliefs. Other characters in this project include Fridge and Nabee. A human participant can speak into a microphone to interact with Sydney, and can use a training stick to guide Sydney's attention, a milkbone box to give rewards and a clicker to signal a reward. Fridge, *dog's best friend*, is the graphical instantiation of the participant on the screen and plays the role of a dog trainer as shown in Figure 9-2. Essentially, the design of the interaction provided through this project is based on the "click-and-treat" training paradigm which is a positive reinforcement training system based on operant conditioning and made popular by animal behaviorists [140, 94].

This chapter explains the details of this project with emphasis on the character-human

Figure 9-2: **Sydney K9.0** implements a virtual dog training session where Fridge, a refrigerator in the virtual world, is the trainer, which is essentially the graphical instantiation of the human participant who can provide input to the system with the aid of various interfaces such as a wearable microphone, a training stick, a milkbone box and a clicker.

interaction and the character's learning that is evoked through the interaction. Analysis follows based on comparisons between the observed phenomena in this project and the real behavior of animals.

9.1 Characters

There are three graphical characters in the virtual world of **Sydney K9.0**. They are Sydney the dog, Fridge the refrigerator who is the dog's best friend, and Nabee the butterfly.

9.1.1 Sydney

Sydney is a virtual dog [1] implemented using the creature kernel framework explained in this thesis. Sydney is shown in various moods in Figure 9-1.

[1] Although the implementation of Sydney was inspired by real dog behaviors, Sydney is a "cartoon" character. We did not aim at exactly imitating a dog but rather at demonstrating the developed creature kernel architecture and learning algorithms associated with it.

Perception. Sydney's perception system is roughly divided into two parts. One is composed of virtual sensors and the other is composed of sensors that are responsible for receiving inputs from the physical world. Virtual sensors are taxonomy sensors which enable Sydney to access all the returnable features of the targeted objects [62]. In particular, the perception network is extended by weighted feature value providers [2] that return levels of beliefs in whether the objects are far, close, etc.

In this project, there are four different interfaces through which the human participant can interact with Sydney: a training stick, a clicker, a milkbone box and voice through a microphone. The user can move the training stick in any direction in the three dimensional space and the movement and resulting location of the stick is depicted by the corresponding positional change of the training stick that Fridge holds. Sydney gets the training stick-related information using his virtual sensor targeted at the virtual training stick in Fridge's hand. The clicker sound is treated as a special sound signal that is perceived as distinct from a speech signal, even though both are auditory signals. A special interface is attached to the system which stores a value from the time when a clicker signal is received through the attached microphone until the value is requested by and returned to the main system. When Sydney's behavior is rewardable, the human participant shakes the milkbone box, which elicits a bone on the screen which Sydney can receive and grab to satisfy his hunger. A shake sensor is installed inside the milkbone box, which is activated whenever the box is shaken. This signal is sent to the bone, not to Sydney, and Sydney only knows that a reward has been given through his perception of the bone through his bone taxonomy sensor. A speech signal is delivered to the system through a microphone. Speech data sampled at 11025 Hz are averaged and thresholded to be converted to vectors of 512 components. Data are converted to their Cepstral coefficient representations for further processing, which is discussed in Section 9.3. The module of the system that

[2] A value provider extends a transducer which is one of the programming primitives for implementing the creature kernel. It takes a set of objects and returns a floating value which is the result of the operation on the taken objects [62]. A feature value provider, in particular, returns a value calculated on certain features of the objects.

deals with this data conversion and acoustic pattern processing is called *Dog Ear* [3].

Motivation. Like any other characters implemented on the basis of the creature kernel framework, Sydney's motivation system is composed of two parts – the drive system and the affect system. Sydney has three main drives, which are hunger, curiosity and fatigue.

Sydney feels hungry if he does not get an adequate amount of food for a long time, and thus the food which satisfies the hunger drive is perceived as an incentive or positive reinforcement for Sydney. The hunger drive provides Sydney the motivation to learn through interaction with the human participant because learning and performance of a selected behavior gives Sydney a piece of milkbone as reward.

Curiosity drives Sydney to explore the world. Drives rise, at different rates, while they are not satisfied and satisfying stimuli are not provided, or while the key stimuli that trigger them are perceived. Curiosity rises if Sydney stays at one place for long, and exploring the world lowers curiosity. Sydney is also interested in knowing and interacting with other creatures such as Nabee (See Section 9.1.3). So, for example, when Nabee is close to Sydney, his curiosity rises and often it surpasses the other drives.

Fatigue brings Sydney back home and causes him to rest until he feels restored. After the fatigue level goes back to normal, Sydney gets up and start the behaviors that satisfy the currently most active drive.

Behavior. Sydney's behavior system is a sum of two main parts. One is the primary behavior part, which is a set of behavior nodes and sequences that can be used to satisfy goals, and the other is the autonomous behavior part, which contains more reflex-like behavior such as ear movement, which is triggered when Sydney detects

[3] Robert C. Burke at the MIT Media Lab, Synthetic Characters Group, did the implementation, so that I could integrate it into Sydney's creature kernel.

sound and *autonomous stand*, which is a low priority default behavior that is activated to avoid situations where no animation is running because no particular behavior has been chosen by the action selection mechanism. The primary behavior part is basically composed of behavior units such as walk, run, lie down, sit down, roll over, beg, etc., which are within the behavior network to trigger movement sequences for satisfying goals and drives.

Motor. Animation files made by animators are provided as the means of enabling Sydney to perform certain skills and display them on the screen. Leaf node behavior units, at the termination of the behavior system's branches, form a one-to-one mapping with animation files, where the behavior system blends into the motor system [4].

9.1.2 Fridge

Fridge is essentially the graphical instantiation of the user and is a very simple minded character. Sydney's hunger drive makes Sydney more eager to get into the interactive or training sessions, and this desire is expressed through his cheerful behavior communicating *"let's play"* in front of Fridge. Fridge does have a desire to interact with Sydney, which rises when there is no interaction and Sydney is far from Fridge. Fridge opens its door and show its gut filled with milkbones.

Fridge holds a training stick in its hand and the movement of the stick is determined by the direct input data from the physical training stick that the human participant controls. Since Sydney is interested in the tip of the training stick, the virtual training stick directs Sydney's attention. This influences Sydney's interactive learning, and is described in section 9.3.

[4] The underlying motor system that takes the affect parameters and blends hand-crafted animations to generate a new animation sequence with a certain affect quality was implemented by Michael Johnson [60], who is also at the MIT Media Lab, Synthetic Characters Group. All the character models and animations were generated by the artists in the group, Jed Wahl and Scott M. Eaton.

9.1.3 Nabee

Nabee, the butterfly, is a very simple minded creature. Its main behavior is to fly back and forth between two points in the world and it is not aware of other characters in the world. Sydney gets distracted no matter what he was doing when this flying creature is nearby.

9.1.4 Cinematographer

There is another character in this project, called Cinematographer or camera character [5]. Although it does not have its graphical bodily appearance on the screen, it does have its own motivation and action selection. Its main goal is to show off the affective states of the main characters and take shots from the best angles and distances to best convey the state of the world to the human observers [132].

9.2 Interaction

From Sydney's stand point, the information that he receives that is relevant to training comes from auditory inputs and visual inputs. Auditory inputs include utterances of the human participant and the clicker sound, and visual input used as part of the training comes from his virtual visual perception of the training stick. Like a real dog, Sydney does not comprehend language and does not have a concept of language. What matters is the acoustic pattern of the speech signal, and thus Cepstral coefficient representation is regarded as sufficient to encode necessary information [6]. It is assumed that Sydney has the ability to distinguish the clicker sound from an utterance due to its distinctive frequency. In practice, this is implemented as two separate channels, i.e., two serial ports, each of which is exclusively used for one of those two

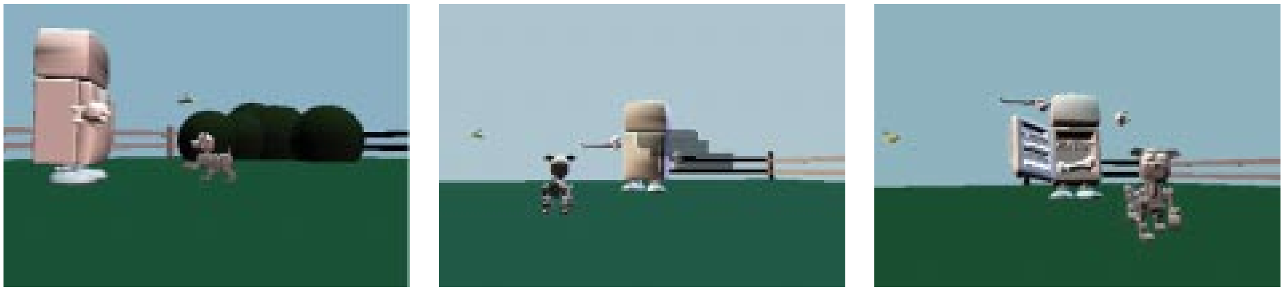
[5] The basis camera creature and the accompanying functionality was implemented by Bill Tomlinson [132] who is also with Synthetic Characters Group, at the MIT Media Laboratory.

[6] This is based on the assumption that dogs use acoustic patterns for distinguishing and matching verbal commands rather than rich representations with the help of grammar primitives such as phonemes or syllabus. The Cepstral coefficient merely takes the log of the power spectrum of the speech signal; it is considered to be a reliable method for distinguishing spectral patterns of voice signals [95]

types of auditory inputs.

Without the human participant's presence, Sydney freely wanders around based on his internal drives and affects aroused endogenously as well as perceptually. At realization of the human's presence, through a speech signal that is perceived as an attentional bid, Sydney runs toward Fridge and waits for a command [7]. Speech signal is either perceived as an attentional bid or a command based on Sydney's state of mind, i.e., what he expects to happen. When he is not expecting an interaction with a human participant, a speech signal that comes in through his auditory channel directs his attention to Fridge, i.e., the human participant, and makes him expect that more of the interaction would happen. Once his attention is on the interaction with the human participant, he considers the speech signal as a behavior command and processes the data to find out whether it matches one of the known command acoustic patterns. As the user's speech command is issued, Sydney does a trick that, he believes, matches the spoken command. If it was right, the human participant shakes the milkbone box in his hand which elicits a milkbone in the virtual world. Sydney runs toward the milkbone and takes it. It reduces Sydney's hunger drive and makes him feel rewarded. No reward is given when Sydney did not do what he was supposed to do. The reward solidifies his belief in the link between the command and the behavior he just performed. When the behavior was not right, he gets no reward which, in turn, weakens his belief in the associative link between the perceived acoustic pattern and the behavior, which becomes less likely to be performed the next time he hears the same acoustic pattern. When there is no input from the human participant for long enough, Sydney loses interest in standing in front of Fridge, waiting for a speech command, and other drives such as curiosity take over and cause Sydney to start doing other things that he is more interested in. Images in Figure 9-3 are screen shots from the **Sydney K9.0** demo and show views of example interactions.

[7] This assumes that Sydney has already learned that staying in front of Fridge is a context that can lead to command-behavior interaction, which then leads to a rewarding situation. Dogs that did not have this interaction before would not show this approaching and waiting behavior without learning, but we have built this into Sydney as if it were an innate pattern.



a b c

Figure 9-3: (a) At perception of the human participant's attention bid, Sydney runs toward Fridge and waits for further interaction. This is because through past interactions, Sydney learned that interaction leads to rewarding, thus it has developed positive attitude toward interaction itself. (b) Fridge is the human participant's graphical instantiation in the virtual world. The motion of a physical training stick, which is controlled by the human participant, is mapped onto the coordinate changes of the virtual stick held by Fridge. The stick movement in the virtual world is perceived by Sydney and it directs his attention. (c) When there is no input from the human participant for a long enough time, Sydney loses interest in Fridge and leaves to do behaviors driven by other motivations.

More on the utterance inputs. There are two challanges regarding implementing a system that responds to a spoken command at a dog's level of intelligence. One of the challenges to the system with this interaction is figuring out a way to distinguish different implications of the speech sound signals that come through the same channel, i.e. auditory perception. Table 9.1 shows an example interaction sequence that may occur in this kind of setting [21]. This sequence demonstrates the fact that dog training is repetitive and requires finding out the functional implication of spoken commands. For example, even though *"Good dog!"* and *"Roll over"* are both spoken by the human participant and perceived through the dog's auditory perception channel, whereas the former marks the end of a behavior and signals a coming reward, i.e. food, the latter conveys the user's intending for the dog to do a certain behavior that matches the command. Without *a priori* knowledge of any of the language or syllabus, dogs manage to find it out and learn to distinguish those two different signals. The implemented system uses the clicker sound to signal reward, instead of verbal approval (although the latter could be implemented in the system also) and nothing is presented when Sydney's behavioral response is not right. How this is implemented is presented in Section 9.3.

Command	Classification	Context
Sydney!	Attentional bid	Dog ignores user and is chasing a butterfly.
Sydney!	Attentional bid	Dog ignores user
Come!	Attentional bid	Dog runs towards fridge.
Here he comes.	Miscellaneous /Other	Ignored by dog.
Roll over.	Command	Dog sits down.
Too Bad!	Discouragement	Dog gets back up.
Roll over!	Command	No response from the dog.
Too Bad!	Discouragement	Dog gets back up.
Roll Over.	Command	No response from the dog.
Roll Over!	Command	The dog rolls over.
Good dog!	Approval	User clicks clicker, rewards dog with bone.

Table 9.1: Within a Sydney K9.0 interaction scheme, the spoken commands are categorized into five groups, *attentional bid, command, approval, discouragement, miscellaneous*, as shown in this table. Different groups have different functional implications, which are figured out through the dog – human participant interaction. This table shows a typical interaction sequence that can be seen in a dog training session.

9.3 Clicker Training

One of the challenges of introducing learning ability in synthetic characters is to maintain a lifelike impression made by the characters on the people interacting with them. While we would like to be able to train characters and personalize them as our friends and companions, we do not want them to passively store knowledge and simply replay it. We want them to learn what they are being taught while staying in character. Even if they are learning, they should have their own drives and affects and show these states through their actions and emotional expressions. This raises the issue of how to train a being that has its own intentions. Pet trainers face a similar problem in their training of creatures that always have their own drives and interests. Among various techniques tried with pets, the clicker training [94, 140] method has proven to be particularly successful, and general enough to be useful for teaching a variety of behavioral responses to pets. Borrowing this idea from pet training, our project implements a clicker training session.

Clicker training is based on operant conditioning and consists of two associative learning processes. By repeatedly giving a reward (i.e., something like a food treat that is motivationally important to the dog) right after sounding a hand-held "clicker" device, the dog forms an associative link between the clicker sound and the positive reinforcer. After multiple repetitions, the clicker becomes a secondary reinforcer that reliably signals that a primary reinforcer will be forthcoming. In addition, the clicker acts as an "event marker" indicating the exact behavior and/or configuration for which the dog is to be rewarded. This solves one of the biggest problems in learning, namely, the credit assignment problem. Figuring out exactly for which behavior it is being rewarded is a huge obstacle for a dog in a learning situation, and the clicker's sharp temporal characteristic makes it easy for the dog to figure it out.

Operant conditioning forms a linkage between a behavior, such as bar-pressing, and a reaction, such as the arrival of food. If the reaction is rewarding, the propensity to engage in the behavior increases [125]. The milkbone provides an unconditioned stimulus that elicits appetitive behavior that leads to chewing and eating behavior. Eating is rewarding, and thus perceiving a milkbone becomes rewarding for the dog. The unconditioned stimulus, food, and the unconditioned response, eating that leads to hunger drive satisfaction, is an innate behavior context chain.

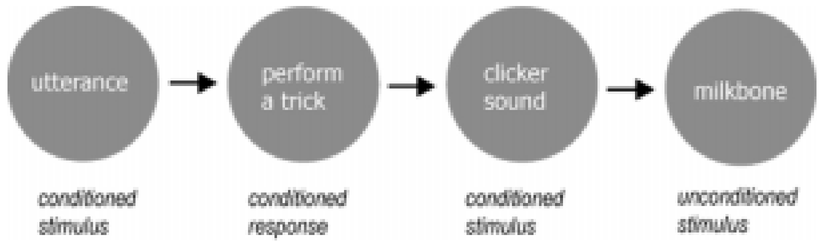
In Figure 9-4 (b), the chain of links marked A corresponds to this. The first thing that a dog trainer does is to form a linkage between the clicker sound and something rewarding. It could be any type of food that the dog enjoys or petting or praise that causes positive affective arousal in the dog. A strong link between the clicker sound and the unconditioned reward stimulus makes the dog regard the clicker sound itself as rewarding, and thus the trainer can use this clicker sound as a reward or positive reinforcer for further training.

This associative link between two perceptual stimuli is marked B in Figure 9-4. This is useful for training because the clicker sound has two properties that are very useful. Firstly, it has a very distinctive sound that is easily discriminable from other auditory inputs. This makes it easy for the learner, i.e., the dog, to discover what it means in the context. Secondly, it can provide a precise mark, signalling the end of a behavior that the trainer has selected, because the click is very brief [8]. For example, verbal praises such as "Good dog" might be used for the same purpose but since the duration of the verbal pronunciation is long, a number of behaviors could be activated while the speech signal is being made, and thus the dog may not get the proper information on what it is supposed to learn. Finally, operant conditioning builds a chain of links labeled C in Figure 9-4. A verbal command is issued and by rewarding with the clicker sound whenever the dog does the behavior it is supposed to perform, it forms an associative link between the command and the behavior which leads to the clicker sound. The sound has a positive reinforcing value due to the expectation of reward that follows. This interaction paradigm can be summarized as Sydney's known behavioral context as shown in Figure 9-4 (a).

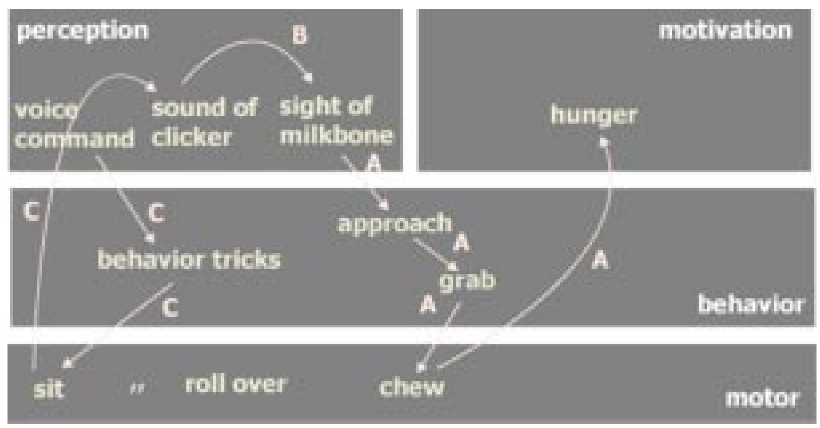
9.3.1 DogEar

The DogEar system refers to a subsystem designed to mediate verbal communication between a human participant and Sydney. It receives the human participant's utterance data, and converts the raw data to a Cepstral coefficient format that is received and stored by Sydney's auditory memory. DogEar samples data at rate of 11025 Hz and captures data that are above certain threshold. Data are averaged over windows of 512 samples. Sampled data are represented as a vector of Cepstral coefficients [95]. To obtain the vocal tract response after removing the pitch ripple, the method of filtering the log-magnitude of the signal with an inverse FFT was adopted. Then the signal coefficients that are beyond the pitch frequency are truncated, and then in the Fourier domain, 10 filters placed linearly on a scale from 100 Hz up to 2kHz are used

[8]To be precise, Sydney does not "know" that the clicker sound says it is the end of a behavior. Instead, since he has an associative link between the clicker sound and the appearance of a milkbone, he shifts his focus, expecting a milkbone reward instead of persisting in the behavior.



a



b

Figure 9-4: **a.** As the result of clicker training, linkages that form a chain of hearing a speech command, perform a behavior trick, hear clicker sound and get milkbone as reward. **b.** Clicker training, effectively, builds in a *context* within Sydney's creature kernel. The chain labeled **A** is an unconditioned stimuli - unconditioned response (US-UR) link that preexists in the creature kernel. During the first phase of the clicker training, an associative link between two stimuli, clicker sound and food, a piece of milkbone in this project's case, is formed, and labeled **B** in this figure. The third associative link chain is formed through an operant conditioning process and this is labeled **C**. The link encompasses the sequence of spoken command → performance of a corresponding trick → clicker sound. The link between a spoken command and a trick is strengthened as the outperforming of the link is rewarded by the clicker sound, which has a positive reinforcing value as it is followed by an unconditioned rewarding stimulus.

for further analysis. See Burke [21] for details.

9.4 Observations and Discussion

Preliminary informal interactive experiences have been provided to experienced programmers, as well as novice users who have never interacted with software creatures before. Participants who had pet-training experience felt comfortable extending their experience to the virtual pet training session. One response we received to a recalcitrant Sydney while adjusting the noise controls in a crowded environment was, "he's acting that way because you aren't treating him gently enough. I know how dogs behave since I have my own pet dog at home." The user went on to (gently) teach Sydney how to roll over on command.

Since the acoustic data processing does not assume a priori knowledge of language or grammar, Sydney's acoustic system has proven robust. At one point during a demonstration, Sydney had learned to respond appropriately to commands in English, French, Japanese and Norwegian. Users were consistently impressed by Sydney's life-like behavior. These observations show the plausibility of the implementation of Dog Ear as a model for an auditory signal processing channel for a doglike synthetic character.

The clicker training method is powerful for producing desired behavior patterns induced by luring and shaping. But the implementation of such interactions heavily rely on the variability of motor actions, which was not the focus of this thesis work. The next obvious step would be to incorporate a more flexible motor system architecture to enable fuller implementation of luring and shaping.

Limitations of this implementation also include the lack of ability to process affect information in the verbal commands. As our participants indicated, real dogs discern contextual information from cues such as prosidy and pitch in utterances. This extension will be implemented by augmenting DogEar to extract these acoustic cues and transfer them to the behavior system.

Chapter 10

Sydney K9.0: How and what does Sydney learn?

As discussed before, the learning in Sydney results in formation of links between nodes in the creature kernel. To be more specific, as shown in Figure 9-4 (b), it is assumed that there is an innate pathway from the perception of food (milkbone) to the appetitive behavior such as searching and grabbing, and then to the consummatory behavior, namely the chewing that gives satisfaction of the hunger drive, an event resulting in reward that comes from the motivation system of the animal. Given this innate structure, the first thing that is learned within the clicker training scenario is the association between the clicker and the appearance of the milkbone. The milkbone has a rewarding value that comes from the satisfaction of the hunger drive, and the reward value gets transferred to the clicker sound as the contingency repeats. Having this link established, then the context of utterance – behavioral response – clicker sound is formed, and this gets attached to the appearance-of-milkbone node through the clicker sound node.

Within the creature kernel framework, learning proceeds as a repeated succession of concept learning and structural learning. Concept learning, involving formation of concepts within the Bayesian framework [1], takes place first in the process of forming the contingency link between the clicker sound and the appearance of a milkbone.

10.1 Review of Bayesian concept learning

This section reviews some terminologies that will be assumed to be known to the reader throughout this chapter. You can skip to the next session if you do not feel a need to review the Bayesian framework, without danger of missing any of the content of this chapter.

Probability $P(h|X)$ is the probability that hypothesis h is a true extension of the concept, given the n examples, $x_1, x_2, ..., x_n$ that we have seen so far. For the purpose of illustration, let's assume that the concept that Sydney is trying to learn is the reliable signal that he can rely on to predict the appearance of a milkbone (AM). Anything that is represented as a node in the creature kernel may be regarded as a hypothesis. For example, events such as movement of fridge (FM), appearance of a butterfly (BA), clicker sound (CS), performance of behavior 1 (PB1) and performance of behavior 2 (PB2) have been happening with a close temporal proximity to the event AM. Those five events are h_i, members of the hypothesis set H with above zero probability of being the reliable cue, i.e., the true extension of the concept.

These probabilities are numbers between 0 and 1 reflecting our degree of belief in h; $P(h|X)$ is near 1 only if we are quite confident that h is the true extension, near 0 if we are quite confident that h is not the true extension, and somewhere in between

[1] In the Bayesian framework, the concept formation process consists of the following four stages [120]. First, a constrained hypothesis space of possible extensions of a concept and a probability distribution over that space, which represents the learner's state of knowledge about which entities a concept refers to, is defined. Second, an informative prior distribution over the hypothesis space reflecting the background and contextual knowledge that the learner brings to this task is considered. Third, the size principle for scoring the likelihood of hypotheses, which favors smaller consistent hypotheses with exponentially greater weight as the number of observed examples increases, is incorporated. Last, the notion of hypothesis averaging is applied, i.e., integrating the predictions of multiple consistent hypotheses, weighted by their *a posteriori* probabilities, to arrive at the probability of generalizing a concept to a new entity.

if we are somewhat uncertain. As probabilities, these degrees of belief are normalized to sum to 1 over the hypothesis space H:

$$\Sigma_{h \in H} P(h|X) = 1.$$

Using the Bayes rule, the probability assignment of $P(h|X)$ is,

$$P(h|X) = \frac{P(X|h)P(h)}{P(X)}$$

and thus depends on the product of the two terms $P(X|h)$ and $P(h)$. The likelihood $P(X|h)$ measures the probability that we would observe examples X if h were in fact the true extension of the concept. The prior probability P(h) measures how probable we think it is that h is the extension of the concept before we have observed any examples. The posterior probability $P(h|X)$ measures our belief in h after we observed the examples X.

P(h) represents the learner's *a priori* belief in a certain hypothesis being the true extension of the concept and it reflects the individual differences – intelligence, experience, etc. For example, capacity of memory is the limiting factor. Omitting a particular hypothesis from the hypothesis space is equivalent to including it by assigning it *a priori* probability of zero.

Now consider the liklihood term $P(X|h)$. To compute the probability of observing the examples in X given that the hypothesis h is a true extension of the concept, we require some assumptions about the process that generates the examples and how it depends on the hypothetical extension. Here, **strong sampling** is assumed where the observed examples of a concept are sampled randomly and independently from the concept's extension. This assumption implies that environmental changes or participants' inputs proceed independently of the learner's understanding of the world. Under this assumption, X denotes a sequence of n randomly sampled examples, and the likelihood of observing this evidence given a particular hypothetical extension h for the concept is simply,

$$P(X|h) = [\frac{1}{size(h)}]^n$$

, if h includes those n objects, and 0 if it does not include one more of them.

The size principle can be seen as a quantitative form of Ockham's razor, "Entities should not be multiplied without necessity." Given the examples $x_1, x_2, ..., x_n$, Ockham prefers the hypothesis with the minimal number of entities necessary to explain their occurence. This is always the set $X = x_1, x_2...., x_n$ itself if $X \in H$.

In the Bayesian framework, we cast the problem of generalization as computing $P(y \in C|X)$, the probability that a new stimulus y belongs to concept C, given the set X of previously observed examples. Formally, the generalization is done by averaging over the set of consistent hypotheses. This can be written as follows. From the conditional independence of $y \in C$ and X, we have

$$P(y \in C|X) = \Sigma_{h \in H} P(y \in C|h)P(h|X).$$

10.2 Concept Learning

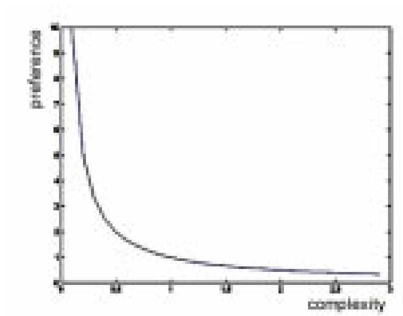
In the creature kernel framework, correlations between entities and dependency and independency evidences are learned through the concept learning procedure. The confidence and beliefs built through this process become the basis of structural learning. In this section, I discuss concept learning and how it fits into Sydney's clicker training scenario.

10.2.1 *A priori*, $P(h)$

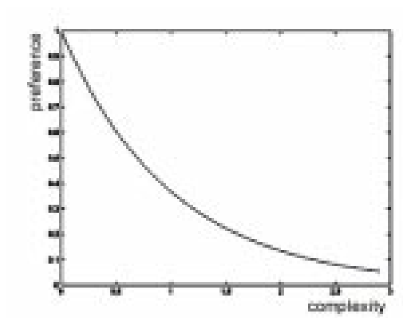
In concept learning within the Bayesian learning framework, an *a priori* distribution represents the learner's knowledge or preexisting belief. The appropriate choice of $P(h)$ when modeling a learner depends on our background knowledge about the learner. The most commonly used *a priori* models are uninformative prior, exponential prior and Erlang prior [58, 3]. Figure 10-1 shows each distribution assuming that the hypothesis h is represented along one axis and its size is s.

Here the preference for a certain hypothesis is represented as a function of complexity, i.e, the description length of the hypothesis. Let's consider the example of Sydney's learning the concept of the reliable cue for predicting the appearance of a milkbone. If he takes the uniformative *a priori*, all hypotheses are equally preferred, for example, two hypotheses, **CS** – a milkbone appears whenever there is a clicker sound – and **FM and PB1 then CS** – a milkbone appears when clicker sound occurs following Sydney's performance of behavior 1 after Fridge's movement – even though the second one is much more complex. On the other hand, if he takes an exponential *a priori,* he prefers the simplest explanation if there is no other relevant information. Taking the Erlang density as *a priori* makes it possible to prefer hypotheses less when they are too simple or too complex.

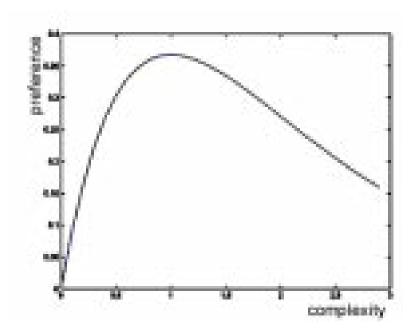
In the implementation of **Sydney K9.0**, the Erlang prior is used with σ being 1.0 to represent the fact that dogs have limited cognitive capacity such that it is difficult for them to consider too many simultaneous events at the same time. For the purpose of illustration, I present a simplified version of Sydney's creature kernel in Figure 10-2. It is assumed that the chain that connects the appearance of the milkbone to the appetitive and consummator behaviors of eating and satisfaction of hunger drive (indicated as **A** in the figure) already exists. Now, as the first pass of clicker training, the clicker sound is made and a milkbone appears, which is a reward to Sydney. Given that various things are happening in the environment at the same time, Sydney initiates concept learning, attempting to figure out a reliable cue for predicting the appearance of the milkbone. For the sake of simplicity of this discussion, assume that Sydney can pay attention to only one thing at a time. Here, the description length, s, is the number of intentional pointers that are concurrently considered to explain a hypothesis. Given the Erlang *a priori* with sigma 1.0 this assumption corresponds



a



b



c

Figure 10-1: In the Bayesian concept learning framework, an *a prioiri* distribution over the hypothesis is taken into account, which reflects the learner's background and the contextual knowledge that the learner brings to the task. Here I show the most prevalently used *a priori* distributions as functions of **s**, the description length, or complexity of the hypothesis. Uninformative prior is the most naive one and is usually chosen when there is no *a priori* reason to prefer one hypothesis over any other. When the expected size of the hypothesis along a certain dimension is known, the apprioriate prior is a maximum entropy density, which takes the exponential function form as shown in (b). Additional information, such as, whether the concept should not be extremely simple or complex, can be taken into account using Erlang density as shown in c. See Jeffreys [58] for more detail.

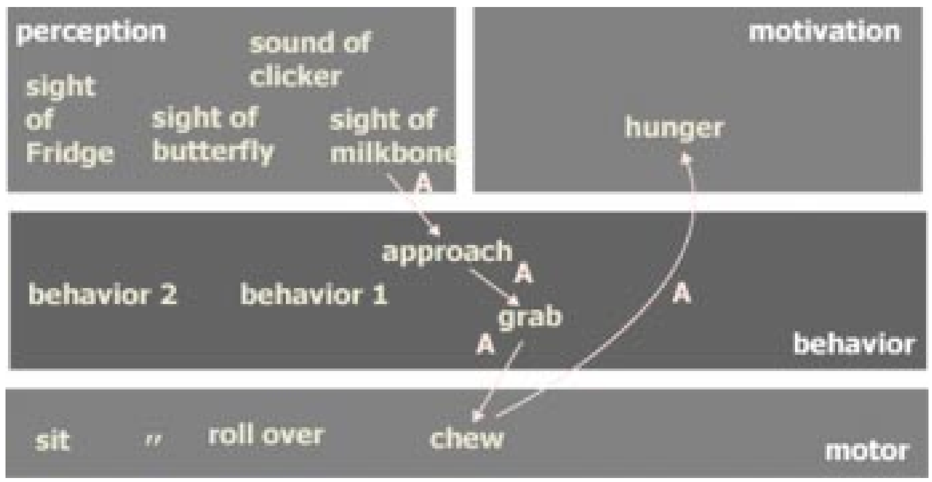


Figure 10-2: This figure illustrates Sydney's creature kernel in a simplified way for the purpose of discussion in this chapter. A few components of each subsystem of the creature kernel are selected and placed. The chain of links, from the apearance of milkbone to appetitive and consummatory behaviors for eating that leads to satisfaction of his hunger drive, is assumed to exist before learning.

s	0	0.3	0.6	0.9	1.2	1.5	1.8	2.1	2.4	2.7	3.0
P(h)	0	0.22	0.33	0.37	0.36	0.33	0.30	0.26	0.22	0.18	0.15

Figure 10-3: For the purpose of concept learning within the Bayesian framework, an Erlang *a priori* distribution is adopted for Sydney. This distribution is based on the assumption that hypothesis description length can be neither too long nor too short. See text for details.

to setting the threshold for hypothesis consideration to be 0.3. See Figure 10-3 for priors with different description lengths.

As shown in Figure 10-2, There are five entities that may taken into account during Sydney's process of concept formation: sight_of_Fridge (movement), sight_of_butterfly, sound_of_clicker, behavior_1 and behavior_2. In the real implementation, the sight_of_Fridge node is expanded to Fridge_moving_the_stick_left, Fridge_moving_the_stick_right, etc, and each of the children nodes of sight_of_Fridge can be considered separately during the concept formation process, but here they are omitted for simplicity of illustration. Given the Erlang *a priori* threshold set at 0.3, the hypothesis space, H, contains only five components, Fridge_moves (FM), butterfly_appears (BA), clicker_sound(CS), performed_behavior1 (PB1) and performed_behavior2 (PB2). All start with *a priori* probability

$$P(h_i) = \frac{P_{Erlang}(h_i)}{\Sigma_i P_{Erlang}(h_i)}$$

, which is the normalized Erlang probability directly calculated from the description length.

The learner, Sydney is modeled as a situated creature looking for the correct hypothesis that activates the sight_of_milkbone node. Figure 10-4 demonstrates simulation results, showing how Sydney builds a concept for given environmental stimuli.

10.2.2 Concept formation

As a learner, Sydney's operation is viewed as a process of alternating forward and backward operations. Forward operation consists of Sydney's getting various stimuli and responding appropriately using his best guess at the situation given the stimuli. His conclusion about a situation and preparation to respond to it appropriately relies on his generalization ability which can be formally represented as

$$P(y \in C|X) = \Sigma P(y \in C|h)P(h|X).$$

Here y is the current set of stimuli and X reflects the past stimuli or events. The term h refers to the individual hypothesis in the set H which is the set of all available hypotheses. And C is the concept, which corresponds to the nature of the hypothesis which H is trying to explain. Backward operation, on the other hand, is the process of updating the creature's beliefs and in this case, the real outcome of the event y is provided and based on its outcome, y is added to X or not and hypothesis h is updated accordingly. For each hypothesis h in H, this update equation is written as

$$P(h|X) = \frac{P(X|h)P(h)}{P(X)}$$

, where X now includes the inputs of past stimuli and the current stimulus that he just encountered. Let us go back to the previous example of Sydney's learning the concept of which is the reliable cue for anticipating the appearance of a milkbone. The hypothesis set includes **FM** (h_1), **BA** (h_2), **PB1** (h_3), **PB2** (h_4), and **CS** (h_5). If what Sydney just observed x_n was the fact that there was a clicker sound and then a milkbone appeared, he updates his example set X to X', which now includes x_n, and the belief in $P(h_5|X')$ becomes higher than $P(h_5|X)$, whereas $P(h_j|X')$ becomes lower than $P(h_j|X)$ for all all $j \neq 5$.

The change in Sydney's belief is shown in Figure 10-4. Given this simple creature kernel, there are five concepts which are taken into consideration. They are FM, BA, CS, PB1 and PB2 as described above. Since the description length for all these five are the same, the *a priori* probabilities are the same for all these. The state of the world is represented as a combination of binary numbers such as (1,0,0,1,1) to say, for example, that Fridge moved, and Sydney did behavior 1 and behavior 2 but butterfly did not fly by and there was no clicker sound. This is shown in the first column of Figure 10-4(a). For this illustration, the event sequence was generated randomly. The binary flag right next to the first column (existence or absence of a vertical bar) shows whether the milkbone appeared or not. It is set to one if CS is one. So, the right answer for Sydney would be figuring out that it is one when CS is one. The

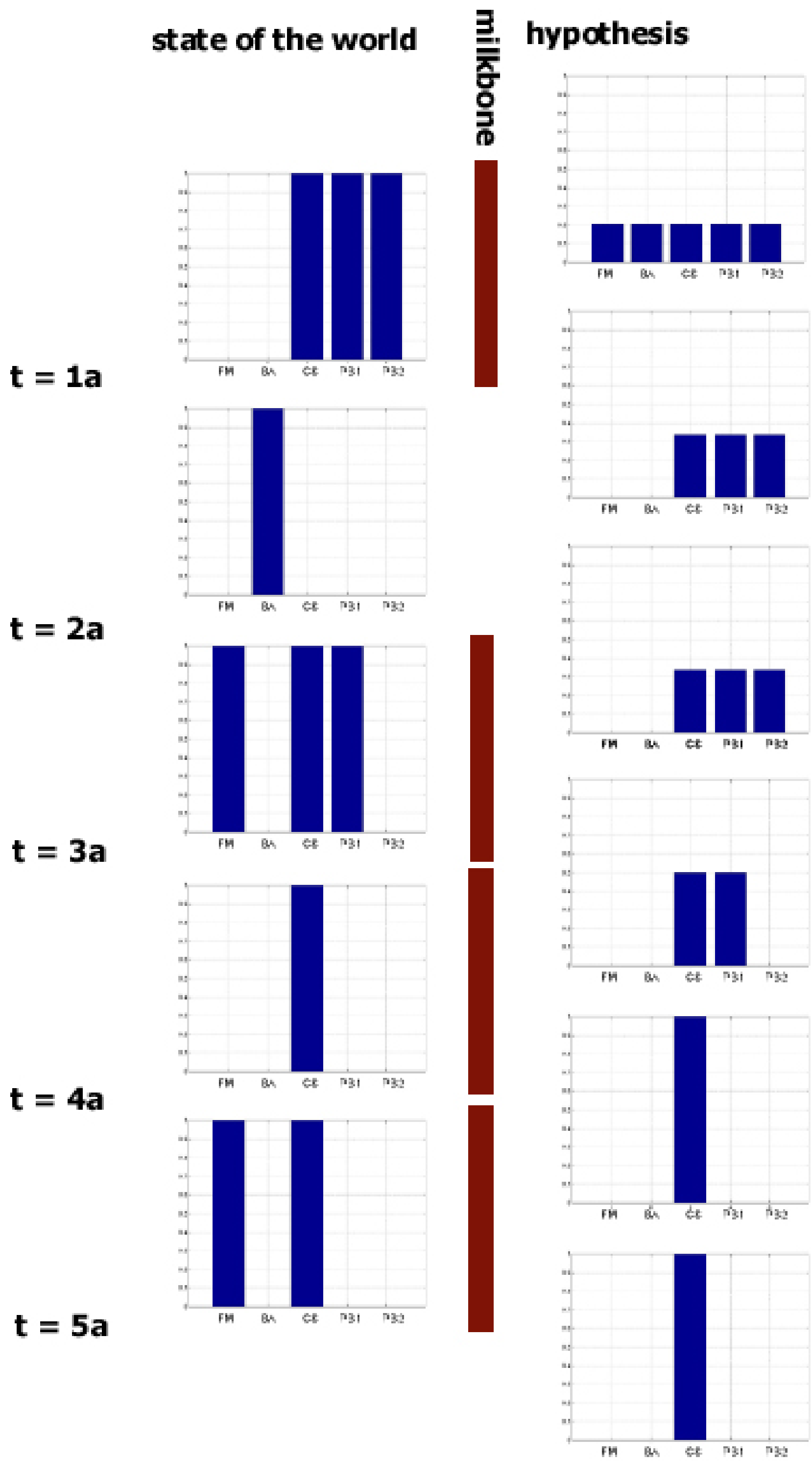


Figure 10-4: Five hypotheses with description length 1 are considered here for purposes of simple illustration. The first column represents the state of the world at each event time, and the existence of a solid bar in the second column shows that there was a milkbone reward. The third column shows progressive change in Sydney's beliefs in each hypothesis. See text for details.

third column represents the progressive change of Sydney's belief in each of the five hypotheses. It is clear that Sydney starts with no particular preference for any of the possible candidate concepts – since the description lengths are all same [2] – but as time proceeds, he becomes more and more confident that the third hypothesis, that the sound of clicker is a reliable cue, predicting the appearance of milkbone. Confidence is not updated when the event of interest, the appearance of milkbone in this case, does not happen as shown in the case of $t = 2a$.

Figure 10-5 illustrates the case where a strict prior threshold is not used. For simplicity of the illustration, it is assumed that Sydney is only interested in three perceptual components – FM, CS and PB1. As he forms the concept for explaining the appearance of the milkbone, combinations of two or more of these components are all considered, but with different *a priori* probabilities as shown in the first bar graph of the third column in Figure 10-5; the number of hypotheses in consideration is seven in this case. The convention of the illustration in Figure 10-5 is the same as that of Figure 10-4 and the inputs are generated randomly.

The above two demonstrations assume a perfect world where there is no aliasing in sensing and the resources such as attention are not limiting components for the operation. Among many other possibilities, two stochastic components are taken into account in the actual simulation. One is the occurrence of inconsistent inputs and the other is the limited processing capacity. Extension of this treatment demonstrates phenomena such as superstitious behavior when certain phenomena coincidently keep happening together with a reliable cue, perhaps because they have a very low threshold for detection, so that Sydney might keep his confidence in such hypotheses high and believe in the false coincidental associations.

[2] This is a naive view taken for the sake of computational simplicity. Species or individual specific preferences do exist in the case of real animals.

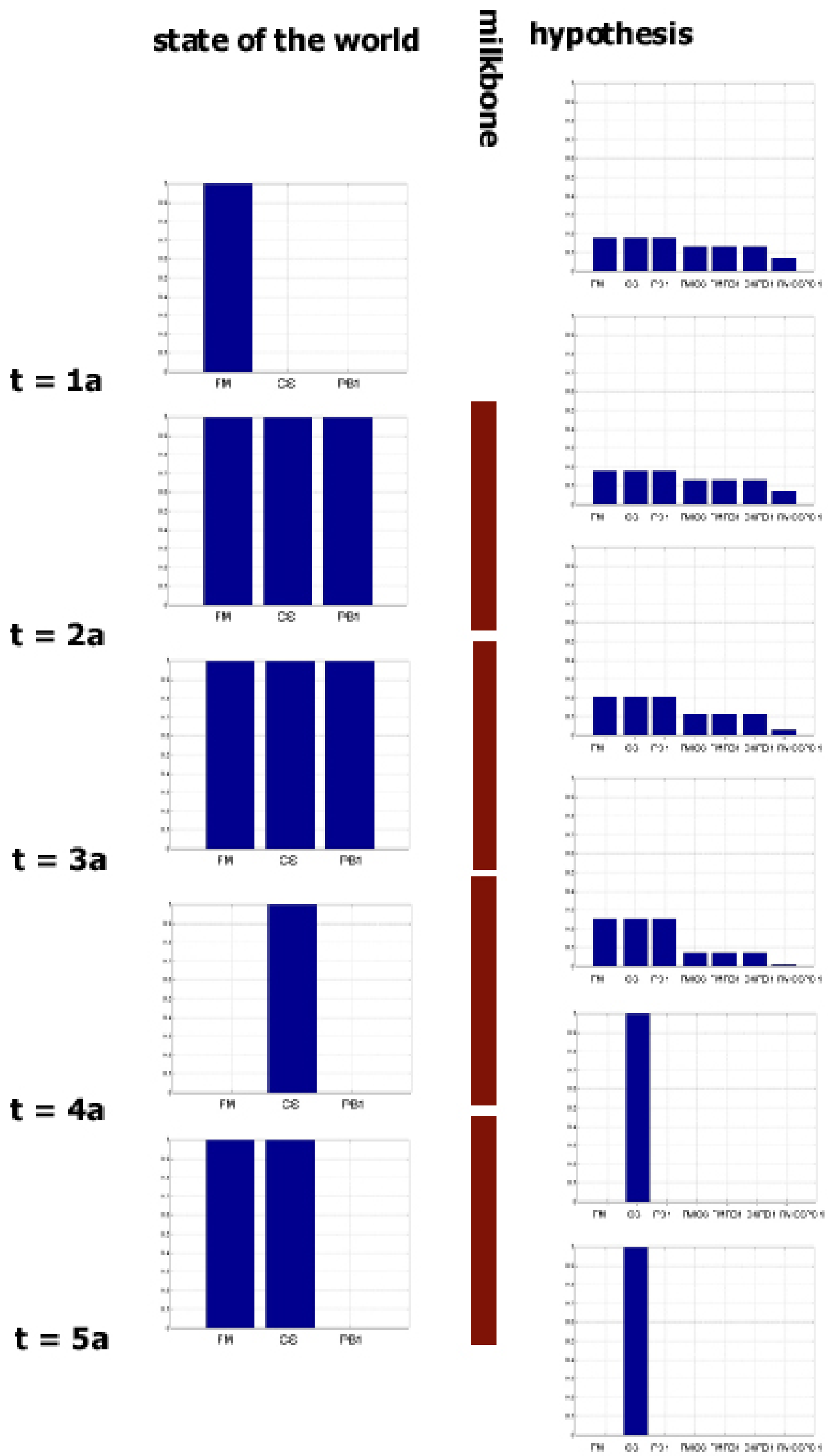


Figure 10-5: For illustrative purposes, only three conceptual components are considered as constituents of a hypothesis. Allowing the description length to be longer than one, seven hypotheses are considered, where *a priori* probabilities are different for each hypothesis. As time progresses, evidence is collected (shown in the left column) and Sydney becomes more confident about one hypothesis versus the others. See text for details.

A note on *time* Many aspects of learning require memory, which in turn brings the concept of time into consideration. The concept of time includes two different usages: one is the time as a pointer and the other is the time as a measure of duration or passage. Examples of the former usage, where time is regarded as (rather) an absolute measure, are 'event A happened when event B happened (here, *when event B happened* is a pointer to a certain point in time),' and 'event A happened *at 3:45 pm on January 24th in year 4821* (here the pointer to a certain point in time is specified by a system of units which has been socially accepted).' The latter usage regards time as a relative measure. Using this meaning of time, statements such as 'event A happened after a certain duration of *time* passed since event B happened' or 'I am hungry since it has been *12 hours* since I had my last meal' can be made. Here, the temporal difference between when two events happened (event A and event B in the first example, and now and when the speaker had the last meal in the second example) is the subject of consideration.

In this system, time is not at all assumed in an absolute sense. There is no system of measure, by which characters can tell 'it already is noon!', for example. However, the relative sense of time, in the subjective manner, has been incorporated. Under this assumption, characters have a sense of the passage of time in an indirect manner. For example, the drive level within a character progressively changes over time and the character can realize that the drive level has changed even though it cannot tell how much time actually has passed since the last realization of the drive level. This affects learning in a similar way. For example, in the case where the associative link between the clicker and milkbone is formed, the temporal difference between the appearance of two incidents affects the strength of the associative link, and the passage of time is measured either in terms of the number of distracters that happened in between, in the backward process, or the change in the level of expectation in the forward process.

Another way the characters can tell the passage of time is through the number of distractions that happened between two events (as briefly mentioned in the last sentence of the above paragraph). The concept of *action time* is adopted here. For example, a character has a certain belief and tries the adequate behavioral response *for a while*, but it gives it up after it thinks that *enough time* has passed. Here, *"enough time"* is measured by the number of action trials made. Some actions take more time than others, so a *five action time* period may mean a very different temporal length in the absolute sense depending on which action that the character is measuring with. But they are considered to raise the fatigue or tiredness level a similar amount, so action time is useful in the system where the passage of time is measured in the relative sense in terms of drive or motivation changes.

From the software standpoint, another timer worth mentioning is the time that the world keeps for updating objects. The tick rate increases with the system speed, and decisions on items that require constant updating such as whether new sensory sampling should be made, are made at every tick.

10.2.3 Inconsistent inputs

Inputs may be presented inconsistently, for example, a milkbone is given to Sydney without a preceeding clicker sound or only the clicker sound is made but food is not given. The learner should be able to tolerate such variations. In real world situations, this variation in inputs may occur partly because the world itself is not perfect or the learner might miss some changes due to his limited sensing capabilities.

In the implementation of Sydney, this problem is dealt with when the confidence in hypotheses is updated. For each hypothesis, violation of a certain hypothesis by a current stimulus does not lead to a complete rejection of the hypothesis, even though this complete rejection approach may be valid in some other situations. For Sydney, updating of confidence in hypothesis h_i, the following formula is adopted.

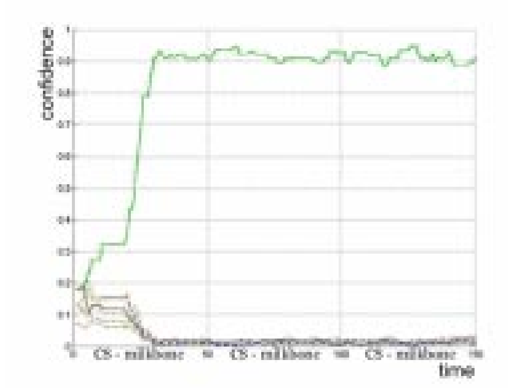
$$P(h_i|X) = \frac{\alpha \cdot P(X|h_i)P(h_i) + \beta \cdot P(h_i)}{\Sigma_i(P(X|h_i)P(h_i) + P(h_i))} + \sigma.$$

Here α and β are free parameters that can be adjusted to assign appropriate momentum to the old belief, and σ is a stochastic term that keeps alternatives from being completely rejected [3]. In effect, this mechanism also ensures that Sydney will not become too fixed in his behavior and will continue to sample alternatives that may be useful in the future as the environment or other conditions change [4].

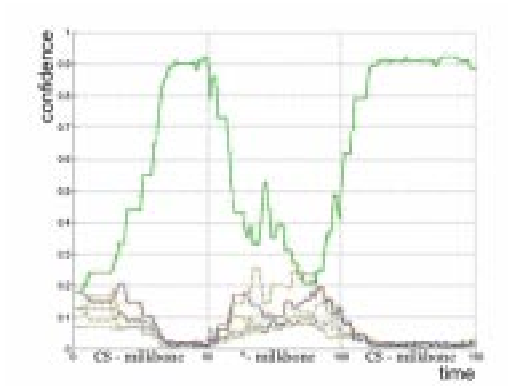
Figure 10-6 illustrates a few simulation runs. In all of these simulations, both α and β are chosen to be 0.5. In all the cases shown in the figure, the hypothesis set H is the same as the case shown in Figure 10-5, where seven alternative hypotheses are considered with different priors. All the graphs in Figure 10-6 contain confidence profiles for all seven hypotheses over time. In particular, the solid green line represents the confidence in the hypothesis CS and the solid red line represents the confidence in the hypothesis FM. Figure 10-6(a) shows the case where all inputs were consistently given, such that a piece of milkbone was given after the clicker sound. After about 30 time units in the experiment, which corresponds to 17 presentations of the milkbone, on average, Sydney builds up confidence in the hypothesis CS, above 0.9. In the case of Figure 10-6(b), after initially consistent CS-milkbone presentations, during the time interval between 50 and 100 the milkbone was given randomly with independent probability 0.5. After the time unit 100 is passed, the milkbone was given again consistently after the clicker sound. Not immediately, but gradually the confidence in hypothesis CS decreases and the situation becomes such that any one hypothesis is not preferred above the others. Then the confidence in hypothesis CS builds up again as the consistent inputs are presented. In the third case, Figure 10-6(c), after the consistent presentations of CS and milkbone during the first 50 time intervals, the milkbone was consistently given after Fridge moved. Consequently, the hypothesis FM becomes gradually more preferred over the hypothesis CS. After the stimulus

[3] These parameters are preset in the current implementation. Systematic update of these parameters reflecting the amount of gained knowledge and experience is another problem that needs to be looked into.

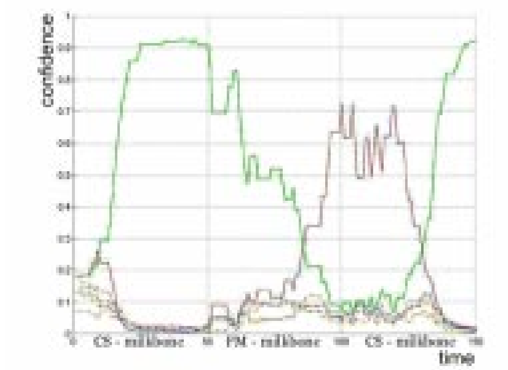
[4] There are some striking exceptions to this. In his descriptions of various learning phenomena, Lorenz [68] classifies such distinctive phenomena as special kinds of learning, such as *imprinting*.



a



b



c

Figure 10-6: While inputs were consistently presented, i.e., clicker sound then milkbone in **a**, inconsistent inputs were provided in cases **b** and **c**. In case **b**, up to experiment time 50, the milkbone was consistently given after the clicker sound. But during the time interval between 50 and 100, the milkbone was given randomly with independent probability 0.5. After the time unit 100 is passed, the milkbone was given again consistently after the clicker sound. In the case of **c**, the initial 0~50 time interval was the same as in the other two cases, but during the time interval 50~100, the milkbone was consistently given after Fridge moved. In all three figures, confidence profiles for all seven hypotheses in the hypothesis set are plotted over time. In particular, the solid green line represents the confidence in the hypothesis CS and the solid red line represents the confidence in the hypothesis FM. See text for details.

presentation schedule is switched back to CS followed by milkbone, confidence in each hypothesis switches back to its former state.

10.2.4 Focus of attention

As the number of components of consideration increases, the number of hypotheses grows exponentially. This growth cannot be dealt with in practice and a focus-of-attention module handles this problem by shifting the focus of attention. In the case of Sydney, a fixed number of hypotheses, less than the capacity limit, ordered in terms of preferences, are considered at every concept learning instance. In addition to that, stochastic choice is made among other possible hypotheses, which are less preferred, due to their complexity in description, or the fact that their confidence got lowered based on prior experiences. In effect, it is possible but takes more time for Sydney to learn more complex concepts, for example, ones with longer description length.

10.3 Structural learning

Having formed the correlative link between the clicker sound and the milkbone, clicker training results in Sydney preferring certain behaviors more than others and then doing those preferred behaviors with the occurrence of the vocal commands. Sydney's learning the right behavioral responses and thus expanding the scenario to a full context which includes the hearing of a verbal command, the performing a trick, hearing a clicker sound, getting a milkbone, and finally eating with consummatory actions, is done through the process of concept formation, and then updating the appropriate part of the subsystem network.

Once the clicker sound is correlated strongly enough with the appearance of a milkbone, the clicker sound itself has a positive stance value that comes from the hunger drive. In other words, when Sydney hears a clicker sound after learning the association between the clicker sound and the milkbone appearance, hearing the sound of the clicker is perceived as being rewarding for the behavior that he just performed, by an amount that is proportional to the multiplication of the connection strength between the milkbone appearance and the clicker sound (which corresponds to the value $P(a\_milkbone\_appears|clicker\_sound)$ and the current level of the drive unit in the motivation system that the clicker sound is pointing to, which is the hunger drive). This is because the rewarding value of the clicker sound is coming from the appearance of a milkbone, and the milkbone's rewarding value is coming from the satisfaction of hunger drive. The next thing that a trainer can do is to watch the dog's behavior and make a clicker sound whenever the dog performs an interesting skill that potentially can be used as a part of a trick. Initially the dog would not know what triggered the clicker sound, but having the performance of behavior as the entity of the concept space that Sydney searches through, he gets to "figure out" the behavior he performed and raise the value, i.e. the intrinsic probability, for performing that behavior since it elicits the clicker sound which is rewarding. This approach can be extended to rewarding a certain combination of behaviors, to teach a preferred behavior sequence rather than a single action, but this will take longer to teach because it is a more complex concept for Sydney based on the fact that the complexity measure we are using is still the description length [5]. Also, here it becomes obvious why a temporally short clicker sound is effective for training. If the marking is precise and short and the clicker sound is made right after one behavior is performed, the search space for Sydney gets smaller and it is easy for him to figure out a reliable concept. These observations correspond to real dog trainers' experiences. A similar procedure proceeds to form the chain of links between a certain pattern of spoken sounds and a behavior trick. In the creature kernel framework, the learned associations are coded in the connection weights among the nodes that constitute the subsystems.

[5]Real trainers accomplishes this task by teaching each behavior separately then fading stimuli. Alternative approach in this system would be rewarding the behaviors that are part of the behavior sequence that is going to be taught so that the probability of performing those behaviors than teach the chaining.

10.3.1 Teaching behavior tricks

In this clicker training framework, there are three different kinds of interactions that enable teaching a dog behavior tricks. Once the associative link between the clicker sound and reward is formed, the clicker can be conveniently used for training since it still carries the intrinsic value that comes from the link. The first interaction is making a clicker sound whenever the dog performs a behavior trick that the trainer intends to teach. This is perceived by the dog as a positive reinforcement for performing the behavior and the general preference given to that behavior increases. The second kind of interaction is one that leads to chaining of behaviors. By rewarding whenever a certain combination of behavior is performed, in proper order, associative links between behavior nodes themselves are formed and from the perspective of the creature kernel's behavior system, it is seen as a known skill of the creature. Finally, whenever the dog seems to be initiating the desired behavior, by issuing a vocal command, the trainer can have the dog proceed to completion of the behavior and then make the clicker sounds so the dog can learn the **context** of vocal command then behavior then reward.

Even if the dog already has the knowledge of the context, if more than one command is to be taught, he needs to learn to discriminate different vocal patterns and learn the corresponding behavior link from the acoustic pattern. In **Sydney K9.0**, these types of learning all happen simultaneously as a process of repeated concept formation and update in weights. Figure 10-7 illustrates the context forming process. Once Sydney has learned the useful context (voice command $\rightarrow$ performing a trick then reward) he needs to learn which among the known tricks is the behavior matching a certain voice command.

Let's assume that there are four known tricks as shown in the figure (Figure 10-7), and the four tricks form the hypothesis set for this learning. *A priori* probabilities are proportional to the probability of occurrence of the clicker sound that Sydney received when he performed each behavior, possibly outside of the voice command

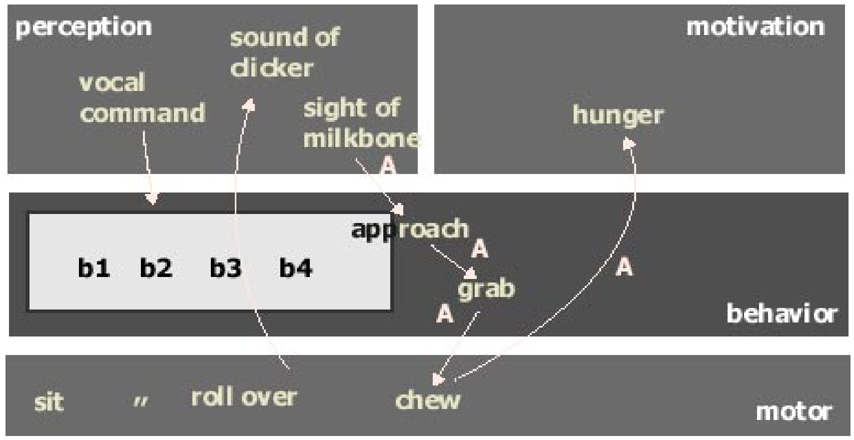
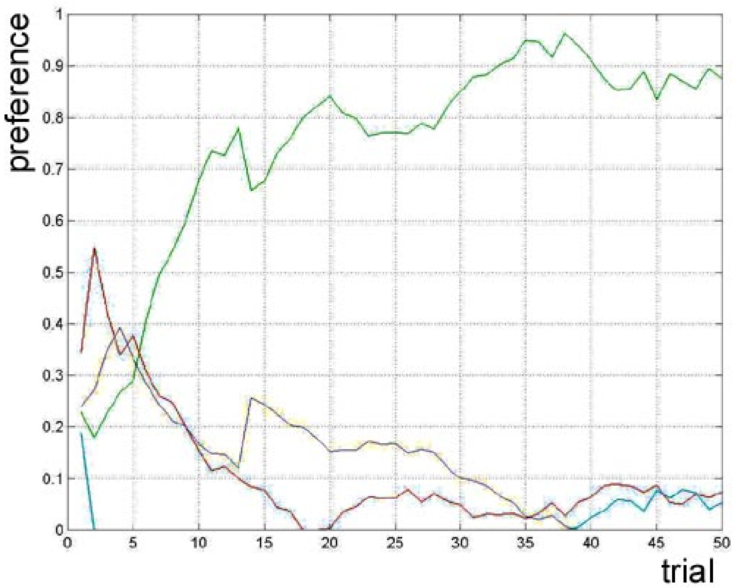


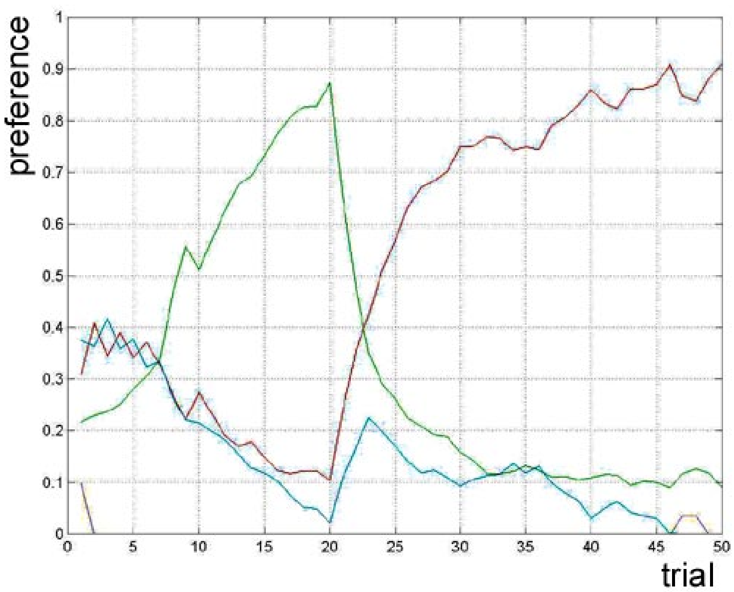
Figure 10-7: This figure shows the creature kernel view of Sydney's context formation process. Having the knowledge that performance of certain behaviors bring the clicker sound and then reward, and this happens in the context of a voice command, he still needs to learn to discriminate the meaning of different acoustic patterns, i.e., respond to a certain command with a certain type of behavior.

context. The formation of context proceeds within the operant conditioning framework where Sydney performs various behaviors in search of the right trick and the completion of the context, and he is rewarded for right behavior. His doing a behavior corresponds to the forward operation part of the process and the probability for doing one of the four tricks under consideration reflects his confidence in that behavior, i.e., $P(y \in C|X)$. Here y is one of the four tricks and concept C can be expressed as *"it is the right behavior that will complete a valid context* [6]*,"* and X is the past experience and gained knowledge including *a priori* beliefs. Through the process of being rewarded for doing a specific behavior, Sydney builds his confidence that one of the tricks will bring reward in the presence of a certain voice command. This confidence is coded as the strengthened connection weight between the auditory memory that stores that acoustic pattern and the corresponding behavior trick. Figure 10-8 shows the behavior preference profile in this situation. In the presence of a certain voice command, one of the four behavior tricks was rewarded for being performed and over time, the preference for that behavior got higher compared to the others. The change in preference for the rewarded behavior is shown in Figure 10-8 **a** in the green line.

[6] Here, valid context would be a context where Sydney knows from his experience that its completion would lead to the consummation that he desires.



a



b

Figure 10-8: These two figures show the change in Sydney's preference given to four behaviors in the trick learning situation. **a.** in the presence of a certain voice command, performing one trick is consistently rewarded and Sydney gradually builds up higher preference for that trick than the others in that context. **b.** When the rewarded behavior is changed after a while, Sydney can change his preference to the new behavior which is now being rewarded. In this case, one behavior was rewarded during the first 20 trials, then another behavior began to be rewarded instead during the rest of the session. The preference change profiles are shown in green and red lines, respectively.

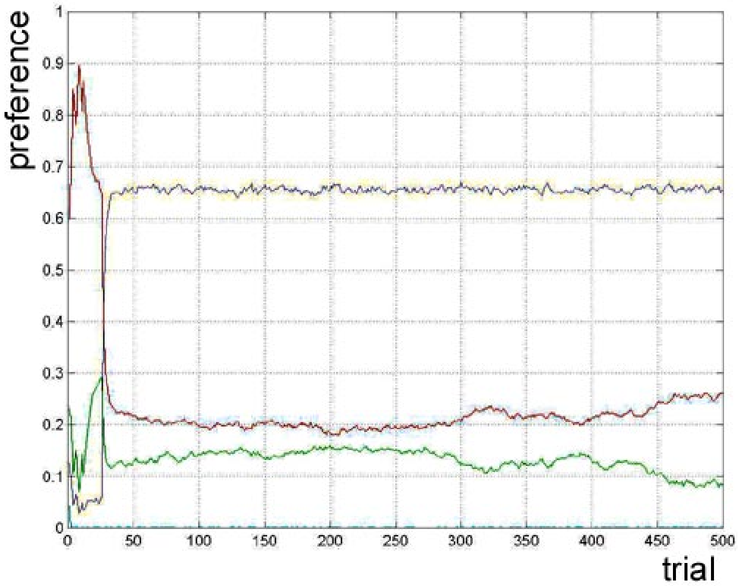


Figure 10-9: Two behaviors were rewarded probablistically in this simulation. One behavior was rewarded for performance in the given context 70 percent of the time, and the other behavior was rewarded 30 percent of the time. The resulting concept shows that Sydney prefers those two behavior proportionately. Preferences to other behaviors remain above zero due to a stochastic component parameter that was set to 0.08. This ensures that Sydney does not give up any of the hypotheses completely.

Sydney can change the concept, and thus the relative preferences given to behaviors. In the simulation result shown in Figure 10-8 b, Sydney was rewarded for performing one behavior during the first 20 trials (the preference profile for the rewarded behavior during this period is shown in green), then another behavior was consistently rewarded afterwards during the rest of the trial runs (the preference change for this newly rewarded behavior is shown in red).

10.3.2 Probability matching

When a behavior is probabilistically rewarded, the resulting concept reflects the reinforcement pattern as shown in Figure 10-9. In this case, one behavior was rewarded 70 percent of the time and another behavior was rewarded 30 percent of the time. The resulting concept shows that Sydney prefers those two behaviors in proportion to the reinforcement frequency. Such matching phenomena have been reported in a number of animal behavior studies [113].

Bitterman reports [88, 5] that in the case of real animals, if the reward is provided with a certain probability, depending on species, some animals match the probability of the behavior or choice with the probablity of the reward or max out, i.e., always choose the case that provides the most reward. Once the concept on the likelihood of the appearance of the reward is formed, this behavioral response can be implemented by corresponding action selection mechanism. If the action selection mechanism is a liner process which forces the animal choose a certain action or make a decision in proportional to the strength of the learned associative connection (Thorndike's law [5]), the animal would show the matching behavior. Whereas if the animal takes the concept and instead of directly applying it to the action selection mechanism, if it anticipate the expected reward and compare it [129], as humans are often thought of doing, it will show the behavior that always chooses the case which gives the larges reward.

10.4 Finding a dog's mind in Sydney

There are various phenomena in addition to those described above, that are reported by people who have experience in dog or animal training. The reports are propagated from mouth to mouth as bits of wisdom to help people better understand animals and get better at training them. In this section, I explain how such phenomena can be explained in the creature kernel based learning frramework.

10.4.1 Blocking

A dog can be taught, for example, to roll over at the hand down gesture. Or a dog can be taught to roll over at the voice command, "roll over." However, once the dog is taught to do the behavior in one way, even if the other cue is consistently presented afterwards, the dog does not pay attention to the other cue, i.e., it does not learn the correlation between the newly presented cue and the appropriate behavioral response.

For example, once the dog has learned that the hand_down gesture is a reliable cue for performing the roll over behavior to get a reward, even if the trainer starts saying 'roll over' consistently while doing the hand down gesture, the dog does not learn that the voice command is a reliable hypothesis predicting that performing the roll over behavior will lead to reward and consummation [7]. In the concept learning framework, this is explained very concisely. During the backward belief update process, which can be written as $P(h_i|X) = \frac{P(X|h_i)P(h_i)}{\Sigma_i P(X|h_i)P(h_i)}$, once $P(h_i)$ is high and outcome y is still consistent with this hypothesis (i.e. the hand down gesture is consistently shown), there is no reason why the creature should prefer other hypotheses over this one. This is shown in Figure 10-10. In this simulation, the hypothesis set is the same as the one considered in Figure 10-4. There are seven hypotheses in this set. During the first 50 time units, CS was presented right before the milkbone. From 50 to 100 time intervals, Fridge moved whenever the clicker sound was made. Then, the stimulus type reverted back to the situation where only the clicker sound was made before the milkbone was given as a reward. As shown in the figure (Figure 10-10), once the concept that the clicker sound is a reliable cue for the appearance of the milkbone, the other variation in the presentation did not affect the confidence levels. The clicker sound hypothesis is shown in a green line that approaches one over time, and the other hypotheses show a trend of converging to zero [8]

[7] In the current implementation, even if hand_down gesture happens before the 'roll over' command, Sydney would not learn that the temporal sequence of hand_down and roll over' is the right stimulus if the already learned 'roll over' consistently happens and is enough to predict that the timing is right to perform the roll over behavior. In real animals, only the concurrent patterns show the blocking phenomena. Even if a dog learned that 'roll over' command tells him that it is time to perform the roll over behavior, if the hand_down gesture begins to always happen before the 'roll over' command, it will expand the reliable stimulus to the sequence of hand_down then the voice command 'roll over.' Also, in the system, a temporal pattern is stored as a form of sequence in the associated memory. How long this sequence can be depends on the short-term memory capacity that we assign to each creature.

[8] In the figure, the confidence level approaches 0.92 instead of 1. This is because there are some stochastic components added to keep the probability of eliciting all the other behaviors from being zero, and the probabilities have been normalized to sum to one over the behaviors in the same behavior group. Characters could not learn that some alternative hypotheses are better when the situation or environment changes, if they completely lose the chance to try them out.

10.4.2 Jackpotting

Schedules of reinforcement affect learning. In the literature of animal learning, there are a number of research reports discussing how various factors such as frequency, size and delay of reinforcement affect the performance of the animal. There are some consistent trends when these factors are combined, for example, animals usually work harder for smaller and immediate rewards than larger and delayed rewards. However, recently pet trainers have realized that a reward procedure called *jackpotting* has very useful effects. Their claim is that instead of a small bit of food, an unexpected large reward is much more effective at inducing the animal to work harder to learn.

There is not much in the operant research literature to support this [20]. But within our creature kernel framework, it can be explained by the fact that within the motivation system, the affect system plays a crucial role in learning by providing reinforcements. In our own experience, unusually large rewards elicit a very strong affective arousal, and those events are stored as special memories as we remember some events with strongly emotional arousal more vividly for a longer time compared to the usual memories of facts and normal events [40]. In the creature kernel, this phenomenon can be implemented by adding a special affective tag to an event when it elicited an affective feedback above a certain threshold, and having this tag refreshed by further experiences. This will strongly affect the creature's action selection and decision making.

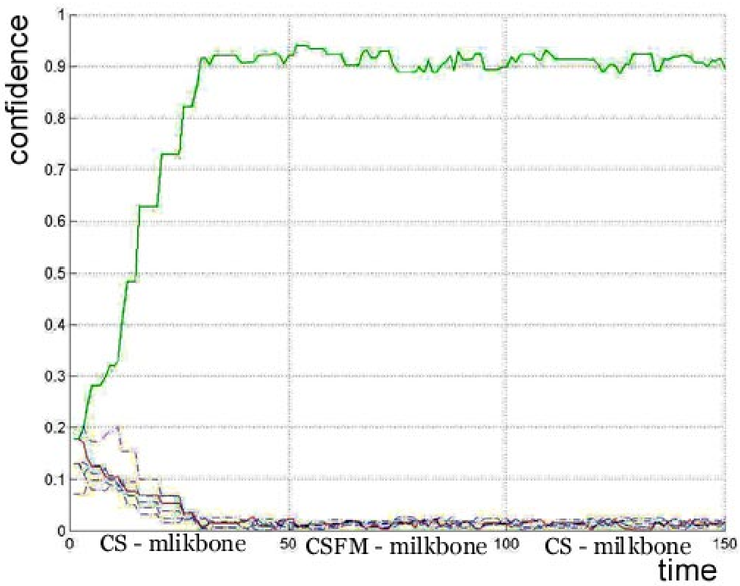


Figure 10-10: Once a concept is learned, if its cue keeps consistently being presented, the extraneous cues are not learned. This phenomenon is called blocking, and the concept learning adopted here explains the phenomenon. See text for details.

Chapter 11

(void): A cast of characters

(void*): A cast of characters implements a dining hall setting. There are three main characters - *Earl the Trucker*, *Fast Eddie (Dude)* and *Elliot the Salesman* (see Figure 11-1 (a)) and each of them can interact with a human user through a buns-and-forks interface (Figure 11-1 (b)). A human participant can "possess" a character by putting the buns-and-forks on one of three plates between her and the screen; each plate represents one of the three characters. This results in a "POSSESSED" signal being sent to the corresponding character. As an acknowledgement of the possession, the character gets up and walks toward the center of the diner hall. From this point on, the character's behavior system receives a strong influence from the interface until it is un-possessed.

11.1 Characters

The three characters have very distinctive and strong personalities. *Earl the Trucker*'s basic emotional state is anger; he represents a conservative character who has a very negative attitude toward dancing. *Fast Eddie*, also called *Dude*, is a happy character. He is interested in doing novel and cool things. Though he does not expect to be possessed before he has any experience of possession, he generally has a very positive attitude toward dancing. And when he figures out that possession leads to dancing, he comes to look forward to being possessed. *Elliot the Salesman*'s underlying emotional

(a) (b)

Figure 11-1: **(void): A cast of characters** implements a diner setting. (a) There are three distinctive characters sitting at the counter. (b) A human participant can come in and possess one of the three characters using a wireless interface mimicking buns-and-forks from Charlie Chaplin's movie *Gold Rush*. From then on, he can force the character to dance in a certain way by wiggling the interface in the way he wishes the legs of the possessed character to move. Characters respond to this possession and forced dancing differently, based on each character's personality, past experience and motivation.

state is fear and nervousness. He is willing to do whatever he is forced to do by authority. Through repeating experiences, he gradually builds a certain attitude toward possession and the dancing experience associated with the possession. Depending on how the user controls the interface, characters can have a fun time dancing or they can have painful experiences falling down on the floor. Happy dancing experiences act as positive rewards for the characters and cause them to look forward to being possessed, whereas too much painful experience makes the characters build up a negative attitude toward dancing and possession, and eventually this could make them upset and walk out of the hall, not listening to the signals from the interface.

In other words, based on how the human participant controls the interface, the possessed character may have fun dancing a certain type of dance that it likes, or it may keep falling down on the floor and have a painful time. This emotional experience is fed back to the character itself, and its attitude toward dancing is updated so it may come to differ from the character's original attitude. The current attitude toward possession and dancing is indicated by its emotional expression. Even if the character is doing the same LEG_LIFT following the human participant's one bun lifting signal, it could do it full of joy or with resistance or hesitance depending on whether the character is enjoying the dance or experiencing too much pain doing it. It may reject being possessed and walk out of the dining hall when the human participant is causing too much pain. If it receives acknowledgement of un-possession, it gains the freedom to act according to its own will. It may go back to its seat or keep dancing. When another character is possessed, it pays attention to the possessed character's behavior and watches that character's dancing and reactions to the human participant. It expresses cheer or sympathy depending on how the possessed character is doing.

11.2 Learning

All these personality based attitude changes and learning take place in the motivation based learning framework implemented around each character's creature kernel. Different kinds of learning phenomena observed during the interactions are as follow.

11.2.1 Organizational Learning

Once a character is un-possessed, it may continue dancing or go back to its seat by its own free choice, since there is no longer any external input which dominates its behavior system as occurred while it was being possessed. The strength of the desire to dance, which is determined by the overall affective feedback from the previous dancing experience, influences the character's decision between going back to the seat or continuing to dance. This attitude or desire to dance is a modifiable parameter that is learned through possession and dancing experience and implemented as the preference learning part of organizational learning.

Upon the decision to keep dancing, it tends to repeat the type of dance that it enjoyed during its previous interaction session. This is because, in its behavior system, the dance types that gave the greater fun to the character gained a higher preference value compared to others and were dynamically assigned as children nodes of the autonomous dancing behavior parent node. This shows an example of strategy learning. A session of free dancing gradually reduces the desire to dance because of fatigue and consummation of its desire, and eventually the character goes back to its seat.

11.2.2 Concept Learning

The main concept that a character builds through an interaction experience is the attitude toward dancing and possession. In this particular installation, the data that these characters collect are the motivational feedback that they experience during the interaction - dancing session. This, in turn, influences the character's attitude toward dancing which then alters the drive system part of the motivation system, changing the desire to dance and the desire to interact with the human participant. The updated attitude toward dancing constitutes a change in the affect part of the motivation system of the character, which is displayed as its emotional state through facial expression and modulation of its motor system at the next time tick.

11.2.3 Affective Tag Formation

Characters have the ability to choose where in the diner to dance. They can choose to dance slightly to the left, to the right or in the middle. This decision is made very quickly, from a combination of personal preference, physical proximity to the point and past experience. The past experience is coded as affective tags attached to locations, which influence the character's decision when it dances the next time. For example, a bad experience in the right corner of the diner makes the character avoid that corner, whereas having lots of fun in one corner increases the probability that the character will choose that corner in the forthcoming session.

11.3 Observations and Discussion

Tight coupling among the four systems that constitute the creature kernel made it possible to build characters that show their affective states transparently. In our implementation, the motivation system continually sends modulating signals to the motor system so the character behaves in the proper emotional way, as signals to the behavior system update the behavioral preferences and concepts. Since these changes can be seen by the human participant while the behavior system is being updated, when the character shows an attitude change the person can easily feel sympathetic toward the character. This is the main response we got from SIGGRAPH 99 participants who had a chance to interact with the **(void*)** characters.

Given this, one of the notable facts is that although the three main characters in the installation showed very different personalities their creature kernels have the same basic structure. Different initial biases toward different desires and preferences for certain types of dance, differently progressing learning rates, etc., made the characters look and behave very differently, as they conveyed strong and easily recognizable archetypes.

From a programmer's point of view, this motivation-based creature design freed us from considering every possible instance or situation a character might encounter. We have been able to create emotionally compelling synthetic characters, and then it was easy to add complicated situations and possible actions and still get realistic and emotional responses from the characters that could elicit sympathetic interest from human participants. That is, we are able to give them internally active and externally expressed feelings that can guide their adaptation to environments not fully anticipated by their creators.

Chapter 12

(void*): A cast of characters – experimental studies

One of the main goals of the **(void*): A cast of characters** project is to build characters that have distinct personalities which are conveyed through the way they express themselves, adapt to the world and learn, and characters that can convey such attributes to the human participants through interactions. In this chapter, I describe how different factors have been considered and integrated into the system to build a character with a certain personality, and then I present results of an evaluation of the success of this approach as judged by users who interacted with the system and thus with the characters.

12.1 Construction process

The process of implementing the characters started with a series of brainstorming sessions for deciding on which personality each character should represent; these sessions included all members of the group who worked on this **(void*): A cast of characters** project. Instead of using terminologies such as extroversion or conscientiousness at the initial stage of conceptualizing personalities, more common words such as coward, cool or easily upset were used for ease of discussion. This process corresponds to annotating the affective nature of animations using plain English words such as happy or sad, which are affect descriptors that fit designations based on the six primary emotions, instead of (valence 1, stance 1) or (valence -1, stance -0.2), which are three-axis representations of affect of the type that is actually used in the creature kernel framework. This convention was taken because it helps people who do not need to know internal representation think easily and intuitively using plain language. The *change of basis* was not difficult, i.e., representations of affect or personality in one basis were readily convertible to the other. After agreement on the general descriptors of the personalities, the behavior designer could decompose the description using the representation that is more easily adapted to the creature kernel framework.

For the **(void*)** project, the following was decided: *Earl the Trucker represents a personality with much anger. He thinks it is irritating to be possessed and to be forced to dance. Elliot the Salesman is a nervous and timid character. He is frightened by possession and dances because he is forced to do so by an authority, but does not quite enjoy it. Fast Eddie(Dude) likes dancing and thus he enjoys being possessed.* This conceptual description can be transformed to a Big Five personality factor representation assigning numbers to each factor that correspond to relative emphasis in each character. The designer's conception of the (void*) character personalities is shown in Figure 12-1. As shown in the figure, the personalities of the (void*) characters can be represented in terms of the amounts of **(extroversion, agreeablness, conscientiousness, emotional stability)**. The representations for Earl, Elliot and Eddie are (medium low, low, medium, medium low), (medium low, medium, medium, medium low) and (very high, medium low, medium low, medium), respectively. To be specific, the **extroversion** factor is mainly reflected in a character's behavior when he is serving the role of a supporting character, i.e., when he is not possessed and one of the other two is the currently possessed character. As a character is more extroverted, he is more ready to express his emotional reactions such as surprise (when Earl dances though he was never expected to dance, for example), or encouragement or sympathy when the possessed character is doing well or being hurt. The **agreeableness** factor is represented as the degree of being unresistant to possession and

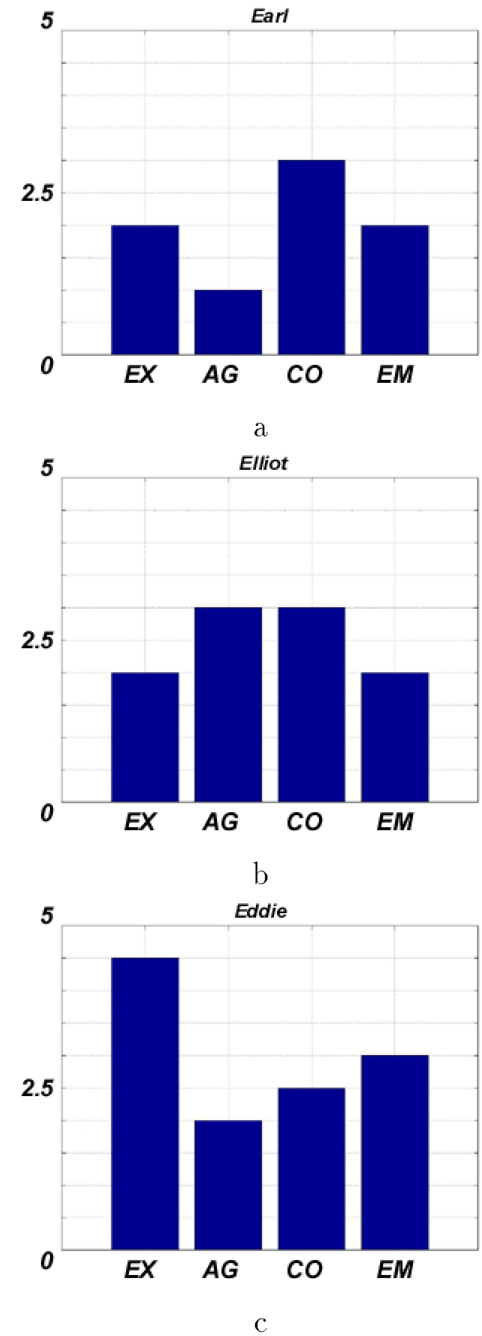


Figure 12-1: Rather general descriptions of character personalities have been transformed to Big Five personality factor representation. From the top, each graph represents the designated personality of Earl the Trucker, Elliot the Salesman and Fast Eddie (Dude). Initials on the *x-axes* of the graphs stand for extroversion, agreeableness, conscientiousness and emotional stability, respectively, and the values are rated between 0 and 5. Factor Five, intellect-imagination, is not included in this representation since it is irrelevant to this study, as discussed in Chapter 8.

other externally influenced or triggered behaviors, combined with confidence in one's own beliefs. Earl, whose agreableness is low, is resistive to possession, initially thinking that being influenced by an external power is not a good thing. Also, although his attitude toward possession does change through the interaction, it procedes very slowly. Eddie's agreeableness is also medium low and is shown by his not being very ready to change his belief that dancing is a cool thing to do though he is accepting of possession and being controlled. Elliot's agreeableness is medium. Although he learns the nature of dancing and possession and is ready to change his initial attitude toward them based on interaction, he does not welcome the possession very much mainly due to his nervousness toward authority. The **conscientiousness** factor is mainly expressed through the way the characters perform actions. This factor is mainly coded in the way each action is performed by each character, i.e., its expression through how the corresponding animation file has been designed and is being played. Animators made Earl's animations in such a way that he performs actions and dances in a kind of jerkey and abrupt manner, Eddie's to make him dance confidently, and Elliot's conscientiously. The **emotional stability** is expressed through the resistance to changes in belief through interactions and the level is medium or medium low for all three characters. Characters may change and update their beliefs in the nature of possession and dancing as the interaction proceeds.

12.2 Experiments and evaluations

The problem with this character behavior design method is that it is not obvious how to evaluate the success of the approach. As an evaluation attempt, however, a number of questions were asked in a questionnaire given to persons who were about to interact with the installation and, thus, the three characters, with further questions given after the interaction. The questions focused on figuring out whether the distinct personalities that the behavior designer tried to convey through behaviors, learning profiles, etc., were well perceived by the novice participants.

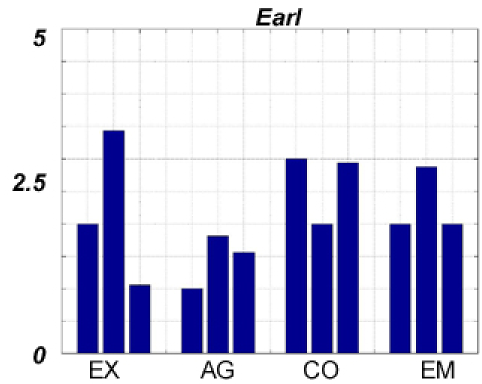
To exclude any bias that might arise from knowing descriptive names of the characters such as Trucker, Salesman,or Dude, only the neutral names such as Earl, Elliot and Eddie were given to the participants. Figures of each character were also included in the survey sheet with these names written on the figures. The survey sheet given to the participants is found in Appendix F.

Overall, all the novice users interacted with the characters without difficulty. All of the participants interacted with all three characters, and prolonged the interaction long enough to witness the emotional arc –the change is attitude toward the possession and dancing– in at least one character. The users' ability to perceive the various personalities was shown by the survey results presented below.

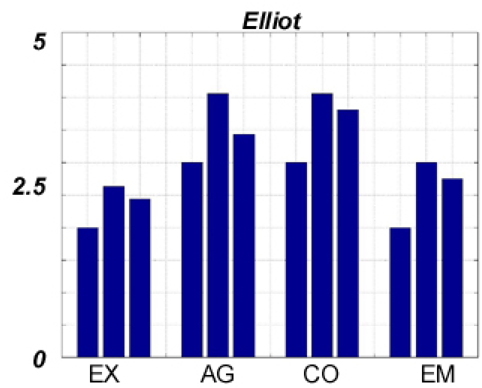
12.2.1 Survey results

Sixteen subjects participated in the interaction experiment. All were either MIT graduate or undergraduate students aged between 19 and 39, with various backgrounds, such as neuroscience, applied math and computer science. The questions given to each subject were mainly about perceived personalities of the three characters. The same set of questions were asked both before and after the real interaction. Subjects were initially asked to answer questions about their perception of the personalities of the characters as judged by just looking at the static figures of the characters. This was to assess how much the visual look of the character affects the apparent personality compared to the details of their actions, learning curves and emotional expressions.

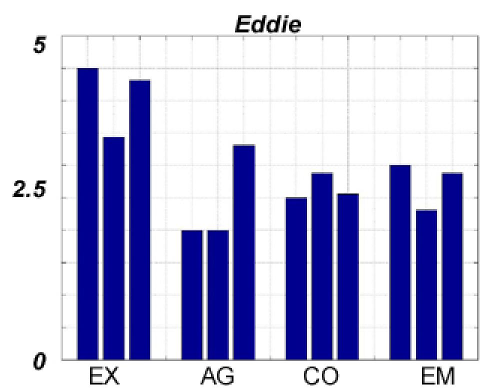
Figure 12-2 illustrates the survey results. Three graphs represent survey results for Earl, Elliot and Eddie, respectively. As in Figure 12-1, initials along the *x-axis* under each column represent four of the Big Five personality factors, with the value, rated on a zero to five scale. There are three bars for each personality factor, and each, respectively, represents the value for that personality factor as conceived by the behavior designer, the mean of the perceived value for the personality factor by the subjects before, and finally after, the interaction.



a



b



c

Figure 12-2: Survey results on perceived personalities of **(void*)** characters' personalities. On the *x-axis*, EX stands for extroversion, AG for agreeableness, CO for conscientiousness and EM for emotional stability. For each case, the first bar represent the designer's perception for each personality component in each character, and the second and third bars represent perceived mean values calculated from subjects' answers, before and after the interaction, respectively.

	extroversion	agreeableness	conscientiousness	emotional stability
Earl	15	14	11	13
Elliot	14	15	12	14
Eddie	14	14	11	13

Table 12.1: Total of sixteen subjects participated in this experiment. In this table, for each personality factor and the character, the numbers of answers that agreed with the designer's conception are shown.

	extroversion	agreeableness	conscientiousness	emotional stability
Earl	0.0002	0.0021	0.1050	0.0106
Elliot	0.0384	0.0002	0.0384	0.0021
Eddie	0.0001	0.0021	0.1050	0.0106

Table 12.2: The probabilities, calculated using binomial test, for getting the results shown in Table 12.1 are shown in 0 to 1 scale.

The major question asked here is whether the personality types that the behavior designer intended to convey through the three different characters were well perceived by the subjects. For each character, the designer's conception for each personality factor is converted from the verbal description to a range of numbers. For example, Earl's agreeableness and conscientiousness are medium and low, respectively. These are converted to the range between 2 and 4 and the range betwen 0 and 2. There are six possible responses for each personality factor of each character. And the number of responses that fall into the designer-intended range was counted (Table 12.1). Then, binomial tests [109] were done over the sampled data to calculate the probability of getting the resulting number (or fewer) of answers within the range. Table 12.2 show the test results. From these results, it can be said that subjects perceived all of the personality factors in all three characters as the designer intended with more than a 95% level of confidence except in three cases, two of which are about the conscientiousness factor in Earl and in Eddie's personality. In the common usage, conscientiousness can refer to how conscientious people are at behaving and performing certain gestures or motor patterns, or it could refer to a high level cognitive component of being conscientious. Since what it meant in the survey was the former as the designer tried to convey the factor through the animations, the subjects were told that conscientiousness should be judged in terms of how motor patterns are performed and behaviors are done, but it seemd difficult for some subjects to distinguish this meaning of the concept and the answers show a high variance as a result.

The next question asked was how much the interactions and the observations of characters' behavior helped subjects gain the impressions of personality beyond the visual appearances. One would expect that the impressions were due to a combination of both factors. In this experiment, the perceived personalities were assessed both before and after the interaction, and the comparison of the results is shown in Table 12.3. The number of answers that were already in the designer's conception range and stayed in the range after the interaction, were counted (shown in the **before** row) and then the answers that got closer to the designer's conception after the interaction, which originally did not belong the designer's conception range, were counted (shown in the **after** row) and the sum is shown in the next row named **total**. It is seen that in the end, the majority of the subjects perceived the character personality as the designer intended. But out of 12 cases, in seven cases, less than 50% of the subjects perceived the character personalities as they were intended before the interaction. It can be concluded that the interactions and observation of behaviors helped the subjects to have a better appreciation of the nature of the characters. Table 12.4 show the effect of interaction on subjects' perception of characters personalities through another comparison. In the table, the rows labeled **closer** show the number of subjects whose perception of personality was not in the designer's conception range before the interaction but got closer to the conception after the interaction. The rows labeled **farther** show the number of subjects who initially did not perceive the personality of characters as the designer intended and the perception got even farther from it after the interaction. Among the twelve comparisons, in nine cases, more subjects perceived the characters' personality closer to what the designer intended after hav-

		extroversion	agreeableness	conscientiousness	emotional stability
Earl	before	2	9	7	5
	after	14	5	5	10
	total	16	14	12	15
Elliot	before	8	7	10	12
	after	4	8	3	4
	total	12	15	13	16
Eddie	before	8	9	10	7
	after	6	7	3	7
	total	14	16	13	14

Table 12.3: For each character and personality type, this table shows the number of answers that accorded with the designer's conception before the interaction, which stayed to be so after the interaction. Second is the number of answers that did not accord with the designer's conception initially but changed to be closer to that after the interaction, and third is the sum of those two cases.

		extroversion	agreeableness	conscientiousness	emotional stability
Earl	closer	14	5	5	10
	farther	0	2	4	1
Elliot	closer	4	8	3	4
	farther	4	1	3	0
Eddie	closer	6	7	3	7
	farther	2	0	3	2

Table 12.4: This table compares the number of subjects who initially perceived the value of the each personality factor differently from that conceived by the designer and then later perceived a personality closer to the designer's conception, after the interaction, and those who came to perceive it farther from the designer's comception.

ing the interaction, rather than farther. There are three cases that tied. From this result, it can be said that the interaction biased the subjects' beliefs in the way that the designer intended with 99.9% level of confidence. Two of the tie cases are about the **conscientiousness** factor. As described above, it seems that it is because in the common usage, conscientiouness includes a high level cognitive factor that is not emphasized in this implementation, so some changes in beliefs corresponded to chance level.

12.3 Conclusion

The greatest obstacle to this study was difficulty in use of consistent terminology. Even if we all understand and even agree on what each word referring to each personality factor means, the descriptions of individual personalities still tend to be context dependent and quite broad. Nonetheless, the personality of the characters perceived by novice subjects showed characteristics that are similar to those that the behavior designer had intended to be conveyed.

Part IV

DISCUSSION AND CLOSING REMARKS

Starting from a consideration of the basic organization of the central nervous system (CNS), this thesis has demonstrated a method of implementing a software kernel for building interactive, intentional virtual creatures with lifelike characteristics such as adaptation abilities, personality and expression of emotions.

The main contribution of this work is the provision of a framework upon which lifelike artifacts can be built. This framework includes specific basic components encompassing different hierarchical levels. The creature kernel, which is the basis of a synthetic character, is composed of four main subsystems – perception, motivation, behavior and motor systems – where the results of the coordinated interactions of these four systems are propagated back to the creature kernel as well as outwardly, to be expressed through specific movements and patterns of action and facial expressions. The feedback signal propagated back to the creature kernel updates the kernel to cope with the dynamics of the world and to make the creature more adaptive. This update includes both quantitative and qualitative changes which are realized through structural learning, concept learning and affective tag formation. These forms of plasticity are reflected by changes in beliefs, in desires relative to various drives, in behavioral or object preferences and in acquired skills. Each synthetic character is assumed to have various built-in properties, like the innate characteristics of real animals, different for different species. These include a number of sensors, innate drives, behavior patterns and motor skill abilities that support the expression of those behavior patterns through the body (**3D** graphics in the cases discussed in this thesis). These behavior patterns can be seen as the equivalent of the behaviors often referred to as *fixed action patterns* by ethologists.

However, the current implementation of learning has limitations in various aspects. One of the inherent problems with the reinforcement learning approach is the curse of dimensionality. Adopting the belief network and the independent relations explicitly coded in the network helps in dealing with this problem, but nevertheless the simultaneous activition of many events could be a problem. In real creatures the attention mechanism solves this problem [99, 40]. We have not implemented an attention mechanism in the proper way yet. Incorporation of an attention mechanism that does not suffer from perceptual aliasing [77] would be the next step.

Feature extraction [32, 106, 119] is another major problem. Often in machine learning, unsupervised learning proceeds to provide the raw material for the next stage, in most cases, which in most cases is supervised learning [4]. In the current implementation, it is assumed that creatures already know what are the characteristic features of all the objects that they are dealing with. For example, they can query a flower and get data such as its color and smell. Generalization depends on handcoded information on hierarchical relations among those features. The fact that **animal** is a bigger category than **lion**, so using the information **animal** is a preferred way of learning things initially, is given as *a priori* knowledge to the creatures. But this is one of the problems that biological creatures depend on their own experiences and knowledge to solve. They learn to attend to specific features, and figure out the inclusion relations among features that describe objects over the course of development and learning. We took the short cut solution for now but it will be an important next area to explore for the implementation of believable characters.

The associative learning comes only from positive examples. A character does learn from negative affective experiences but still the examples are used only as **positive** examples of those that elicit negative affective reponses. The ability to learn from negative examples may speed up the learning, but the current system does not include this ability.

Since the whole system is built within a probability framework, there are two fundamental flaws that come from this approach. One is the fact that the order of the appearance of learning examples does not affect a creature's learning, which is not found to be true in cognitive science literature [38, 42]. This can only be true when the examples are all independent, which is not always the case in the real world. The other is the question of where all the *a prioris* are coming from. Preferences for certain colors, smells and tastes have to be handcoded before running the system. This brings up the nature-versus-nurture debate again, and where to draw the boundary line is another question to think about.

The role of the motivation system, which includes the affect system, has been emphasized throughout this thesis. The desired characteristics of the synthetic characters that were major aims of the thesis work were adaptability and compelling lifelike impressions. Real-time interactions with the external world, which includes both the virtual environment where the synthetic characters are placed and the real world where some of the sensing inputs arise, demands adaptability. This is because installing all the necessary components for dealing completely with all situations that might occur is intractable from a resource standpoint, and it is simply too difficult for the designer to predict and implement them all. Achieving a compelling lifelike impression requires believability in the characters, which should be supported by the intentionality they display, which must be conveyed in a clear and understandable way. The motivation system plays the crucial role in both cases, and thus has been emphasized in the implementation of the creature kernel.

In particular, once the motivation system in the model was expanded to incorporate an affect system with the major role in learning, then various phenomena of adaptation often thought of as distinct were all found to be natural outcomes.

The thesis started with an overview of the organization of the CNS and the functional needs that drove the evolution of the hierarchical system. In the natural world, as the need for showing more complex behaviors and coping with a more dynamic environment arose, demand for higher intelligence also arose. Intelligence emerges either from additions of higher levels in the hierarchy, i.e., by adding structures that can deal with qualitatively different problems, or from expansion of the number of the units at the same level in the hierarchy. The latter would correspond to a creature with a greater number of different motor skills that includes finer detail so that, in effect, it would be able to manifest very appropriate motor control for various situations through bottom up choices. This could be seen to mimic the addition of a high-level control mechanism that can carefully control little details in a top down manner.

The simulation examples shown in this thesis demonstrate characters that function adequately in the given environments. But they have only a few basic drives, and the number of motor skills is small and the granularity of control is also rough. Expansion of the creature kernel in both horizontal and vertical directions will make the implementation of more intelligent characters possible, and interesting expansions of adaptation ability will emerge. This system contains only four basic component systems, and the functions of other parts of the CNS that are also crucial for the survival of real creatures have been given only minimum consideration. Sensory systems and special memory systems are the two most prominent examples among them. Learned knowledge is all stored in the form of connection weights and there is no explicit memory. The structure of memory and the way the contents are accessed affect learning, and adaptive changes in memory and specific learning strategies need to be allowed in future development of the software. Sensors are all treated with equal importance, and it is assumed that characters do not need to deal with the notorious binding problem, which is not always true in real animals. Efficient strategy for proper binding is also a subject of learning and thus should accelerate the efficient development of attention mechanisms.

Though multiple characters of similar complexity were developed and placed in the (void*) situation, not too many interactions were allowed among the characters themselves. Different personalities, experiences and *a priori* parameter settings and strategies cause differential development of each character, and thus change the nature of interactions. This is what novels are about in the world of literature [1]. Characters

[1] I watched a Star Trek, *Voyager* episode recently. In this episode, there was an interaction called

Figure 12-3: Shakespeare's play, *Hamlet* is decomposed into nine story elements as shown in this figure. In the preliminary simulation, characters' interactions with the human participant changes relative strengths of drives and desires of the characters, and this causes them go through the story elements in an order different from the original plot. A story emerges as the interaction proceeds, in a way that depends on an individual character's personality and past experiences.

experience various things in the world through their interaction with the world and other characters, and as a result, build motivations for initiating certain actions or responding in a certain way when the next event occurs. With properly updatable motivation systems and adaptability of characters being implemented, an interactive emergence of a story can be realized. A preliminary implementation was done using the story elements of Hamlet (See Figure 12-3) and the results proved that a seemingly coherent plot can emerge differently every time, through interactions between human participants and characters.

Also, building artificial creatures provides us with a new kind of opportunity to understand biological organisms better, in addition to enabling us to build synthetic characters that are fun and beneficial to interact with. Artifacts are built using our best knowledge of what the biological creatures would be like, put in a similarly complex and dynamic environment. Depending on whether the simulated organism succeeds or not, the results tell us whether our assumptions may be right and sufficient, or whether there must be some missing factors that we have not thought of. Thus, the synthetic creatures are a kind of embodied theory of behavior and its underlying control system.

Holonovel where the human can interact with holographic humanoid figures, which are not distinguishable in their realistic impression from the participant himself. The person keeps interacting with those holographic figures and changing the world. The holographic figures respond and behave following their own minds, and a story emerges as the interaction proceeds. At one point, one of the participants was put in a dangerous situation and people who were watching him outside of the *Holonovel* changed the personality parameters of the holographic figures instead of resetting the scene or event. It was the only proper way of changing the way the story proceeds, and thereby the person was rescued.

Part V

APPENDIX

Appendix A

Derivation of the learning rule for organizational learning

Let's consider a preference update process of a behavior group with a parent and N children behaviors. v_i, the valence asociaed with child behavior i can be any value between -1 and +1, where -1 corresponds to the *worst* experience, and +1 corresponds to the *best* experience. For simplicity, we approximate it to a discrete case where v_i can be any value that can be represented as $-1 + \frac{n}{(\frac{k-1}{2})}$ where n is an integer between 0 and $k-1$ and then, k, here corresponds to the level of discretization, which can be adjusted arbitrarily.[1]

Let us assume that a parent behavior has N children behaviors. When one of those N behaviors is activated at the next tick, the possible number of events that can happen is $N \times k$. Table A.1 illustrates one possible simple case, where there are three children behaviors and valence value is discretized to take one of five possible values between -1 and 1. For example, the entry at second row and second column represents the event when the child behavior 2 is activated and the valence as the result of that activation is -0.5. Each entry at $i-th$ row and $j-th$ column, e_{ij} has its own probability of occurrence $p(e_{ij}|b_i)$, with $\Sigma_{ij} p(e_{ij}|b_i) = 1$. The preference to

[1] when k = 1, it is assumed that v_0 equals a certain constant and does not change, which means that b_i has fixed valence value and it is not updated over time.

behavior 1	(-1,0,0)	(-0.5,0,0)	(0,0,0)	(0.5,0,0)	(1,0,0)
behavior 2	(0,-1,0)	(0,-0.5,0)	(0,0,0)	(0,0.5,0)	(0,1,0)
behavior 3	(0,0,-1)	(0,0,-0.5)	(0,0,0)	(0,0,0.5)	(0,0,1)

Table A.1: Table of all possible events when $N = 3$ and $k = 5$.

behavior i would be then, the summation of the nonzero entry of each row multiplied by this probability i.e.

$$E(R_i) = \Sigma_j p(e_{ij}|b_i) v_{ij}, \tag{A.1}$$

where v_{ij} is the valence value associated with the entity e_{ij}.

The initial value of the preference value reflects the *best guess* of the parent behavior based on its *a priori* knowledge, ξ, on probability of the occurrence of each event e_{ij}, and thus the valence v_{ij} that it would experience as the result of the activation of that behavior i. Formally, it can be written as follows,

$$IP_i = \Sigma_j p(e_{ij}|\xi) \cdot v_{ij} \tag{A.2}$$

where IP_i represents the initial preference given to behavior i.

Here we can define Θ to be a variable that corresponds to the underlying probability of each "event" and use $P(\Theta|\xi)$ to represent uncertainty about Θ given our background konwledge ξ. Another variable D represents the experience of the creature, which encompasses the behavior chosen and the experienced valence as the result of it. Using the Bayes' rule, now we can obtain the probability distribution for Θ given experience D, and background knowledge ξ:

$$p(\Theta|D, \xi) = \frac{p(\Theta|\xi) p(D|\Theta, \xi)}{p(D|\xi)} \tag{A.3}$$

where

$$p(D|\xi) = \int p(D|\Theta, \xi) p(\Theta|\xi) d\Theta. \tag{A.4}$$

Given Eqn A.3, the expected valence for each behavior i is equal to the average over possible values of θ_{ij} :

$$p(v_i|D,\xi) = \int v_{ij} \cdot p(e_{ij}|\theta_{ij},\xi)p(\theta_{ij}|D,\xi)d\theta_{ij}. \tag{A.5}$$

Here the observable variable v_{ij} is discrete and has one of k possible values between -1 and +1, and the likelihood function is given by

$$p(e_{ij}|\theta_{ij},\xi) = \theta_{ij}, j = 1, ..., k. \tag{A.6}$$

where $\Theta = \theta_{11}, \theta_{12}, ..., \theta_{Nk}$ are parameters with a multinomial distribution. The simple conjugate prior used with multinomial distribution is the Dirichlet distribution [46]:

$$p(\theta_i|\xi) = Dir(\theta_i|\alpha_{i,1}, ..., \alpha_{i,k}) \equiv \frac{,(\alpha_i)}{\Pi_{j=1}^{k} ,(\alpha_{ij})}\Pi_{j=1}^{k}\theta_{ij}^{\alpha_{ij}-1}, \tag{A.7}$$

where $\alpha_i = \Sigma_j \alpha_{ij}$ and $\alpha_{ij} \geq 0$. The posterior distribution $p(\theta_i|D,\xi) = Dir(\theta_i|N_{i1} + \alpha_{i1}, ..., N_{ik} + \alpha_{ik})$. Given this conjugate prior and experience D, the normalized probability for e_{ij} is given by

$$p(e_{ij}|D,\xi) = \int \theta_{ij} Dir(\theta_i|N_{i1} + \alpha_{i1}, ..., N_{ik} + \alpha_{ik})d\theta_i = \frac{N_{ij} + \alpha_{ij}}{N_i + \alpha_i}, \tag{A.8}$$

where $N_i = \Sigma_j N_{ij}$ and $\frac{N_{ij}}{N_i}$ equals the implicit *a priori* probability of e_{ij}. Given the experience set D, *a posteriori* probability for e_{ij} is $D(e_{ij}|D,\xi)$ and expected valence EV_i equals $\Sigma_i P(e_{ij}|D,\xi) \cdot v_{ij}$. After each experience of v_{lj}, the expected valence for each behavior is updated as

$$EV_i(t+1) = EV_i(t) \cdot \frac{N+\alpha}{N+\alpha+1} + \frac{v_{lj}}{N+\alpha+1} \tag{A.9}$$

if $i = l$, and

$$EV_i(t+1) = EV_i(t) \cdot \frac{N+\alpha}{N+\alpha+1} \tag{A.10}$$

if $i \neq l$. Similar learning algorithm can be applied to stance to give the normalized expected stance at time t, $\frac{ES_i(t)}{\Sigma_i ES_i(t)}$. Preference for behavior i is, then, calculated as the weighted sum of normalized expected valence and normalized expected stance: $PR_i = w_{vi} \cdot \frac{EV_i(t)}{\Sigma_i EV_i(t)} + w_{si} \cdot \frac{ES_i(t)}{\Sigma_i ES_i(t)}$.

Appendix B

Learning in the creature kernel

This appendix section brings together and summarizes the learning mechanisms found in various parts of the dissertation.

Action selection is the mechanism that resolves conflicts in the resource. In the case of animals or synthetic characters, the most obvious resource in conflict is the motor system. This is known as the behavioral final common path [80] in the animal behavior literature. The behavior system of the creature kernel is composed of a cascaded structure of groups of mutually exclusive behaviors, and action selection is the mechanism for deciding relative priorities within the behavior system to solve this conflict [1].

The creature kernel is a value-based system, and the action selection mechanism utilizes the provided values. The value update equation adopted in this thesis work is mainly based on the one suggested by Blumberg [7] as introduced in Chapter 6, which is shown below (Eqn. B.1).

$$V_{it} = PR_i \cdot [LI_{it} \cdot Combine\,\{(\Sigma_k f(RM_{kt})), (\Sigma_n f(IV_{nt}))\}], \tag{B.1}$$

[1]In fact, the motor system is not the only limiting factor. The perception and the motivation system have their own limitations in terms of the number of processes that can be dealt with at the same time. Attention mechanisms resolve this problem in real animals.

where V_{it} is the value of behavior i at time t, LI_{it} is the level of interest in behavior i at time t, RM_{kt} is the value of the releasing mechanism k at time t, which includes outputs from the perception system such as P_i or other behaviors, B_j. IV_{nt} is the value of the internal variable n at time t, which mainly comes from the motivation system.

Given this, the default way of converging on one behavior in each behavior group is the mutual inhibition process, which was also formulated by Blumberg [7] and introduced in Chapter 6 (See Eqn. B.2).

$$V_{i,t+(n\times\Delta)} = V_{i,t+((n-1)\times\Delta)} - (\Sigma_{m\neq i} N_{m,t+(n\times\Delta)} \cdot V_{m,t+(n\times\Delta)}) \tag{B.2}$$

This action selection mechanism is embedded in the creature kernel, which is composed of four subsystems – perception, motivation, behavior and motor systems, each of which consists of basis units. When it comes to the operation, at every tick t, two types of operations take place in the creature kernel (See Figure 4-5.), they are the forward operation and the backward operation.

Given the above equations, the forward operation is straightforward. Each subsystem makes a diagnosis of the current state of the world and the creature itself, if necessary, and calculates appropriate values. The modules in the perception system sense the state of the world and calculate Boolean values for either presence or absence of certain objects, visual stimuli, sounds, etc., as well as incidental information, such as the distance and affective state of another creature. The resulting values returned by the perception system can be represented as $f_q(p_r)$ where f_q is a function that takes an object returned by a sensor (p_r) and returns a value [2], where $q \geq r$. The motiva-

[2]Current system has not been expanded to include a proper implementation of the attention mechanism. At this sensing and sampling stage, the values from the motivation system calculated at the previous tick can bias the creature to focus primarily on a certain portion or aspect of the world. Where the attention mechanism should be placed in the creature kernel depends on assumptions and implementation requirements. The easiest way of incorporating it into the current system would be to have sensors carry a multiplicative flag determined by two factors. One of the factors would be a threshold level, above which level the sensor should be activated no matter what

tion system does a similar thing during the forward operation phase. It samples and calculates the levels of all the drives and affective components based on the behavior or sensing of the outcome from the previous tick. At its calculation, the motivation system also incorporates influences from the perception, behavior and motor systems.

The outcome from both of these systems in combination with the values from the behavior system is fed into the behavior system for its value calculation. The value of each behavior unit is calculated using Equation B.1 and an action selection mechanism determines which of the behaviors win(s) at the behavioral final common path to make it send the actual action command to the motor system.

Then the process turns into the backward process through which the result of the previous action is reflected in the creature kernel. This process can be viewed as learning. The learning may happen in all four systems. In fact, learning in this system is characterized as a distributed process, but here I explain it primarily using the behavior system as the example.

Plasticity is imposed on all the variables that appear in Eqn B.1.

- PR_i. This is the preference from the parent behavior. It depends on the outcome of the previously performed behaviors. The preference is updated in proportion to the experienced valence or stance while performing the actions that involve the behavior. Like in habit formation [57], satisfaction from the behavior just performed enhances the preference given to that behavior and suppresses that given to the others in the same behavior group, i.e., the alternatives. In effect, this makes the creature prefer the effective behavior, or form a habit even if the performed behavior does not have any particular advantage over the others except for the fact that it happened to be peformed prior to others. The preference update rule is written for the behavior just performed

the current focus of attention is. The other would be a Boolean flag signaling whether the sensor is within the focus of attention range, in which case, the sensor should return a value.

as follows.

$$PR_i(t+1) = PR_i(t) \cdot \frac{N+\alpha}{N+\alpha+1} + \frac{I \cdot v_i}{N+\alpha+1} \tag{B.3}$$

and the update rule for the other behaviors in the same behavior group is as follows.

$$PR_i(t+1) = PR_i(t) \cdot \frac{N+\alpha}{N+\alpha+1}. \tag{B.4}$$

- **$f(RM_k)$.** Each releasing mechanism, RM_k is a pointer to either one of the behavior units or to perception units. After a certain behavior activation and action, recently activated perception units and behavior units are fed into the concept learning process to compute the relevance to the behavior of curent interest for learning, B_i. As discussed in Chapter 6 and illustrated in Chapter 10, the confidence in belief that those units in consideration are relevant to a certain outcome of the current behavior is updated using

$$P(h|X) = \frac{P(X|h)P(h)}{P(X)} \tag{B.5}$$

  where X is the outcome of the event, h is the hypothesis under consideration, i.e, each of RM_k currently involved in the update in belief process. Once the relevance is figured out, the nature of the effect is reflected in the function, completing the update in $f(RM_k)$.

- **$f(IV_n)$.** Internal variables, IV_n are pointers to units in the motivation system. The learning rule for this case is the same as that for the releasing mechanisms. Changes in certain drive or affect levels are scrutinized and concept learning is applied to update the belief in the relevance of the behavior to those motivation system components.

Affective tag formation. The pronomes that were encountered through the activation of the behavior gets tags associated with affective components, which later are perceived as affective tags. Each object of interest is implemented as an instance of a class with affective component variables. During the backward process, the encountered pronomes have their affective tags updated, which later can be retrieved when the creature makes a decision for another forward process. See Chapter 6 for details.

All these types of learning happen also in the other parts of the creature kernel, not just in the behavior system, and result in various phenomena such as S-S associations (the concept learning within the perception system), development of a secondary drive/motivation (learning within the motivation system), etc. After this whole update, the creature is ready for the next tick, and another forward operation begins.

B.1 More on learning algorithm

This section explains the motivation behind the computational framework introduced so far.

Uncertainty is one of the inherent problems faced by synthetic characters as well as biological creatures. They need to survive in an unpredictable world with limited information processing and sensing capabilities. Both time and processing power are not unlimited resources that can be exploited to any desired degree of precision. Nonetheless, at every tick of the simulation update, creatures are required to make a decision that – as far as their bounded processing abilities allow – maximizes the expected reward. Among performable behaviors B_j, the one that would maximize the expected reward (ER) is chosen to be performed [3]. See Eqn. B.6.

$$i = argmax ER(B_j) \tag{B.6}$$

[3]The concept of reward here is a broad one. It does not only refer to the immediate reward that would prohibit exploration of a new regime in the behavior repertoire space. The reward here includes especially the expected gaining of knowledge that will allow the creature to go beyond the exploitation of current knowledge.

A probabilistic network is a way of representing uncertain *knowledge* in a systematic way [90]. Its independence relationship (and conditional independence relationships) explicitly marked through absence of links makes coherent reasoning or decision making possible. In the creature kernel implementation, each subsystem is modeled as a probabilistic network.

A clear alternative would be a logic-based approach. However, this representation without a network makes it difficult to deal with the existence of *invisible facts*, and thus leads to unintentional neglect of relevant information in the belief calculation. It has been the cause of counterintuitive results that early implementations of expert systems suffered from [90, 108].

Reinforcement learning is a learning framework for a creature trying to survive in an unknown environment [118]. A synthetic character is modeled as a situated learner figuring out its own strategy for survival. Reinforcement learning has been widely used for situated agents [50, 1, 104, 65, 66]. This is because, unlike stochastic learning algorithms that require a large number of examples, or extensive processing that cannot be done on line [44, 63, 64, 100], or directive teaching signals in the case of supervised learning, the assumptions made in the reinforcement learning fit the situation that agents (real or synthetic) are faced with.

Given a policy, a reward function, a value function and a model of the environment, reinforcement learning algorithm can be applied. In our implementation of the behavior system, at each backward operation process, the preference value of the performed behavior is updated as,

$$EV_i(t+1) = EV_i(t) \cdot \frac{N+\alpha}{N+\alpha+1} + \frac{r}{N+\alpha+1}. \tag{B.7}$$

Here r is the sum of rewards from the motivation system through stance (s) and valence (v) values multiplied by the arousal level (I). (See Appendix A for more details of the equations.) Assigning different mixing weights w_v and w_s, Eqn. B.7 can be rewritten as,

$$EV_i(t+1) = EV_i(t) \cdot \frac{N+\alpha}{N+\alpha+1} + \frac{I \cdot w_v \cdot v}{N+\alpha+1} + \frac{I \cdot w_s \cdot s}{N+\alpha+1}. \tag{B.8}$$

The *valence* and *stance* values here include both the direct feedback from the just performed behavior and the disappointment factors that are proportional to the difference between the expected *valence* and *stance* and the actually experienced ones. Further refining the equation, Eqn. B.8 can be rewritten as,

$$\begin{aligned} EV_i(t+1) = EV_i(t) \cdot \frac{N+\alpha}{N+\alpha+1} \\ +\frac{I \cdot (w_{v1} \cdot v + w_{v2} \cdot (E(v) - v))}{N+\alpha+1} \\ +\frac{I \cdot (w_{s1} \cdot s + w_{s2} \cdot (E(s) - s))}{N+\alpha+1}. \end{aligned} \tag{B.9}$$

Assuming both w_{v1} and w_{s1} are equal to w_1, and w_{v2} and w_{s2} are equal to w_2, Eqn. B.9 can be, again, rewritten as,

$$EV_i(t+1) = EV_i(t) + \frac{1}{N+\alpha+1}((I \cdot (w_1 - w_2))(v+s) + (I \cdot w_2)(E(v)+E(s)) - EV_i(t)). \tag{B.10}$$

Assuming $E(v)+E(s)$ is close enough to $EV_i(t+1)$, Eqn B.10 is similar to the $TD(0)$ algorithm, which is the simplest of the well-known reinforcement algorithms [118].

In general, in computational neuroscience, associative learning is modeled as Hebbian update of correlative links [45, 48]. This idea is expanded in the thesis work to include updates in directional links between any units in the kernel, i.e., any of the basis units in the behavior, motivation and perception systems. In the creature kernel, simultaneous activation of two basis units results in an update in the links that connect them. In particular, appearance of a certain object associated with an affective experience attaches an affective tag to the object.

The concept of Hebbian update is expanded to be incorporated in this Probabilistic architecture in the form of concept learning. Beliefs in causality and relatedness relations are updated in the probabilistic manner as proposed by Tenenbaum [121]. The study of concept learning has a long tradition [19, 107, 52, 84]. Which model explains the concept formation phenomenon the best is subject to debate and is an area still being studied. Modeling of behavior in terms of Bayesian probability theory is well received nowadays due both to its plausibility in explaining empirical results [139, 86, 13] and to its computational applicability. We have stretched the notion of *concept* to include ideas of what is often called *context* in the animal behavior and learning literature. Confidence is correlative and causal relations are learned and, in turn, are used when making decisions for performing actions. This confidence is not explicitly coded as any single item in the system, but is stored in the form of connection weights in a distributed manner.

Appendix C

preference.m

```
EV = [0.2  0.3 0.5];
% actual reward
vAlue = [-4.15 0.8 -0.1];
% variance
vAriation = [1 1 1];
N = 10;
alpha = 0;
rEward = 0;
c = 0;
figure; hold on;
maxIter = 50;
oNe = zeros(maxIter,1); tWo = zeros(maxIter,1); tHree = zeros(maxIter,1);
for world=1:maxIter
   alpha = world - 1;
   r = rand;
   if(r < EV(1)) c = 1;
   else if((r>=EV(1))& (r<(EV(1)+EV(2)))) c = 2;
      else c = 3;
      end
   end
```

```
if(c == 1)
   %% reaction from the world
   rEward = (vAriation(1)*randn)+vAlue(1);
   %% creature's reaction to the consequence
   sUrprise = abs(EV(1) - rEward);
   %% plot when surprise occurs
   if(sUrprise > 0.5)
      b = world*ones(1,1);
      hist(b);
   end
   %% update expected values
   EV(1) = (EV(1) * ((N+alpha)/(N+alpha+1))) + ((sUrprise * rEward)/(N+alpha+1));
   EV(2) = (EV(2) * ((N+alpha)/(N+alpha+1)));
   EV(3) = (EV(3) * ((N+alpha)/(N+alpha+1)));

else if(c == 2)
   %% reaction from the world
   rEward = (vAriation(2)*randn)+vAlue(2);
   %% creature's reaction to the consequence
   sUrprise = abs(EV(2) - rEward);
   %% plot when surprise occurs
   if(sUrprise > 0.5)
      b = world*ones(1,1);
      hist(b);
   end
   %% update expected values
EV(1) = (EV(1) * ((N+alpha)/(N+alpha+1)));
   EV(2) = (EV(2) * ((N+alpha)/(N+alpha+1)))+ ((sUrprise * rEward)/(N+alpha+1));
   EV(3) = (EV(3) * ((N+alpha)/(N+alpha+1)));
```

```
    else if(c==3)
     %% reaction from the world
     rEward = (vAriation(3)*randn)+vAlue(3);
     %% creature's reaction to the consequence
     sUrprise = abs(EV(3) - rEward);
     %% plot when surprise occurs
     if(sUrprise > 0.5)
        b = world*ones(1,1);
        hist(b);
     end
     %% update expected values
     EV(1) = (EV(1) * ((N+alpha)/(N+alpha+1)));
     EV(2) = (EV(2) * ((N+alpha)/(N+alpha+1)));
     EV(3) = (EV(3) * ((N+alpha)/(N+alpha+1))) + ((sUrprise * rEward)/(N+alpha+1));
    end
  end
end
% clamp : since it is the probability of choice
EV(1) = max(0.0,min(EV(1),1.0));
EV(2) = max(0.0,min(EV(2),1.0));
EV(3) = max(0.0,min(EV(3),1.0));
% normalize
sUm = EV(1) + EV(2) + EV(3);
EV(1) = EV(1) / sUm; EV(2) = EV(2) / sUm; EV(3) = EV(3) / sUm;
oNe(world,1) = EV(1); tWo(world,1) = EV(2); tHree(world,1) = EV(3);
   %% plot(world,EV(1),'r'); plot(world, EV(2),'b'); plot(world,EV(3),'g');
end
plot(oNe,'r:'); plot(tWo,'b.'); plot(tHree,'g-.');
axis([0 maxIter 0 1]);
xlabel('# events'); ylabel('preference from parent');
```

```
figure; hold on;
plot(oNe,'r:'); plot(tWo,'b.'); plot(tHree,'g-.');
xlabel('# events'); ylabel('preference from parent');
```

Appendix D

concept.m and other related routines

```
% concept.m
a = 0.1*randn(100,1) + 1.5;
b = 0.1*randn(100,1) + 0.5;
rr = rAdius(a);
narrowingFactor = 8;

figure; plot(a,b,'.r'); hold on;
axis([0 4 0 1.02]);

c = 0:0.2:4;
maxIter = 5;
k = zeros(max(size(c)),maxIter);

for i=1:maxIter
   for j=1:max(size(c))
      k(j,i) = 1 / (1+(narrowingFactor*(dIstance(c(j),a)) / rr)^(i-1));
   end
end
```

```
plot(c,k(:,2),'k:');
plot(c,k(:,3),'k-.');
plot(c,k(:,5),'k');

xlabel('data points'); ylabel('confidence in belief');

%rAdius.m
%returns the max distance between two data within a one dimensional array, a

function r = rAdius(a)

s = size(a);
currentMaxDistance = 0;

for i=1:s
   for j=1:s
      kk = abs(a(i) - a(j));
      if(kk > currentMaxDistance)
         currentMaxDistance= kk;
      end
   end
end

r = currentMaxDistance;

%dIstance.m
%get the min distance between a data point and an array of data
function k = dIstance(c, a)
```

```
ss = size(a);
currentMinDistance = 99999;

for i=1:ss
   t = abs(a(i) - c);
   if(currentMinDistance > t)
      currentMinDistance = t;
   end
end

k = currentMinDistance;
```

Appendix E

An example behavior weight file (Dude)

Following list is an example of behavior_weight_file initialization. This list is taken from Dude's behavior_weight_file and it shows the relative hierarchical structure among behavior units and initial organization of behavior groups.

```
CHILD PRIMARY_BEHAVIORS SIT 1.0

CHILD SIT WALK_TO_DOOR_INSIDE_FIRST_TO_SIT 1.0

CHILD SIT GO_TO_SEAT_TWO 1.0

CHILD SIT SIT_SIT_SIT 1.0

CHILD SIT_SIT_SIT NOBODY_SIT 1.0

CHILD NOBODY_SIT STAY_SIT1 1.0

CHILD NOBODY_SIT STAY_SIT2 1.0
```

```
CHILD NOBODY_SIT STAY_SIT3 1.0

CHILD NOBODY_SIT STAY_SIT4 1.0

CHILD SIT_SIT_SIT TRUCKER_SIT 1.0

CHILD TRUCKER_SIT OCC_POIN_AT_TRUCKER 1.0

CHILD OCC_POIN_AT_TRUCKER LK_AT_TRUCKER_G_LEFT 1.0

CHILD OCC_POIN_AT_TRUCKER LK_AT_TRUCKER_G_RIGHT 1.0

CHILD OCC_POIN_AT_TRUCKER LK_AT_TRUCKER_G_MIDDLE 1.0

CHILD TRUCKER_SIT LK_AT_TRUCKER_R 1.0

CHILD LK_AT_TRUCKER_R LK_AT_TRUCKER_RH 1.0

CHILD LK_AT_TRUCKER_R LK_AT_TRUCKER_RF 1.0

CHILD LK_AT_TRUCKER_R LK_AT_TRUCKER_RG 1.0

CHILD TRUCKER_SIT LK_AT_TRUCKER_L 1.0

CHILD LK_AT_TRUCKER_L LK_AT_TRUCKER_LH 1.0

CHILD LK_AT_TRUCKER_L LK_AT_TRUCKER_LF 1.0

CHILD LK_AT_TRUCKER_L LK_AT_TRUCKER_LG 1.0
```

```
CHILD SIT_SIT_SIT SALESMAN_SIT 1.0

CHILD SALESMAN_SIT OCC_POIN_AT_SALESMAN 1.0

CHILD OCC_POIN_AT_SALESMAN LK_AT_SALESMAN_G_LEFT 1.0

CHILD OCC_POIN_AT_SALESMAN LK_AT_SALESMAN_G_RIGHT 1.0

CHILD OCC_POIN_AT_SALESMAN LK_AT_SALESMAN_G_MIDDLE 1.0

CHILD SALESMAN_SIT LK_AT_SALESMAN_R 1.0

CHILD LK_AT_SALESMAN_R LK_AT_SALESMAN_RH 1.0

CHILD LK_AT_SALESMAN_R LK_AT_SALESMAN_RF 1.0

CHILD LK_AT_SALESMAN_R LK_AT_SALESMAN_RG 1.0

CHILD SALESMAN_SIT LK_AT_SALESMAN_L 1.0

CHILD LK_AT_SALESMAN_L LK_AT_SALESMAN_LH 1.0

CHILD LK_AT_SALESMAN_L LK_AT_SALESMAN_LF 1.0

CHILD LK_AT_SALESMAN_L LK_AT_SALESMAN_LG 1.0

CHILD PRIMARY_BEHAVIORS POSSESSED 1.0

CHILD POSSESSED WALK_TO_DOOR_INSIDE_FIRST 1.0
```

```
CHILD POSSESSED WALK_TO_DANCE_POINT_LEFT 1.0

CHILD POSSESSED WALK_TO_DANCE_POINT_MIDDLE 1.0

CHILD POSSESSED WALK_TO_DANCE_POINT_RIGHT 1.0

CHILD POSSESSED POSSESSED_GESTURE_SIT 1.0

CHILD POSSESSED POSSESSED_GESTURE_STAND 1.0

CHILD PRIMARY_BEHAVIORS DANCE 1.0

CHILD DANCE HAD_TO_CALM_DOWN 1.0

CHILD DANCE DEFAULT 1.0

CHILD DEFAULT CONFUSED 1.0

CHILD DANCE TWIRL_LEFT 1.0

CHILD DANCE TWIRL_RIGHT 1.0

CHILD DANCE TWIRL_BOTH 1.0

CHILD DANCE LIFT_FORWARD_LEFT 1.0

CHILD DANCE LIFT_FORWARD_RIGHT 1.0

CHILD DANCE LIFT_FORWARD_BOTH 1.0
```

```
CHILD DANCE LIFT_SIDEWAYS_LEFT 1.0

CHILD DANCE LIFT_SIDEWAYS_RIGHT 1.0

CHILD DANCE LIFT_SIDEWAYS_BOTH 1.0

CHILD DANCE CROSSOVER 1.0

CHILD DANCE KNEE_LIFT_LEFT 1.0

CHILD DANCE KNEE_LIFT_RIGHT 1.0

CHILD DANCE KNEE_LIFT_BOTH 1.0

CHILD DANCE DANCE_WALK 1.0

CHILD DANCE SHUFFLE 1.0

CHILD DANCE TWIST_LEFT 1.0

CHILD DANCE TWIST_RIGHT 1.0

CHILD PRIMARY_BEHAVIORS UNPOSSESSED 1.0

CHILD UNPOSSESSED UNPOSSESSED_GESTURE 1.0

CHILD PRIMARY_BEHAVIORS WALK_OUT 1.0

CHILD WALK_OUT WALK_TO_DOOR_INSIDE 1.0
```

```
CHILD WALK_OUT WALK_TO_DOOR_OUTSIDE 1.0

CHILD WALK_OUT STAND_SNAP 1.0

CHILD WALK_OUT STAND_SNAP3 1.0

CHILD AUTONOMIC_BEHAVIORS LOOK_AT_CURRENT_PRONOME 1.0

CHILD AUTONOMIC_BEHAVIORS MAINTATAIN_BALANCE 1.0

CHILD AUTONOMIC_BEHAVIORS POST_EMOTION 1.0

CHILD AUTONOMIC_BEHAVIORS FACIAL_EXPRESSION 1.0

CHILD AUTONOMIC_BEHAVIORS SET_DANCE_POINT_TAGS 1.0

CHILD AUTONOMIC_BEHAVIORS BLINK 1.0
```

Appendix F

Survey sheet given to (void*) participants

In this experiment, participation is voluntary and no coercion to participate will be involved. You are free to withdraw your consent and to discontinue participation in the project or activity at any time. You may decline to answer any questions and confidentiality and anonymity are assured. The data collected will be reported in such a way that your identity is protected. Your sincere answers will help us build characters that are more lifelike and fun to interact with. Thanks for your cooperation.

Date_______________ Time _________________ Age ________________Gender __________

(Void *): A cast of characters

1. Information on the person who is participating in this survey.

Figure F-1: Figures of all the characters in the (void*) project and their neutral names were given to the survey participants to exclude any prejudice that might come from descriptive names that were internally used for programming purposes.

(1) Are you familiar with computer games?

(minimal) 0 --- 1 --- 2 --- 3 --- 4 --- 5 (maximal)

(2) How much are you interested in each of the following areas? (Please, write down 0

Psychology

Visual arts

Computer science

Biology

Sociology

Dancing

Other ___________________

2. Reactions / impressions:

[Before the interaction]

(1) How much extrovert would you expect Earl to be?

(minimal) 0 --- 1 --- 2 --- 3 --- 4 --- 5 (maximal)

(2) How much would you expect Earl to listen to other people's opinions?

(minimal) 0 --- 1 --- 2 --- 3 --- 4 --- 5 (maximal)

(3) How conscientious would Earl be in his behavior?

(minimal) 0 --- 1 --- 2 --- 3 --- 4 --- 5 (maximal)

(4) How easily would you expect Earl's emotional state to change?

(minimal) 0 --- 1 --- 2 --- 3 --- 4 --- 5 (maximal)

(5) Which of the following can describe Earl's personality?

stern friendly dominant submissive angry joyful amenable outgoing gentle

(6) How much extrovert would you expect Elliot to be?

(minimal) 0 --- 1 --- 2 --- 3 --- 4 --- 5 (maximal)

(7) How much would you expect Elliot to listen to other people's opinions?

(minimal) 0 --- 1 --- 2 --- 3 --- 4 --- 5 (maximal)

(8) How conscientious would Elliot be in his behavior?

(minimal) 0 --- 1 --- 2 --- 3 --- 4 --- 5 (maximal)

(9) How easily would you expect Elliot's emotional state to change?

(minimal) 0 --- 1 --- 2 --- 3 --- 4 --- 5 (maximal)

(10) Which of the following can describe Elliot's personality?

stern friendly dominant submissive angry joyful amenable outgoing gentle

(11) How much extrovert would you expect Eddie to be?

(minimal) 0 --- 1 --- 2 --- 3 --- 4 --- 5 (maximal)

(12) How much would you expect Eddie to listen to other people's opinions?

(minimal) 0 --- 1 --- 2 --- 3 --- 4 --- 5 (maximal)

(13) How conscientious would Eddie be in his behavior?

(minimal) 0 --- 1 --- 2 --- 3 --- 4 --- 5 (maximal)

(14) How easily would you expect Eddie's emotional state to change?

(minimal) 0 --- 1 --- 2 --- 3 --- 4 --- 5 (maximal)

(15) Which of the following can describe Eddie's personality?

stern friendly dominant submissive angry joyful amenable outgoing gentle

(16) Would it be fun to interact with Earl?

(minimal) 0 --- 1 --- 2 --- 3 --- 4 --- 5 (maximal)

(17) Would you expect Earl to like to interact with you?

(minimal) 0 --- 1 --- 2 --- 3 --- 4 --- 5 (maximal)

(18) Would it be fun to interact with Elliot?

(minimal) 0 --- 1 --- 2 --- 3 --- 4 --- 5 (maximal)

(19) Would you expect Elliot to like to interact with you?

(minimal) 0 --- 1 --- 2 --- 3 --- 4 --- 5 (maximal)

(20) Would it be fun to interact with Eddie?

(minimal) 0 --- 1 --- 2 --- 3 --- 4 --- 5 (maximal)

(21) Would you expect Eddie to like to interact with you?

(minimal) 0 --- 1 --- 2 --- 3 --- 4 --- 5 (maximal)

[After the interaction]

(1) Which character(s) did you interact with?

Earl   Elliot   Eddie

(2) How much extrovert would Earl to be?

(minimal) 0 --- 1 --- 2 --- 3 --- 4 --- 5 (maximal)

(3) How much would you expect Earl to listen to other people's opinions?

(minimal) 0 --- 1 --- 2 --- 3 --- 4 --- 5 (maximal)

(4) How conscientious would Earl be in his behavior?

(minimal) 0 --- 1 --- 2 --- 3 --- 4 --- 5 (maximal)

(5) How easily would Earl's emotional state change?

(minimal) 0 --- 1 --- 2 --- 3 --- 4 --- 5 (maximal)

(6) Which of the following can describe Earl's personality?

stern   friendly   dominant   submissive   angry   joyful   amenable   outgoing   gentle

(7) How much extrovert would Elliot to be?

(minimal) 0 --- 1 --- 2 --- 3 --- 4 --- 5 (maximal)

(8) How much would you expect Elliot to listen to other people's opinions?

(minimal) 0 --- 1 --- 2 --- 3 --- 4 --- 5 (maximal)

(9) How conscientious would Elliot be in his behavior?

(minimal) 0 --- 1 --- 2 --- 3 --- 4 --- 5 (maximal)

(10) How easily would Elliot's emotional state change?

(minimal) 0 --- 1 --- 2 --- 3 --- 4 --- 5 (maximal)

(11) Which of the following can describe Elliot's personality?

stern friendly dominant submissive angry joyful amenable outgoing gentle

(12) How much extrovert would Eddie to be?

(minimal) 0 --- 1 --- 2 --- 3 --- 4 --- 5 (maximal)

(13) How much would you expect Eddie to listen to other people's opinions?

(minimal) 0 --- 1 --- 2 --- 3 --- 4 --- 5 (maximal)

(14) How conscientious would Eddie be in his behavior?

(minimal) 0 --- 1 --- 2 --- 3 --- 4 --- 5 (maximal)

(15) How easily would Eddie's emotional state change?

(minimal) 0 --- 1 --- 2 --- 3 --- 4 --- 5 (maximal)

(16) Which of the following can describe Eddie's personality?

stern friendly dominant submissive angry joyful amenable outgoing gentle

(17) Was it fun to interact with Earl?

(minimal) 0 --- 1 --- 2 --- 3 --- 4 --- 5 (maximal)

(18) Did Earl initially like to interact with you?

(minimal) 0 --- 1 --- 2 --- 3 --- 4 --- 5 (maximal)

(19) Did Earl like to interact with you later?

(minimal) 0 --- 1 --- 2 --- 3 --- 4 --- 5 (maximal)

(20) Was it fun to interact with Elliot?

(minimal) 0 --- 1 --- 2 --- 3 --- 4 --- 5 (maximal)

(21) Did Elliot initially like to interact with you?

(minimal) 0 --- 1 --- 2 --- 3 --- 4 --- 5 (maximal)

(22) Did Elliot like to interact with you later?

(minimal) 0 --- 1 --- 2 --- 3 --- 4 --- 5 (maximal)

(23) Was it fun to interact with Eddie?

(minimal) 0 --- 1 --- 2 --- 3 --- 4 --- 5 (maximal)

(24) Did Eddie initially like to interact with you?

(minimal) 0 --- 1 --- 2 --- 3 --- 4 --- 5 (maximal)

(25) Did Eddie like to interact with you later?

(minimal) 0 --- 1 --- 2 --- 3 --- 4 --- 5 (maximal)

(26) Which character is most like yourself?

(27) Any comments (on this experiment, the project, the characters, etc.)?

Bibliography

[1] C. W. Anderson. Strategy learning with multilayer connectionist representations. In *Proceedings of the Fourth International Workshop on Machine Learning*, 1989.

[2] Joseph Bates. The nature of character in interactive worlds and the oz project. Technical Report CMU-CS-92-200, School of Computer Science, Carnegie Mellon University, October 1992.

[3] J. O. Berger. *Statistical decision theory and Bayesian analysis.* Springer-Verlag, 1985.

[4] C. M. Bishop. *Neural networks for pattern recognition.* Oxford University Press, Oxford, UK, 1995.

[5] M. E. Bitterman. Thorndike and the problem of animal intelligence. *American Psychologist*, 24:444–453, 1969.

[6] B. Blumberg, P. Todd, and P. Maes. No bad dogs: Ethological lessons for learning. In *From Animals to Animats, Proceedings of the Fourth International Conference on the Simulation of Adaptive Behavior*, Cambridge, MA, 1996. MIT Pres.

[7] Bruce Blumberg. *Old Tricks, New Dogs: Ethology and Interactive Creatures.* Ph. d in media arts and sciences, The Media Lab, Massachusetts Institute of Technology, Cambridge, MA, September 1996.

[8] Bruce Blumberg and Tinsley Galyean. Multi-level direction of autonomous creatures for real-time virtual environment. In *Computer Graphics Proceedings, SIGGRAPH 95*, Los Angeles, California, August 1995.

[9] Bruce Blumberg, Michael Patrick Johnson, Michal Hlavac, Christopher Kline, Ken Russell, Bill Tomlinson, Song-Yee Yoon, Andrew Wilson, Teresa Marrin, Joseph Paradiso, Zoe Teegarden, and Dan Stiehl. Swamped!: An interactive installation with a sympathetic interface. In *Emerging Technology, SIGGRAPH '98*, Orlando, FL, August 1998.

[10] R. C. Bolles. *Theory of motivation*. Harper and Row, New York, NY, 2 edition, 1975.

[11] V. Braitenberg. *Vehicles: experiments in synthetic psychology*. MIT Press, Cambridge, MA, 1984.

[12] C. Breazeal(Ferrell). A motivational system for regulating human-robot interaction. In *Proceedings of AAAI '98*, Madison, WI, 1998.

[13] M. R. Brent and T. A. Cartwright. Distributional regularity and phonotactic constraints are useful for segmentation. *Cognition*, 61:93–125, 1996.

[14] Alf Brodal. *Neurological anatomy in relation to clinical medicine*. Oxford University Press, New York, NY, 1981.

[15] Rodney Brooks. A robust layered control system for a mobile robot. *IEE Journal of Robotics and Automation*, 2(1), March 1986.

[16] Rodney Brooks. Challanges for complete creature architecture. In *From Animals to Animats: Proceedings of the First International Conference on Simulation of Adaptive Behavior*, London, England, August 1990. A Bradford Book, The MIT Press.

[17] Rodney A. Brooks, Cynthia Breazeal, Matthew Marjanovic, Brian Scasselati, and Matthew Williamson. The cog project: Building a humanoid robot. *To appear in a Springer-Verlag Lecture Notes in Computer Science Volume.*

[18] J. Brown. Some tests of the decay theory of immediate memory. *Quarterly Journal of Experimental Psychology*, 10:12–21, 1958.

[19] J. A. Bruner, J. S. Goodnow, and G. S. Austin. *A study of thinking.* Wiley, 1956.

[20] Mary R. Burch and Jon S. Bailey. *How dogs learn.* Howell Book House, New York, NY, 1999.

[21] Robert C. Burke. Towards a synthetic ear for a virtual canine. Technical report, MIT 9.520, Cambridge, MA, Fall 1999.

[22] Ann B. Butler and William Hodos. *Comparative Vertebrate Neuroanatomy: Evolution and Adaptation.* John Wiley and Sons, 1996.

[23] M. Cabanac. Pleasure: the common currency. *Journal of Theoretical Biology*, 155:171–200, 1992.

[24] Anthony Chavez, Alexandros G. Moukas, and Pattie Maes. Challenger: A multiagent system for distributed resource allocation. In *Proceedings of the International Conference on Autonomous Agents*, Marina Del Ray, CA, 1997.

[25] C. M. Chew and G. A. Pratt. A minimum model adaptive control approach for a planar biped. In *Proceedings of IEEE/RSJ International Conference on Intelligence Robots and Systems*, Kyongju, Korea, October 1999.

[26] Noam Chomsky. *The Minimalist Program.* MIT Press, 1995.

[27] Antonio Damasio. *Descartes' Error; Emotion, Reason, and the Human Brain.* Putnam Publishing Group, 1994.

[28] Charles Darwin. *The expression of the emotions in man and animals.* University of Chicago Press, Chicago, Il, 1965.

[29] T. L. Dean and M. P. Wellman. *Planning and Control.* Morgan Kaufman Publishers, San Mateo, CA, 1991.

[30] D. Dennett. *The Intentional Stance.* MIT Press, Cambridge, MA, 1987.

[31] R. Descartes. *Treatise of man.* Harvard University Press, Cambridge, MA, 1972.

[32] R. O. Duda and P. E. Hart. *Pattern Classification and Scene Analysis.* Wiley, 1973.

[33] I. Eibl-Eibesfeldt. *Ethology: The biology of behavior.* Holt, Rinehart and Winston, New York, NY, 2 edition, 1975.

[34] I. Eibl-Eibesfeldt. *Human Ethology.* Aldine de Gruyter, Hawthorne, NY, 1989.

[35] P. Ekman. *Emotion in the human face.* Cambridge University Press, Cambridge, UK, 2 edition, 1982.

[36] P. Ekman. Facial expression and emotion. *American Psychologist*, 48:384–392, 1993.

[37] M. S. El-Nasr, T. R. Ioerger, and J. Yen. A web of emotions. In *Workshop on Emotion based Agent Architecture (EBAA '99)*, Seattle, WA, 1999.

[38] R. Elio and J. R. Anderson. The effects of information order and learning mode on schema abstraction. *Memory and Cognition*, 12:20–30, 1984.

[39] A. J. Fridhund. *Human Facial Expression: An Evolutionary View.* Academic Press, 1994.

[40] Michael S. Gazzaniga. *Cognitive Neuroscience.* W. W. Norton, New York, NY, 1998.

[41] L. R. Goldberg. An alternative description of personality: The big-five factor structure. *Journal of Personality and Social Psychology*, 59:1216–1229, 1990.

[42] A. I. Goldman. *Epistemology and Cognition.* Harvard University Press, Cambridge, MA, 1986.

[43] A. M. Graybiel, T. Aosaki, A. W. Flaherty, and M. Kimura. The basal ganglia and adaptive motor control. *Science*, 265:1826–1830, 1994.

[44] Simon S. Haykin. *Neural Networks*. Prentice Hall, Upper Saddle River, NJ., 1999.

[45] Donald O. Hebb. *Essay on mind*. L. Erlbaum Associates, Hillsdale, NJ, 1980.

[46] David Heckerman. Learning bayesian networks. In *The Ninth Annual Conference on Neural Information Processing Systems*, Denver, CO, November 1995.

[47] Charles Judson Herrick. *The thinking machine*. The University of Chicago Press, Chicago, IL, 1929.

[48] John A. Hertz, Richard G. Palmer, and Anders Krogh. *Introduction to the Theory of Neural Computation*, volume 1 of *Santa Fe Institute studies in the sciences of complexity*. Addisson-Wesley Pub. Co., Redwood City, CA, 1991.

[49] Jerry A. Hogan. *Causal Mechanisms of Behavioural Development*, chapter 1. The concept of cause in the study of behavior, pages 3–15. Cambridge University Press, 1994.

[50] John H. Holland. *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control and Artificial Intelligence*. A Bradford Book, The MIT Press, Cambridge, MA, 1 edition, 1992.

[51] C. L. Hull. *Principles of Behavior*. Appleton-Century-Crofts, New York, NY, 1943.

[52] E. B. Hunt. *Concept learning, an information processing problem*. Wiley, 1962.

[53] C. E. Izard. *Human Emotions*. Plenum, New York, NY, 1977.

[54] C. E. Izard. *Emotions in personality and psychopathology*. Plenum Press, New York, NY, 1979.

[55] C. E. Izard. *Measuring emotions in infants and children.* Cambridge University Press, Cambridge, UK, 1982.

[56] C. E. Izard. *The Psychology of Emotion.* Plenun, 1991.

[57] William James. *The principles of psychology*, volume 2. Holt, New York, NY, 1890.

[58] H. Jeffreys. *Theory of Probability.* Clarendon Press, 3 edition, 1961.

[59] F. V. Jensen and F. V. Jensen. *Introduction to Bayesian Networks.* Springer-Verlag, New York, NY, 1996.

[60] Michael P. Johnson. Multi-dimensional quaternion interpolation. In *ACM SIGGRAPH 99 Conference Abstracts and Applications*, page 258, August 1999.

[61] R. P. Kesner, B. L. Bolland, and M. Dakis. Memory for spatial locations, motor responses, and objects: triple dissociation among the hippocampus, caudate nucleus, and extrastriate visual cortex. *Experimental Brain Research*, 93:462–470, 1993.

[62] C. Kline and B. Blumberg. The art and science of synthetic character design. In *Proceedings of the AISB 1999 Symposium on AI and Creativity in Entertainment and Virtual Art*, Edinburgh, Scotland, 1999.

[63] John R. Koza. Genetic programming: A paradigm for genetically breeding populations of computer programs to solve problems. Technical Report STAN-CS-90-1314, Department of Computer Science, Stanford University, Stanford, CA, 1990.

[64] John R. Koza. *Genetic Programming.* MIT Press, Cambridge, MA, 1992.

[65] Long-Ji Lin. Self-improve reactive agents based on reinforcement learning, planning and teaching. *Machine Learning*, 8:293–321, 1992.

[66] Michael L. Littman. An optimization-based categorization of reinforcement learning environments. In *From Animals to Animats 2, Proceedings of the Second International Conference on Simulation of Adaptive Behavior*, pages 262–270, Cambridge, MA, 1992. MIT Press.

[67] K. Lorenz. Evolution of ritualization in the biological and cultural spheres. *Philosophical Transactions of the Royal Society of Briatain*, 251:273–284, 1966.

[68] K. Lorenz. *The Foundation of Ethology.* Springer-Verlag, New York, NY, 1981.

[69] W. Lyons. Emotion. Cambridge Studies in Philosophy Series. Cambridge University Press, 1980.

[70] N. J. MacKintosh. *Conditioning and Associative Learning.* Oxford University Press, Oxford, UK, 1983.

[71] P. Maes, T. Darrell, B. Blumberg, and A. Pentland. The alive system: Wireless, full-body interaction with autonomous agents. *Communications of the ACM, Special Issue on New Horizons of Commercial and Industrial AI*, 38(19), November 1995.

[72] Pattie Maes. How to do the right thing. *Connection Science Journal*, 1(3):291–323, 1989.

[73] N. R. F. Maier and T. C. Schneirla. *Principles of Animal Psychology.* McGraw-Hill Book Company, New York, NY, 1935.

[74] G. Mandler. *Mind and Emotion.* Wiley, New York, NY, 1975.

[75] Maja J. Mataric. Coordination and learning in multi-robot systems. *IEEE Intelligent Systems*, pages 6–8, March/April 1998.

[76] Michael Mateas. An oz-centric review of interactive drama and believable agents. Technical Report CMU-CS-97-156, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, June 1997.

[77] A. K. McCallum. Learning to use selective attention and short-term memory in sequential tasks. In *From Animals To Animats, Proceedings of the Fourth International Conference on the Simulation of Adaptive Behavior*, Cambridge, MA, September 1996. MIT Press.

[78] R. R. McCrae and P. T. Costa. Validation of the five factor model of personality across instruments and observers. *Journal of Personality and Social Psychology*, 52:81–90, 1987.

[79] R. J. McDonald and N. M. White. A triple dissociation of memory systems: Hippocampus, amygdala and dorsal striatum. *Behavioral Neoroscience*, 107:3–22, 1993.

[80] D. J. McFarland and R. M. Sibly. The behavioral final common path. In *Philosophical Transactions of the Royal Society of London*, volume 270 of *Series B. Biological Sciences*, pages 265–293. 1974.

[81] J. K. McGaugh. Time-dependent processes in memory storage. *Science*, 153:1351–1358, 1966.

[82] Neil McNaughton. *Biology and emotion*. Cambridge University Press, 1989.

[83] Marvin Minsky. *The Society of Mind*. Simon and Schuster, New York, NY, 1988.

[84] T. M. Mitchell. Generalization as search. *Artificial Intelligence*, 18:203–226, 1982.

[85] O. H. Mowrer. *Learning Theory and Behavior*. John Wiley and Sons, Inc., New York, NY, 1960.

[86] S. Narayanan and D. Jurafsky. Bayesian models of human sentence processing. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, volume 20. L. Erlbaum, 1998.

[87] Walle J. H. Nauta and M. Feirtag. *Fundamental neuroanatomy*. Freeman, New York, NY, 1986.

[88] John A. Nevin. *The study of behavior*. Scott, Foresman and Company, Glenview, IL, 1973.

[89] J. Panksepp. Toward a general psychological theory of emotions. *Behavioural and Brain Sciences*, 5:407–420, 1982.

[90] Judea Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann Publishers, INC., San Francisco, CA, 2 edition, 1988.

[91] L. R. Peterson and M. J. Peterson. Short-term retention of individual verbal items. *Journal of Experimental Psychology*, 58:193–198, 1959.

[92] R. Plutchik. *Emotion, theory, research, and experience*. Academic Press, San Diego, CA, 1980.

[93] R. Plutchik. *The emotions*. University Press of America, Lanham, MD, 1991.

[94] Karen Pryor. *Clicker Training for Dogs*. Sunshine Books, Inc., Waltham, MA, 1999.

[95] Lawrence Rabiner and Biing-Hwang Juang. *Fundamentals of Speech Recognition*. Prentice-Hall, New York, NY, 1993.

[96] S. Reilly. *Believable Social and Emotional Agents*. PhD thesis, School of Computer Science, Carnegie Mellon University, Philadelphia, PA, 1996.

[97] R. A. Rescorla. Behavioral studies of pavlovian conditioning. *Annual Review of Neuroscience*, 11:329–352, 1988.

[98] M. Z. Rosenzweig, E. L Bennett, P. J. Colombo, and P. A. Serrano. Short-term, intermediate-term and long-term memory. 57:193–198, 1993.

[99] Mark R. Rosenzweig, Arnold L. Leiman, and S. Marc Breedlove. *Biological psychology: an introduction to behavioral, cognitive, and clinical neuroscience.* Sinauer Associates, Inc., Sunderland, MA, 1999.

[100] D. E. Rumelhart and J. L. McClelland. *Parallel distributed processing.* MIT Press, 1986.

[101] J. Russell. A circumplex model of affect. *Journal of Personality and Social Psychology*, 39(6):1161–1178, 1980.

[102] Michael K. Sahota. Action selection for robots in dynamic environments through inter- behaviour bidding 138. In *From Animals to Animats 3*, Brighton, England, August 1994.

[103] G. Saucier. Effects of variable selection on the factor structure of person-descriptors. *Journal of Personality and Social Psychology*, 73:1296–1312, 1997.

[104] J. Schmidhuber. Reinforcement learning in markovian and non-markovian environments. In *Advances in Neural Information Processing Systems*, volume 3, pages 500–506. Morgan Kaufmann, 1991.

[105] J. P. Scott. The emotional basis of social behavior. *Annals of the New York Academy of Sciences*, 159:777–790, 1969.

[106] R. N. Shepard and P. Arabie. Additive clustering: Representation of similarities as combinations of discrete overlapping properties. *Psychological Review*, 86:87–123, 1979.

[107] R. N. Shepard, C. I. Hovland, and H. M. Jenkins. Learning and memorization of classifications. *Psychological monographs*, 75, 1961.

[108] Edward Hance Shorliffe. *Computer-based medical consultations, MYCIN.* Artificial Intellifence Series; Elsevier computer science library. Elsevier, New York, NY, 1976.

[109] Sidney Siegel. *Nonparametric Statistics for the Behavioral Sciences.* McGraw-Hill Book Company, Inc., 1956.

[110] B. F. Skinner. *The behavior of organisms.* Appleton Century Crofs, New York, NY, 1938.

[111] R. C. Solomon. The logic of emotion. *Nous*, (11), 1977.

[112] Emmet H. Spier. *From Reactive Behaviour to Adaptive Behaviour: Motivational models for behaviour in animals and robots.* Ph.d thesis, Balliol College, University of Oxford, 1997.

[113] J. E. R. Staddon. *Adaptive behavior and learning.* Cambridge University Press, Cambridge, UK, 1983.

[114] N. L. Stein and L. J. Levine. *Thinking about feelings: the development and organization of emotional knowledge*, volume 3, pages 165–198. Lawrence Erlbaum Associates, Hillside, NJ, 1987.

[115] N. L. Stein and T. Trabasso. *Children's understanding of changing emotional states.* Cambridge University Press, New York, NY, 1989.

[116] Nancy L. Stein and Linda J. Levine. *Making sense out of emotion.* Lawrence Erlbaum Associates, Hillside, NJ, 1990.

[117] K. T. Strongman. *The psychology of emotion.* Wiley, New York, NY, 2 edition, 1978.

[118] Richard S. Sutton and Andrew G. Barto. *Reinforcement Learning: An introduction.* MIT Press, Cambridge, MA, 1998.

[119] Joshua B. Tenenbaum. Learning the structure of similarity. In *Advances in Neural Information Processing Systems*, volume 8, pages 3–9, Cambridge, MA, 1996. MIT Press.

[120] Joshua B. Tenenbaum. *A Bayesian framework for concept learning.* Ph. d in brain and cogniive sciences, Massachusetts Institute of Technology, Cambridge, MA, February 1999.

[121] Joshua B. Tenenbaum. Bayesian modeling of human concept learning. In *Advances in Neural Information Processing Systems*, Cambridge, MA, 1999. MIT Press.

[122] D. Terzopoulos, X. Tu, and R. Grzeszczuk. Artificial fishes: Autonomous locomotion, perception, behavior, and learning in a simulated physical world. *Artificial Life*, 1:327–351, December 1994.

[123] Guy Theraulaz, Simon Goss, Jacques Gervet, and Jean-Louis Deneubourg. Task differentiation in polistes wasp colonies: A model for self-organizing groups of robots. In *From Animals to Animats, Proceedings of the First International Conference on Simulation of Adaptive Behavior*, London, England, August 1990. A Bradford Book, The MIT Press.

[124] Frank Thomas and Ollie Johnston. *The Illusion of Life: Disney Animation.* Abbeville Press, New York, NY, 1981.

[125] Edward L. Thorndike. *The fundamentals of learning.* Teachers College, Columbia University, New York, NY, 1932.

[126] M. Tinbergen. *The Study of Instinct.* Oxford University Press, New York, NY, 1951.

[127] N. Tinbergen. Derived activities: Their causation, biological significance, origin and emancipation during evolution. *Quarterly Review of Biology*, 27:1–32, 1952.

[128] N. Tinbergen. *The herring gull's world.* Collins, London, UK, 1953.

[129] Edward Chance Tolman. *Purposive behavior in animals and man.* Appleton Century Crofts, New York, NY, 1967.

[130] S. S. Tomkins and R. McCarter. What and where are the primary affects? some evidence for a theory. *Perceptual and Motor Skills*, 18:119–158, 1964.

[131] Silvan S. Tomkins. *Affect, cognition and personality: empirical studies*. Springer Pub. Co., New York, NY, 1965.

[132] William M. Tomlinson. Interactivity and emotion through cinematography. Masters thesis in media arts and sciences, The Media Lab, Massachusetts Institute of Technology, Cambridge, MA, May 1999.

[133] E. Tulving. Memory: Performance, knowledge, and experience. *european Journal of Cognitive Psychology*, 1:3–26, 1989.

[134] J. Velasquez. A computational framework for emotion-based control. In *Proceedings of the Grouding Emotions in Adaptive Systems Workshop, SAB '98*, Zurich, Switzerland, 1998.

[135] R. Ventra and C. Pinto-Ferreira. Emotion based agents: Three approaches to implementation (preliminary report). In *Workshop on Emotion based Agent Architecture (EBAA '99)*, Seattle, WA, 1999.

[136] K. von Frisch. Honeybees: Do they use direction and distance information provided by dancers? *Science*, 158:1073–1076, 1967.

[137] W. Grey Walter. *The Living Brain*. W. W. Norton and Company INC., New York, NY, 1953.

[138] Barbara Webb. Robotic experiments in cricket phonotaxis 45. In *From Animals to Animats 3*, Brighton, England, August 1994.

[139] Yair Weiss. *Bayesian motion estimation and segmentation*. PhD thesis, Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA, September 1998.

[140] Gary Wilkes. *Click and Treat Training Kit Version 1.2*. www.clickandtreat.com, Mesa, AZ, 1995.

[141] Steward W. Wilson. Classifier systems and the animate problem. *Machine Learning*, 2:199, 1987.

[142] J. Yamaguchi, N. Kinoshita, and I. Karto. Development of a dynamic biped walking system for humanoid: Development of a biped walking robot adapting to the humans' living floor. In *Proceedings of IEEE International Conference on Robotics and Automation*, pages 232–239, 1996.

[143] P. T. Young. The role of hedonic processes in motivation. In *Nebraska symposium on motivation*, pages 187–188. Lincoln: University of Nebraska Press, 1955.