Emotion and Learning

Emotion and Learning

A Computational Model of the Amygdala

Jan Morén

Lund University Cognitive Studies 93
2002

Typography and layout by LaTeX.
Printed in Sweden by JustNu, Lund

ISBN 91-628-5212-4
ISRN LUFDA/HFKO-1010
ISSN 1101-8453
Lund University Cognitive Studies 93

To procrastinators everywhere

Acknowledgements

A thesis is a strange publication. Though it is presented as the work of a sole author, in reality it is frequently only through the help and support of many people that it has come into existence. So it is with this one as well.

There are a number of people that in large or small part have helped through this process. Foremost among them, though is my advisor and colleague Dr. Christian Balkenius. He was my advisor for my Master's project in computer science and encouraged me to pursue doctoral studies rather than leaving academia for the industry. The thesis itself is the fruit of a joint project undertaken by me and Dr. Balkenius; without his help, suggestions, encouragement and support over the years, this thesis simply would not exist.

I would also like to thank Prof. Peter Gärdenfors who founded and leads the department of cognitive science, both for creating and maintaining such a stimulating and creative environment in which to work, as well as for believing that a somewhat disorganized and confused computer science graduate would have a future as a doctoral student at the department.

The other researchers and doctoral students at the department have also been of invaluable help. It would be no exaggeration to say that I've learned more about this field during our seminars and discussions than I could ever have been able to learn in any other way. In particular, I'd like to thank Lars Kopp, Petra Björne and Jens Månsson for many interesting discussions and arguments during this work. I'd also like to thank the people at the department for all the feedback and comments I've received during this work.

My friends and family have been very supportive. I would especially like to thank my friend Iréne Stenfors, with whom I have had any number of discussions and who has acted as a sounding board more often than I could count. I'd also like to thank Louise Gripenhov who proofread some of the chapters, giving me invaluable feedback on my use of language. And a big thank you to my parents and siblings and to all my other friends for showing such patience and understanding when I've let this work take precedence over their company.

Lund, August 02, 2002 Jan Morén

Contents

Chapter 1

Introduction

Emotions, and by extension, emotional conditioning, is becoming increasingly important not only as a part of the general study of learning, but as an important subject in its own right. Whereas emotions have previously been seen as a low-level reaction system at best, and as irrelevant to cognition at worst, today it has risen to prominence as an integral part of any biological and biologically inspired system.

In this thesis, we will look at the neurophysiological basis of emotions and at the features needed for a learning system using classical conditioning, and attempt to bring the needed requirements together. The goal is to develop a functional model of conditioning inspired by the neurophysiology and of classical conditioning, and test it in simulation. But first, we need to put emotions into perspective.

How does a biological organism benefit from having emotions? Clearly, there are great evolutionary benefits to having emotions, or the majority of animals would not have them, as is now the case. Indeed, emotions are a system that is present in all but the simplest of organisms. The easy answer is that emotions encodes information to the animal about what features in the world to like (and thus interact with in one way or another), and what not to like (and thus avoid).

Of course, every animal has a number of innate structures that guides its behavior. Most animals can identify their typical foods, predators and partners, and has likewise instinctive behaviors that will tend to guide the animal towards the proper interaction with desirable stimuli and away from undesir-

able stimuli. But not even these simple mappings between stimuli and reactions will be enough to keep the organism from encountering problems. While an instinct telling the animal to avoid large, speedy fish will keep it away from sharks, it will likewise have the animal fleeing tuna fish or even large codfish – a situation that will keep the animal constantly at its toes, and will cause it to expend a large amount of precious energy avoiding non-existent dangers.

One answer is of course to allow evolution to devise more specific detection algorithms to avoid dangers, and only the dangers. The problem is of course that many of the dangers are evolving at the same speed as the organism in question, getting steadily better at camouflage and deception. Another problem is that some dangers are so transient that there is no possibility for an organism to adapt over an evolutionary timescale. Thus an organism needs to adapt to its surroundings during its lifetime, and not just over an evolutionary timescale.

Consequently, it needs learning. With learning, animals are able to adapt in a matter of seconds, rather than generations. This speed for both learning and relearning also means that animals can adapt to highly specific, unusual and transient events.

To function, any learning system needs some kind of evaluation of the current situation, and feedback on whether the results of the learning really were beneficial or not. To some extent, these evaluations are built in; food and mates are good, pain and illness are bad. Most animals have a fairly large array of such inborn evaluations that are able to guide it over the course of its lifetime.

Of course, now we have a new problem. We can learn appropriate actions to take in real time, based on the innate evaluations, but the evaluations themselves are still developed in evolutionary time. We need a way to learn new evaluations during the animal's lifetime just as it can learn the proper reactions to them.

This is where the ability to condition emotional reactions in real time comes in. By being able to associate innate emotional stimuli with other stimuli, they can be given an emotional significance when needed. Just as importantly, these evaluations can be learned at a much greater level of specificity; they can be constrained to be valid only for a specific place; or a given time of day; or only when accompanied with other, specific stimuli.

The ability to learn emotional reactions is important for other reasons as well. In artificial systems there is no evolutionary development; the system is purposefully designed from the start. All adaptability for an artificial system must

thus be explicitly built in. This aspect of artificial systems – as well as the recurrent problems of flexibility for those systems – have given a focus on the concept of autonomy that studies of natural systems have not. We rarely ask how 'autonomous' an ant is, or if it is less autonomous than a beetle. We do ask these questions about artificial systems. We believe that the concept of autonomy is closely interconnected with the ability to adapt to changing evaluations as well as changing circumstance.

One way to look at this aspect of adaptability in terms of artificial systems is by looking at *needs* and *goals*. We propose that autonomy is the ability to generate goals internally, based on the needs of the system. These needs may, of course, be the result of design which makes it possible to control an autonomous system. The difference between defining the needs of a system and defining its goals may seem slight. However, we argue that there is a major difference between these two approaches. Whereas goals (as we speak of them) are task-specific, needs are a minimal set of objectives needed for the system to successfully exist in its environment. It is the difference between "wanting to find food" and "being hungry".

Defining goals for a system implies setting a predetermined prioritation among the possible activities the system is expected to perform. Any flexibility regarding the appropriateness for pursuing a given goal must be explicitly or implicitly built in by the designer. Another problem associated with explicit goals is that although the system may be able to generate subgoals, it cannot generate entirely new goals if the situation demands it. Setting goals means sacrificing flexibility and adaptability for control; the system is not fully autonomous.

Giving the system needs, then allowing the systems to generate internal goals, on the other hand, means that we give the system maximal opportunity to fulfill its mission in any way it sees fit. As we describe below, the needs are evaluated together with an emotional evaluation to generate an objective, which will subsequently drive action selection. Of course, as needs are rather more abstract, the designer would sacrifice some control over the system to achieve a greater degree of autonomy.

Traditionally, cognitive science and computer science have studied learning at the expense of its lesser-known counterparts (LeDoux, 1995). Emotion, learning and motivation can not be so easily separated, however. They are intertwined and depend heavily on one another; it can in fact be difficult to determine their boundaries at times. Motivation is what drives a system to actually do anything; without any motivation one way or another, there is no reason to act – or to learn. Emotions, on the other hand, indicate whether a chosen course of action was successful or not, and what maybe should have

been done instead. It thus gives constant feedback to the learning systems. Learning, of course, is the mechanism by which the emotional and motivational subsystems are able to adapt to an ever-changing environment.

1.1 Computational Modeling

There is currently a trend to work with computational models as a means of investigating phenomena. This raises the question of how computational models differ from other kinds of modeling, and what place modeling in general has in investigations.

A model is a simplified description of a phenomenon. It can be as tangible as a scale model of a physical object (such as an architectural model), or as vague as a conceptual description of a process. The pertinent point of a model is always that it brings forward some aspects of the modeled phenomenon at the cost of others. A model that doesn't throw away any properties would be useless; it would be identical with the phenomenon studied itself.

What properties to throw out and what to keep is of course dependent on what aspects of the phenomenon the investigator wants to study. For an architect, the fact that her paper model is not usable as a building and not constructed of the same materials as the real building is immaterial; she wants to convey and study its shape, and for this the paper model is sufficient. Likewise, a weather simulation has very little in common with 'real' weather, an economic model has no actual economic actors running around transferring real goods or money to one another, and a wind tunnel model of an airplane is utterly incapable of transporting passengers, or even fly by itself.

The nature of a model thus depends on what purpose the investigator has in designing it. An architect is concerned with form; a model railroader might be concerned with visual appearance or with simulating timetables for real railroads; a meteorologist is concerned with large-scale physical trends in the behavior of the atmosphere.

A computational model is a model described in such a way that it can be mathematically analyzed or implemented in a computer simulation. Note that while most computational models are virtual (ie. expressed as mathematics or as code), they do not need to be; economic models have sometimes been implemented as a physical system of reservoirs valves and pipes, with the flow of wealth represented as water flowing through the system . While seemingly unorthodox, this certainly qualifies as a computational model.

In the context of describing the functions of brain areas, there are really two kinds of models: descriptive and computational. A descriptive model is just that; it is a conceptual description of the functionality of each element (whether the element is a single neuron, or major brain systems) and their interconnections, with inferences based on this description of how the elements interact. The strength of this kind of model is that it neatly encapsulates the critical features of the system in question to make its functioning easy to grasp. The drawback is of course that it isn't easily testable for veracity; it is all too easy to convince oneself that the model explains a given phenomenon when in fact it does not.

A computational model, on the other hand, is designed to be testable, either through mathematical analysis, or in simulation. Each element is thoroughly specified in a mathematical notation, or in a manner that can easily be translated to mathematics without loss of meaning. The interconnections between elements is also fully specified. The advantage of this approach is that it gives investigators a way to 'empirically' test their theories in a controlled manner.

There are a number of caveats with using computational models, however. Writing or building, and running simulations are very satisfying activities with immediate, tangible results, and it is easy to be seduced by this ease and seeming relevance. The first problem is to control what it is you actually simulate. When designing a computational model it is all too easy to tailor the model to the medium, rather than to the phenomenon you wish to model. Any medium imposes constraints on its expression, and especially in the case of computer simulation, cutting corners or changing the dynamics of the model can happen even without the investigators realizing it.

With an 'accurate' model – ie. a model expressed in simulation the way it was originally intended – there are still potential pitfalls. The model will produce beautiful numerical results, expressed to as many decimals as the investigators want, and it is a common – and understandable – failing to interpret this as being *accurate* to all these decimals. It generally is not, of course; the model itself is only an approximation of the phenomenon under study.

Finally, the model can be confused with the phenomenon itself. As the phenomenon under study is often complex or abstract – good reasons to work with models in the first place – it is easy to look at the model as providing better data than it in reality can do. Especially with complex models, there is a danger of over-interpreting the results, and see features where none exist. All these aspects must be taken into consideration when working with models as an investigative tool.

1.1.1 System Level Modeling of the Brain

When modeling a phenomenon, a choice has to be made at what level the model should function. In the case of simulation of psychological or neurophysiological phenomena there are several levels to work with, from biophysical simulations of cell-level dynamics all the way up to large-scale behavioral models. Exploration at any of these levels is of course worthwhile, but there are drawbacks as well.

Low-level models are most often focused on the neurophysiological function and its biological underpinnings. While these models have great explanatory power over the physiological structure, they are usually focused on such small, specific structures that they leave questions of the structures' role in the larger system unanswered. The constraints used to build the model come almost solely from neurophysiological data, leaving behavioral and psychophysical data behind.

In contrast, high-level models of phenomena like learning, attention, spatial navigation and memory are often constrained only by psychological or psychophysical data. They are frequently not concerned with the neurophysiological or anatomical structures that implement the underlying functionality.

We attempt to explore the viability of using simulations at an intermediate level to study the processes implementing emotional conditioning. As we will see, the simulations used are constrained both by neurophysiological and anatomical data, as well as behavioral data. It is thus neither purely at a physiological, nor at a behavioral level. Instead, this is an instance of system level modeling. The goal is to study behavior using a physiologically constrained model. At this level we try to take into account both the functionality of individual areas, and the interactions between areas, as defined from empirical data. The modeling of an individual area is strictly functional; at this level of analysis, the specifications of individual cell assemblies is not of interest. These functional modules are however interconnected in very much the same way as the real areas that they model.

1.2 Emotions

One major problem in dealing with emotional processes is the confusing terminology associated with the area. Terms properly associated with emotions have been expropriated for use in other fields and, not uncommonly, the same phenomenon have been given different labels depending on the field in question.

When we talk about emotions in this text, we do not discuss the subjective feeling that we experience, but a reaction to a stimulus as being emotionally charged. Such emotions are either *primary* or *secondary*.

Primary emotions are generated by stimuli or contexts directly and intrinsically related to the needs of the system. These can be things like the smell of food, pain or sexual signals. These stimuli do not need to be associated with other emotional stimuli and are resistant to any change in their effectiveness. However, their expression can be inhibited by other systems.

Secondary (or higher order) emotions are stimuli that are not emotionally charged in themselves, but that becomes emotional through association with a primary emotional stimuli (where external reinforcers would count as such), or with secondary stimuli already emotionally charged. This association enables these stimuli to act the same as primary stimuli, being used for motivation, as well as directing attention. In effect, secondary emotions predict the possible occurrence of primary emotions.

Although the emotional system will react to both pleasant and unpleasant stimuli, most of the work in this area has been focused on fear (LeDoux and Fellous, 1995). Fear can be defined as the reaction to a signal that predicts punishment. This signal is said to be *aversive*. Fear in this sense is often equated with anxiety (Gray, 1982). Fear as an emotional state gives rise to *avoidance* behaviors. When an animal feels fear it will try to avoid whatever made it fearful.

We have to distinguish between passive and active avoidance: passive avoidance is to refrain from doing something, as the consequences would be negative; active avoidance is to actively behave so as to avoid a negative consequence. This is a larger difference than it may seem. With active avoidance, the animal learns what to do to avoid a negative outcome – there is a well-defined course of action to take. When an animal learns to avoid doing something for concern of the consequences, as in passive avoidance, it receives no guidance on what to do instead.

A real life example would be a child that wants to light a match. If the parent just yells ”No!” when she lights a match, she does not learn how to correctly approach her goal. All she’s learned is that lighting a match is a bad idea. She will not know how or when it is all right to do so, and will not know *why* lighting the match was a bad thing to do. She can not map this knowledge in any useful way to other situations. If, on the other hand, the parent tells the child not to light the match, but to always ask a parent instead, she will have a positive course of action whenever she wants to accomplish something that may require the lighting of a match.

When a system does not receive an expected reinforcer (or positively charged stimulus), the result is *frustration*, and anger is a reaction to this (Rolls, 1995). If an agent bumps into a wall unexpectedly, for instance, this may make it frustrated since it was not able to move closer to its goal which could result in an aggressive reaction towards the wall or anything else that happens to be around at the moment.

Positive emotions are less well defined but can be called hope or anticipation (Panksepp, 1981). These emotions are reactions to *appetitive* stimuli that are rewarding or predict reward (Rolls, 1995), such as the smell or sight of food, or a potential mate.

Stimuli can also become emotionally charged in the absence of primary stimuli, if they are unexpected (Gray, 1982). They will arouse interest and direct attention to them for further evaluation. The emotional reaction to novelty is both aversive and appetitive; this is an approach-avoidance conflict caused by a lack of experience with the stimulus (Lewin, 1936).

Common terms associated with emotions are *motivation* and *drive*. Although sometimes thought of as the same thing, they are in fact rather distinct. Hebb (1955) popularized the drive concept and described a model accounting for its function. He saw a drive as motivational energy driving the organism towards activity while leaving the nature of the needed activity largely unspecified. He also relates the concept of drive to reward and punishment. An optimal drive level is rewarding while a too high or low drive is punishing. Since novelty increases drive, a moderate level of novelty is rewarding.

Motivation is in our view a much more focused concept than drive; it is the combination of internal needs, emotional state and context (Balkenius, 1995). It does not only describe what to accomplish, but also wholly or in part how to do it. Unlike drive, motivation is directed toward specific stimuli; drive may indicate that the animal is hungry, whereas motivation will indicate that a specific, present, foodstuff is very desirable right now.

So, what about the feelings we associate with emotions? A common theory, described by LeDoux (1996) among others, states that the subjective feeling we experience is generated as a result of our reacting to a pleasant or unpleasant stimuli. There is some evidence that it in fact is the *reaction* itself that elicits this subjective sensation (Schachter, 1964). We will not discuss the subjective aspects of emotion in this thesis, however.

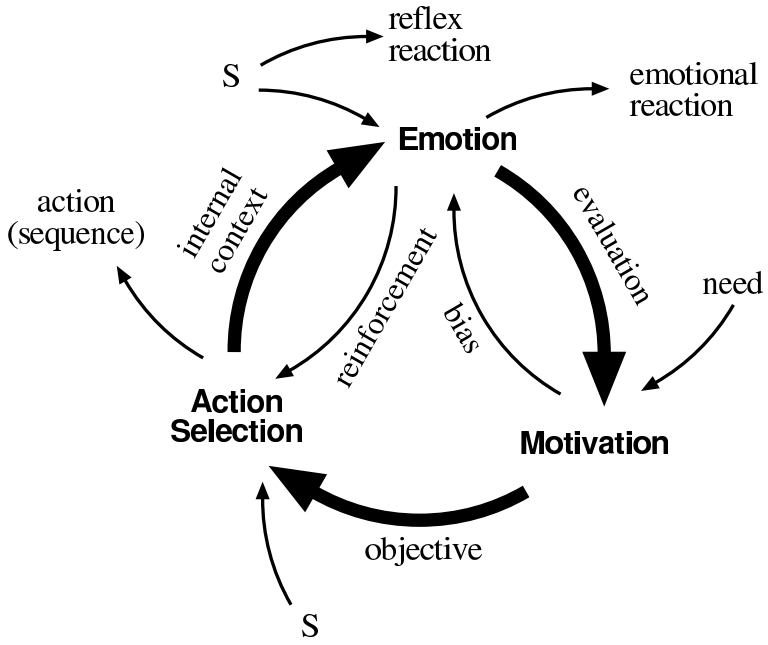

Figure 1.1: A model for generating actions based on internal needs and emotional evaluation of stimuli. This model has three interconnected main subsystems, incoming data from sensory subsystems and internal context, and outgoing paths leading to other subsystems (not described here).

1.2.1 A Model of Emotional Integration

We will not discuss details of the emotional system itself in this chapter; for such a discussion, see the next chapter, or see Balkenius and Morén (1998a); LeDoux and Fellous (1995); Rolls (1986). Instead we give a system-level description of the proposed architecture for an autonomous, emotionally driven system. This system-level model is not designed to be mappable directly on to the neurophysiological substrate in which these mechanisms are implemented. Several areas, and parts of areas, are part of the functions we describe here, and some areas would be mapped to more than one function. This model is rather meant to give an abstract framework to describe the role of emotion in the context of other high-level functions.

In figure 1.1 we have a principal view of the emotional subsystem as a part of a larger autonomous system (we have here omitted the subsystems dealing with the emotional aspects of attention and long term memory acquisition, concentrating instead on action selection). This description focuses on three interacting subsystems. The *motivational system* compares the internally gen-

erated needs with the emotional evaluations and generates *objectives* for action selection.

The *action selection system* uses the objectives to generate action sequences, changing the internal state of the system in the process. This system also receives information about outside stimuli and can create complex actions. If the objective is too abstract, the changes in internal context will enable the emotional and motivational subsystems to create more concrete objectives to accomplish the main objective.

Using the external stimuli and the external and internal contexts, the *emotional system* evaluates the stimuli for the motivational subsystem (Balkenius, 1995). This subsystem can also generate emotional reactions directly without involving the action selection system.

This model is a form of two-process model, as proposed by Mowrer (1973). One process, the emotional subsystem, learns an evaluation of stimuli thus forming an opinion of the desirability or undesirability of the stimulus. The motivational subsystem subsequently generates an objective (or general response) to deal with this stimulus, which the action subsystem carries out. Also, the emotional system generates a reinforcement signal directly to the action selection system.

1.2.2 Motivation

The first of the three components is the motivational subsystem. This system receives the internal needs and the emotional evaluation of the present stimuli and context. This evaluation influences the relative importance of the present needs and allows for the motivational system to generate an objective for the action system. The needs are internally generated, and correspond very roughly to Hebb's drives (Hebb, 1955).

The comparison of the internal needs and the emotional evaluation of stimuli enables the motivational subsystem to take into account the current situation when choosing an objective; a system that went into the kitchen because it was hungry might take a sip of water as well when close to a faucet, even though it was nowhere near as thirsty as it was hungry. This is an example of opportunistic behavior (Balkenius, 1993). On the other hand, when a need is completely fulfilled, the system would normally not react to a positive emotional evaluation unless it was very strong. When satiated, an animal would not eat a sandwich lying in front of it, but it might eat a small piece of candy.

We use the term 'objective', rather than 'goal'. This is mainly because 'goal' connotes a concrete, closed outcome, rather than a more abstract desire or state. An objective *can* be very concrete ("eat that sandwich") and will generate concrete actions, but can also be a high-level desire that in turn will determine one or several other objectives that eventually will fulfill this one.

The motivational system also outputs a bias signal back to the emotional system which is used in the evaluation. For example, food is only important if you are hungry. The emotional system is influenced to evaluate some stimuli higher, as the motivation forms part of the context in which it functions.

1.2.3 Action Selection

Action selection uses the motivational objective, the present stimuli and the context to generate actions to resolve the objective. These actions can be highly structured and context dependent; this subsystem is able to do a great deal of planning within the present context. The outcome of this system can be twofold. First, the system generates an action sequence if it is able to; these actions will of course in turn change the present stimuli and the externally generated context. If, however, the objective is too abstract or dependent on long-term memory, no explicit actions will be generated.

The system will also generate an internal context that in turn will influence both the emotional system and the action system itself. This internal context consists of needs, short term – or active – memory (including cognitive structures), bodily states and emotional state.

The action system will also get an evaluation directly from the emotional system. This evaluation is used as reinforcement when the action system learns to perform motor sequences.

1.2.4 Emotions

The emotional subsystem will make use of both external and internal stimuli and contexts when evaluating a stimulus. External stimuli are any features out in the world, and the external context is formed from these. Similarly, internal stimuli are internal bodily states, such as hunger and thirst, hormone levels, or activated memories.

The emotionally charged stimuli are used in several ways. First, as we have described here, the evaluation is used as *incentive motivation* by the motivational system to produce objectives. It is also used as a reinforcer for motor-learning in the action-selection system. Additionally, they are used for long term memory and attention, as we will see in the next section.

The emotional evaluation is a *reinforcer*. This means that the evaluation is able to increase or decrease the probability of something else happening. The emotional evaluation is in one of two forms: first order and higher order emotional evaluation.

First order – or primary – evaluation occurs with stimuli that are intrinsically emotionally charged (such as pain), but unexpected stimuli are also charged, as they can have a potentially large impact on the system. Many of these stimuli will by themselves generate an emotional reactions directly independently of action selection. Primary stimuli can also generate reflex reactions such as withdrawal even before they enter the emotional system.

With higher-order evaluation you have secondary stimuli that by themselves do not elicit a reaction, but acquires emotional content through association with a primary or another secondary stimulus. There is ample evidence that the learning being performed here can be described as classical conditioning (LeDoux, 1992).

The emotional system uses the current context to be able to rightly evaluate emotional stimuli, and the systems' needs are certainly a part of the internally generated context.

1.3 Learning and Attention

Mowrer (1973) established a two-process model whereby an emotional system evaluated stimuli and the evaluation then being used in the learning system proper. By not only advocating this role of emotions in learning, but also suggesting how such a system could be implemented, this work spawned a good deal of interest and the development of several new models based on this idea (see Gray (1975); Klopf (1988)).

One of the primary functions of emotions is the capability of evaluating stimuli. When a previously unknown or unremarked stimulus occurs in association with an emotionally charged stimulus, the emotional system will asso-

ciate this new stimulus with the same or a similar emotional content. Traditional learning methods all rely on some form of reinforcer, presumably generated from the outside. In reinforcement learning methods that allow for internally generated reinforcement, it is still very directly linked to the external reinforcer and the problem has been reduced to one of credit assignment (Kaelbling et al., 1996), as the reinforcers become directly linked with the specific actions taken by the system at the time.

In real life, of course, the situation is much more complicated; specifically, solving this as a credit assignment problem will not enable the system to transfer hard-won knowledge between contexts. Once a stimuli is evaluated by the emotional system, this evaluation can then be used as a basis both for evaluation of other stimuli and for evaluation of the contexts themselves.

The second function of the emotional system is to focus the system's attention where it would do the most good. The world is too complicated, and the sensory subsystems too variegated, for the system to be able to spend time and other resources on it all. By using the emotional systems' capability for evaluation and prioritizing, this sensory barrage can be sifted through so only the most relevant stimuli receives any attention. Very closely related to this is of course the need to decide what events to retain as long term memories, and we believe this is accomplished by the same mechanism.

1.4 The Thesis

This thesis is divided into six chapters, each dealing with a different aspect of the subject matter. Much of the content has been previously published as papers during the last few years and collected and reworked for the thesis. Though a good deal of material has been moved and reworked to make for a better reading experience, most of the chapters are still heavily based on one or two papers. The present chapter is based on Morén and Balkenius (2000b).

Chapter 2 will discuss the neurobiological foundations of the brain areas involved in emotional conditioning. This includes the amygdala, the orbitofrontal cortex and the hippocampus. The discussion will not be an exhaustive overview of these areas, but will be centered on those areas in relation to the functionality that is involved in emotional conditioning. Also, the perspective is computational and functional, rather than physiological.

Chapter 3 discusses the phenomenon of emotional learning from a psychological and experimental perspective. Although there are many forms of learn-

ing, the focus will be on classical and instrumental conditioning. We will discuss the basic mechanisms of classical conditioning, and look at some computational models implementing this functionality. We will also look at instrumental conditioning from a two-process perspective and look at some of the issues this learning mechanism. The section on conditioning models has previously been published as (Balkenius and Morén, 1998a).

Chapter 4 introduces our model of the amygdala, including a functional description, physiological mappings and results from simulations. The model is tested both in the presence and absence of a simple model of the orbitofrontal cortex. An early version has been published in (Balkenius and Morén, 1999; Morén and Balkenius, 2000a).

Chapter 5 will discuss the hippocampal model in a similar way as the amygdala model in the previous chapter. Both simulations in isolation and together with the amygdala model are presented. We will look at what capabilities the addition of a context processing ability gives the overall system. This chapter is based in part on (Balkenius and Morén, 2000b).

Chapter 6 discusses the model as a system-level implementation of a two-process model. We will see how it performs when parts of it are 'lesioned', or disabled, and we will compare the results obtained with the model to the other conditioning models discussed in chapter 3.

Chapter 2

Neurobiology of the Amygdala

It has recently been suggested that the association between a stimulus and its emotional consequences takes place in the brain in the amygdala (LeDoux, 1995; Rolls, 1995, 1999). In this region, highly analyzed stimulus representations in the sensory cortices, as well as coarsely categorized stimuli in the thalamus are associated with an emotional value. Evidence suggests that the process involved is classical conditioning (LeDoux, 1995; Rolls, 1995). The result of this learning is subsequently sent to other brain structures, including the hypothalamus, which produces the emotional reactions. Rolls (1986, 1995) has suggested that the role of the amygdala is to assign emotional value to each stimulus that is paired with a primary reinforcer.

There is little doubt that at least fear conditioning occurs in the amygdala; Fanselow and LeDoux (1999) reviews the data on this. Another review by Medina et al. (2002) shows that fear conditioning and eye blink conditioning occurs in different structures – the amygdala and the cerebellum, respectively. Also, Tsvetkov et al. (2002) show the expression of LTP (long-term potentiation) in the lateral amygdala during auditory fear conditioning.

In this chapter we will describe the neurobiological aspects of the amygdala, the orbitofrontal cortex, the hippocampus and other associated areas from a functional and computational perspective. The discussion will be fairly brief and focuses on the aspects of these areas that are relevant to the specifics of emotional conditioning. Also, this account is very much taken from a func-

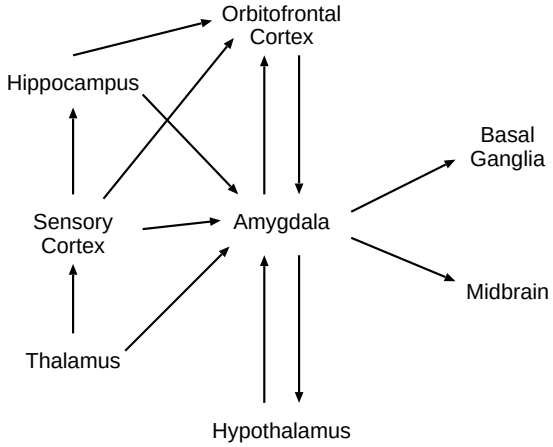

Figure 2.1: A schematic representation of the main areas and pathways connecting the amygdala to other areas.

tional and computational perspective, rather than from a neurophysiological one. As we draw the structure of our model (presented in chapter 4) mainly from the anatomical organization of these areas, the focus will be on this aspect.

We describe the main areas involved in emotional learning, and how they interrelate. This account is centered on the amygdala, as that area is the focus of attention for us in this text. As we can see in figure 2.1, there are quite a few other areas associated with this functionality, especially the thalamus, the hippocampus and the orbitofrontal cortex, and we describe those areas as part of the input and output structures projecting to and from the amygdala. We concentrate on the areas and features relevant to our conditioning model, rather than giving a complete description of all known areas and the connections between them.

In the last section, we take a look at the functional aspects of these interconnections. This will enable us to gain some understanding of what the system does, as well as how it does it.

2.1 Amygdala

Central to this thesis is the amygdala, where the primary affective conditioning occurs. This small, almond-shaped subcortical area is very well placed to receive stimuli from all the sensory cortices and other sensory areas. It is, to-

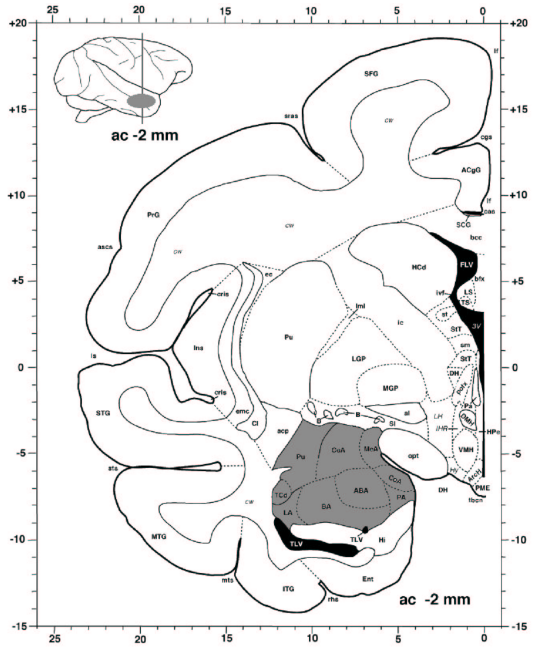

Figure 2.2: Location of the amygdala in the Macaque monkey. The greyed out area in the cross-sectional slice is the right amygdala. On the overview in upper left, the location of the slice is marked with a line, and the position of the amygdala in the brain as a whole is marked with the grey oval in the inset to the left. From (NeuroNames, 2002).

gether with the hippocampus, considered a part of the limbic system, which consists of various deep-lying areas in the cerebral cortex.

The amygdala – like most structures – is actually present in both hemispheres, with selective contralateral interconnections between them. There is some evidence (Öhman and Mineka, 2001) that the two structures respond to somewhat different stimuli. This is probably in part an effect of the fact that the two hemispheres have somewhat different functionality, and thus that the two structures receive different data to work with. In any case, the differences are not important at the level at which our model is working, and will be ignored from now on.

The amygdala consists of a number of distinct nuclei (figure 2.2). At least 5 main regions can be identified – the lateral, basal, accessory basal, central and the medial nuclei – and these can be further divided into subnuclei (Amaral et al., 1992; Pitkänen, 2000). In addition, there are several other areas that could be regarded as nuclei in themselves.

The lateral nucleus is the main input area for sensory information (Amaral et al., 1992; Pitkänen, 2000). From there information is spread to all the other nuclei of the amygdala (but note that other nuclei also receive substantial inputs from various other parts of the brain). The cortical signals enter the dorsal part of the lateral nucleus and continues to the ventral and medial parts. There are few or no backprojections to the dorsal area, or projections between the medial and ventral areas.

The structure of the lateral nucleus is topographical. The rostral part is the terminal for sensory stimuli from the somatosensory (touch), gustatory (taste) and visceral (intestinal) cortices. The caudal part receives its projections from the auditory and visual cortices (Pitkänen, 2000). Tsvetkov et al. (2002) shows that some aspects of auditory fear conditioning occurs in the lateral nucleus.

The two other deep nuclei are the basal nucleus and the accessory basal nucleus (Amaral et al., 1992). Both these structures receive inputs from the lateral nucleus and can be seen as intermediate processing stages. Note, however, that especially the basal nucleus also seems to be the primary output structure for control of higher-order conditioning (Whitelaw et al., 1996).

Finally the information reaches the central and the medial nuclei that serve as the main output region of the amygdala. On the surface of the amygdala lies the paralaminar nucleus and the periamygdaloid cortex. The latter is a cortical area for olfactory processing.

Although the lateral nucleus is mainly an input structure and the central and medial nuclei are output structures all nuclei receive both inputs from other parts of the brain and send outputs to them (Amaral et al., 1992; Pitkänen, 2000). These connections are described in the following subsections.

2.1.1 Amygdaloid Connections

The amygdala receives input from all levels of sensory processing. From the thalamus it receives early sensory signals that have not yet been highly analyzed , (LeDoux, 1995, 2000, p. 294). A more thorough analysis of a stimulus is done in the sensory cortex that also projects to the amygdala (Amaral et al., 1992; Rolls, 1995). Furthermore, the amygdala receives input from olfactory (McLean and Shipley, 1992) and gustatory areas as well as from the hippocampus (Amaral et al., 1992). Also, there are interconnections with the orbitofrontal cortex and the hypothalamus (Pitkänen, 2000).

It is useful to distinguish between three different types of input signals to the amygdala. The first is signals that code parts of the current sensory situation. What am I looking at? What am I hearing? Such signals are initially neutral but can acquire emotional properties though learning. The second type of input have innate significance. These carry information about the value of a stimulus: Is it appetitive or aversive? Can it be eaten? Does it present a threat? Is it a potential mate? The third type of input informs the amygdala of the current motivational state of the organism. Am I hungry, satiated, or sexually aroused?

There are three main sensory inputs to the amygdala that codes for the current situation at different levels of detail. These inputs originates in the thalamus and other subcortical areas, sensory cortex and prefrontal cortex. In addition, there are of course a number of reciprocal connections back to these areas, as well as to other areas such as the basal ganglia and the midbrain.

Thalamus

The thalamus is a subcortical structure that lies next to the basal ganglia. It is a part of the diencephalon together with the hypothalamus. The thalamus is not a homogeneous structure, but is composed of a number of smaller areas that seem to function somewhat independently. Its overall role seems largely to be a way-station between subcortical and cortical structures. Most sensory information (including somatosensory, auditory and visual information) is relayed from the peripheral sensory systems to the sensory cortices through various parts of the thalamus (Kelly, 1991). The thalamus also relays motor signals from the motor cortex. Interestingly, the olfactory system bypasses the thalamus altogether, and it has its own processing areas, largely separate from the rest of the amygdala, though there are some interconnections. There are thalamic sensory inputs to the amygdala, and as discussed by LeDoux (1996) and Öhman and Mineka (2001), these thalamic inputs probably mediate intrinsically emotionally charged stimuli as well as coarsely resolved stimuli in general.

The basal and especially the lateral nuclei of the amygdala are input structures that receive projections from the sensory cortical areas (Rolls, 1995; LeDoux, 1995). They receive connections from a large number of sensory structures in the brain, including the early sensory stages in the thalamus and the most complex sensory areas like inferior temporal area (IT) in the visual cortex, as described in the section on sensory cortices below. From the thalamus we find connections from the auditory analysis areas in the inferior colliculus through the medial geniculate nucleus (LeDoux, 1992; Weinberger, 1995). These con-

nections also terminate in the lateral amygdala. The role of these early connections may be to allow the amygdala to generate emotional responses with very short latency and prepare the organism for fight or flight (Gray, 1995; LeDoux, 1996). This initial reaction can subsequently be modulated by the higher sensory areas. Another possibility might be that these signals are used to prepare the emotional system to more efficiently process the detailed sensory data soon to come from the sensory areas. Either way, it is clear that emotionally significant information reaches the amygdala from lower structures and these are likely to be used as reward and punishment in the learning process.

Similar connections from the lateral geniculate nucleus through which visual information travels have not been reported. However, Morris et al. (1999) show that the pulvinar area of the thalamus is activated during backwards masked presentation of emotional faces. Desimone (1991) reports the presence of amygdaloid cells that respond selectively to faces in monkeys.

The importance of these low-level inputs to the amygdala has been disputed. For example, Rolls (1995, 2000) states that the earlier stages of sensory processing only plays a minor role in the activation of the amygdala. On the other hand, LeDoux (1995, 2000) assigns an important role to the signals from the auditory thalamus. One explanation for the divergent conclusions is that the animals used as objects of study are different; Rolls works with macaque monkeys, while LeDoux works with rats. It is not inconceivable that these early connections really are more important in rodents than in primates. Also, LeDoux primarily works with auditory stimuli whereas Rolls works with vision, and the connections between MGM and the amygdala is a far better documented pathway between thalamic sensory areas and the amygdala.

There are also connections from the ventroposterior medial nucleus of the thalamus that contains fibers that carry gustatory and visceral information (Amaral et al., 1992). This may be an early route through which the amygdala can learn about the consequences of ingesting a certain food substance. These may function as primary reward and punishment in the learning process in the amygdala. Information from the somatosensory pain system is likely to enter at this level also (Davis, 1992). Shi and Davis (1999) show that somatosensory pain enters the lateral amygdala from the posterior intralaminar nuclei of the thalamus.

Hypothalamus

The hypothalamus lies below the thalamus, and seems to be connected to various functions that regulate the endocrine system (especially the pituitary gland), the autonomous nervous system, as well as primary behavioral survival functions such as hunger, thirst and sex drive; see (Schachter, 1970) for an engaging review of obesity in hypothalamically lesioned rats.

There are connections from the medial, central and anterior cortical nucleus of the amygdala to the lateral hypothalamus. In addition, the medial and anterior cortical nuclei project to the anterior hypothalamus, and the medial and accessory basal nuclei project to the ventromedial hypothalamus (Pitkänen, 2000). These are thought to be involved in motivational control of the structures in the hypothalamus (Rosenzweig and Leiman, 1982; Thompson, 1980).

Some parts of the hypothalamus connected to the amygdala are involved with the control of eating. For example, the medial nucleus of the amygdala appears to inhibit ventromedial hypothalamus which in turn controls satiety. The effect is to stimulate eating behavior. The basal lateral amygdala, on the other hand, inhibits lateral hypothalamus and excites ventromedial hypothalamus and thus has an inhibitory influence on eating behavior (Rosenzweig and Leiman, 1982).

There are also projections from the hypothalamus to various parts of the amygdala. There are light projections to the basal and accessory basal nuclei from a number of hypothalamic nuclei. The lateral, ventral and ventromedial hypothalamus project back to the central nucleus of the amygdala, and the ventromedial nucleus projects to the lateral nucleus of the amygdala. The most diverse inputs from the hypothalamus to the amygdala terminates in the medial amygdala. There are a large number of heavy projections from many hypothalamic areas to the medial nucleus, and the projections from the medial nucleus of the amygdala back to the hypothalamus are as heavy and diverse.(Pitkänen, 2000).

Hippocampus

The hippocampus is a twisting, vaguely horseshoe-shaped structure in the same subcortical region as the amygdala. The structure is quite complex, with a three-dimensional organization that makes illustration difficult. The hippocampus is seen as critical for many functions, including (but not limited to) spatial navigation, the laying down of long-term memory and the forma-

tion of contextual representations. All of these roles have been assigned to the hippocampus in different theories and models. The perhaps most influential theory of the hippocampus is the cognitive map theory of O'Keefe and Nadel (1978). They suggested that the hippocampus is responsible for the mapping of the environment mainly based on environmental cues.

Other suggestions include the hippocampus as a memory for sequences or events (Solomon, 1979; Rawlins, 1985; Olton, 1986), working memory (Olton and Samuelson, 1976) or configurational codes (Solomon, 1980). It has also been suggested that the representation of a location of a stimulus and the stimulus itself that are segregated in neocortex are bound together in memory by the hippocampus (Mishkin et al., 1983).

The idea of hippocampus as a place-encoded memory system is supported by data that shows that hippocampal cells in the rat react to the place the animal is situated at (O'Keefe, 1990), but is complicated somewhat by indications that these cells in primates (that are more visually oriented) react to places the animal *observes*, rather than the place the animal itself is situated at (Rolls and Treves, 1998, p. 100). As Rolls discusses, however, this is not a problem for this view in practice. Rather than seeing it as place cells for the animal, they code for place of the objects around it. The rat, being rather more reliant on smell and touch than on vision, would generate data that would superficially look as if the cells are reacting to the animals position in space, rather than on the objects in its vicinity.

Another function associated with the hippocampal system is the comparison between stored regularities and actual stimuli (Gray, 1995). The role of the hippocampus in contextual control of memory and learning is also well known (Hall and Pearce, 1979).

Rolls and Treves (1998) suggest that the hippocampus is organizing disparate sensory information into one episodic instance. This can be used in various ways; Rolls and Treves discuss mainly its use to consolidate long-term memory and as cues for action. In this text, we consider this to be a contextual representation (or code).

Rudy and O'Reilly (1999) explores the role of the hippocampus in contextual fear conditioning, where the animal learns to associate an unpleasant experience with the surroundings in which it happened. They show that the hippocampus is primarily responsible for this function, though some effect remains even with a lesioned hippocampus, probably through 'normal' feature-based association. They also discuss the effects of preeexposure to the context for this effect. Rats that receive an unpleasant stimulus (foot-shock) immediately upon being placed in the environment show markedly diminished con-

textual fear, whereas rats that have had the opportunity to explore their surroundings have an increased effect. The interpretation is that the hippocampus encodes a stimulus configuration to be associated with the unpleasant stimulus. In the first case, the rat never has time to sample their environment enough to create a configuration, whereas in the second case, a stable configuration able to be activated by just a small number of features is created.

There are three major areas of the hippocampus: the dentate gyrus, CA1 and CA3. Closely associated with the hippocampus is the subiculum and the entorhinal cortex with the parasubiculum. The main pathway through the hippocampus is from the entorhinal cortex, to the dentate gyrus, continuing to the CA3, then to CA1 and out through the subiculum (Rolls and Treves, 1998).

Rolls and Treves (1998) view the CA3 area as an autoassociator, capable of associating disparate sensory and other information with each other (the effect is to be able to recall the entire set when presented with only a subset of it). The function of the dentate gyrus seems to be to orthogonalize the input for the CA3 associations to work efficiently. CA1 is organizing the output set from CA3 into contextual units (Rolls and Treves, 1998).

Hippocampal lesions are well known to produce anterograde amnesia, or inability to form new long-term memories. However, this does not impact skill learning or short-term memory (Kupfermann, 1991).

The amygdala is heavily interconnected with the hippocampus. There are moderate to dense connections from the subiculum of the hippocampal formation to all major areas of the amygdala (Pitkänen, 2000). The CA1 area of the hippocampus also projects to the lateral, basal, accessory basal, medial and central amygdala. There are reciprocal projections back to the subiculum and CA1 from all those areas except the central amygdala (which does not project to the hippocampus at all). The lateral, basal, accessory basal and medial amygdala also projects heavily to the parasubiculum. The basal amygdala projects heavily to CA3 as well.

The subiculum is a source of multimodal inputs to the amygdala (LeDoux, 1995) which is probably involved in the representation of stimuli over time intervals larger than 250-300 ms after their termination (Clark and Squire, 1998). It is likely that these connections also mediate representations of the temporal and spatial context in which emotional learning occurs. Bonardi (2001) shows that hippocampus-lesioned rats show learning deficits in classical conditioning, but only when the cue is localized (ie. a light inside the food tray). When the cue is unlocalized (a general increase in illumination level), there is no impairment as compared to the controls.

Sensory Cortex

The sensory cortex receives its input through the thalamus. These areas are responsible for much of the higher perceptual processing for the animal. They receive sensory information from the outlying sensory areas through the thalamus and then process this information very extensively for various purposes. The function of these areas is an extensive subject all by itself, and no effort will be made here to describe this in any detail.

The amygdala receives highly analyzed input from all the sensory cortices. These signals enter the amygdala in the lateral and basal nuclei (Amaral et al., 1992; Rolls, 1995; LeDoux, 1995; Pitkänen, 2000). The visual input includes signals from the inferior temporal cortex (IT) with the highest level of visual analysis (Rolls, 1995). Cells have been found in the IT that react to complex visual stimuli such as objects and faces (Perrett et al., 1992; Desimone et al., 1984). The role of these connections in this system appears to be to supply the amygdala with highly analyzed signals that can be given emotional significance.

The cells in the inferior temporal cortex that react to faces are especially interesting. Some of these cells react to specific persons regardless of the orientation of the face while other cells react to any face given that it has a specific orientation in space or a certain facial expression (Perrett et al., 1992; Desimone et al., 1984). These representations are probably important for assigning emotional value both to specific persons and to emotional expressions and gestures. The accessory basal amygdaloid nucleus also contains cells that react to the presentation of faces (Leonard et al., 1985). It is likely that these cells receive input from the regions of the inferior temporal cortex that react to faces and facial expressions. Consequently, it has been reported that lesions of the amygdala causes deficiencies in social behavior (Kling and Steklis, 1976). Animals with lesions in the amygdala are no longer able to interact with the other members of their group.

The auditory cortex is also well interconnected with the lateral amygdala. There are extensive interconnections from temporal cortex area 2 and 3 in to the caudal part of the lateral amygdala (LeDoux, 1987; Weinberger, 1995; Pitkänen, 2000). Connections from the auditory regions of the superior temporal area have also been reported (Amaral et al., 1992).

In addition to the inputs from the monomodal sensory regions, the amygdala also receives multimodal inputs from the entorhinal cortex (Gray et al., 1981; Amaral et al., 1992). In this respect, the amygdala is similar to the hippocampus which also receives massive projections from this area. These connections seem to be used for sensory integration.

The amygdala also transmits information back to the sensory cortices (Rolls, 1989a; LeDoux, 1995; Weinberger, 1998). There are two kinds of outputs to the sensory cortices. The first type of outputs is likely to be used for priming sensory stimuli as part of an attentional system (Rolls, 1999). The second type is likely to be a part of a system of memory consolidation, where the emotional evaluation triggers the storing of long-term memories in, for example, the visual cortex (Tabert et al., 2001; Cahill and McGaugh, 1998). We will describe this further below.

Orbitofrontal Cortex

Fuster (1997) sees three interrelated functions for the prefrontal cortex: working memory, preparatory set and inhibitory control. His concept of working memory is a representation of current events and actions, as well as such events in the recent past; this is not unlike the concept of context. A preparatory set is the priming of other structures in anticipation of impending action. He also calls this *motor attention*. Inhibitory control is the selective suppression of areas that may be inappropriate in the current situation. It appears that the amygdala is involved in the initial learning of an emotional response while the orbitofrontal cortex is necessary for extinction (Rolls, 1995).

An interesting view of the frontal cortex is that its role is to inhibit the more posterior structures to which it connects (Shimamura, 1995; Fuster, 1997). According to this view, the difference between the various frontal regions comes primarily from what structures they inhibit. Taking this perspective on the orbitofrontal cortex suggests that it inhibits earlier established connections when they are no longer appropriate, either because the context or the reward contingency has changed (Rolls, 1986, 1990, 1995). It has been argued that extinction is controlled by the inhibition from this area (Rolls, 1995; Balkenius and Morén, 2000a). Similarly, habituation can be seen as the active process of inhibiting the orienting reaction to stimuli that are of no value to the animal (Gray, 1975; Balkenius, 2000). The prefrontal cortex has also been implicated in this process (Fuster, 1997). It is likely that the frontal cortex receives information about the current context from the hippocampus. Working together, the hippocampus and prefrontal cortex could be responsible for the inhibition that occurs in habituation and extinction (Rolls, 1995; Fuster, 1997).

The orbitofrontal cortex appears to be especially involved in this function. This can be seen when reinforcement contingencies are changed. Rolls (1995, 2000) suggests that the orbitofrontal cortex reacts to omission of expected reward or punishment and controls extinction of the learning in the amygdala. This extinction is suggested to be the result of an inhibitory influence from the

orbitofrontal region. Cells have been found in the orbitofrontal cortex that are sensitive to sensory stimulation and that code for specific stimuli (Rolls, 1992). This makes it reasonable to consider this a sensory area. The reaction of these cells are more complex than those in the earlier sensory cortices, since they also reflect the history of reinforcement that the stimulus has encountered. These cells have also been found to reverse their activity when reinforcement is changed (Rolls, 1995).

Apart from inhibitory control, the prefrontal cortex has also been suggested to take part in short-term working memory and preparatory set (Fuster, 1997). For emotional processing, these aspects of the prefrontal system are somewhat different from its motor functions. Apart from the orbital regions, the dorsolateral and ventromedial areas are also believed to be involved in emotional processing (Davidson and Irwin, 1999). Patients with ventromedial damage are impaired in the anticipation of future reward or punishment but are still influenced by immediate consequences of their actions. The dorsolateral prefrontal cortex appears to be involved in working memory. Damage in this area makes patients unable to sustain emotional reactions over longer times (Davidson and Irwin, 1999).

Lesions of the frontal cortex result in an inability to change behavior that is no longer appropriate (Shimamura, 1995; Kolb and Whishaw, 1990). For example, in the Wisconsin card-sorting test, subjects are asked to first figure out how to sort cards according to a simple criteria such as color. When the subjects have succeeded, the criteria is changed and the subjects have to find the new rule to sort the cards. Frontally impaired patients are often unable to do this. They may be able to verbalize that the rules have changed but they will persevere in their incorrect behavior.

Inputs to the amygdala from the prefrontal cortex enter the amygdala in the lateral, basal and accessory basal nuclei (Pitkänen, 2000; Rolls, 1995; LeDoux, 1996). The amygdala also projects to prefrontal cortex (Rolls, 1995; Schoenbaum et al., 1998).

Other Areas

The amygdala projects to a number of subcortical areas that affect autonomous functions like heart rate, hormone levels and other autonomous reactions. The central amygdala is the primary structure to project to these areas. It also projects to the periaqueductal gray and tegmentum in the midbrain (Pitkänen, 2000), as well as to the autonomic areas of the medulla oblongata (Rolls, 1995). The purpose of this projection is likely to prepare the body for swift action if required.

The basal ganglia is a collection of subcortical areas that mediate voluntary and planned movement. It consists of the striatum (composed of the caudate and putamen) that seems involved in the instrumental conditioning of motor actions; the globus pallidus and substantia nigra (that seem to be the main output areas back to the thalamus); and the subthalamic nucleus. Both substantia nigra and globus pallidus are divided into two areas. The internal segment of globus pallidus and substantia nigra pars reticulata share many similarities with each other. In fact, they are more similar to each other than with the other part of each structure; they can functionally be considered one area, and is the output area for the basal ganglia (Côté and Crutcher, 1991).

Like the amygdala, the basal ganglia is modulated by the prefrontal cortex (Fuster, 1997). This system in the brain may be involved with the learning of response reward associations (Houk et al., 1995) and the inhibition from frontal cortex could be used to select among competing motor programs.

There are connections from substantia nigra to the lateral, basal and central areas of the amygdala. There is a strong connection from the central nucleus to substantia nigra. Recently it has been found that neurons in substantia nigra (and ventral tegmentum in the midbrain) react to mismatches between expected and received reward (Waelti et al., 2001).

Other low level inputs comes from all the olfactory cortical areas including the periamygdaloid cortex (Amaral et al., 1992). The amygdala also receives direct input from the accessory olfactory bulb which carry information about pheromones from the vomeronasal organ (McLean and Shipley, 1992).

As we have seen, the amygdala receives sensory information at a number of levels of analysis. Each higher level can correct the emotional learning that has taken place using information from the earlier stages. The multimodal convergence at the amygdala could be responsible for the association between a neutral stimulus with an innate evaluation based on for example somatosensory, gustatory or visceral information. Additionally, the hypothalamus contributes with information about the current motivational state of the organism. The hippocampus, directly, and through the orbitofrontal cortex, controls the context-dependent aspects of this processing and its expression.

2.1.2 Lesions of the Amygdala

Lesions of the amygdala produce striking effects on behavior (Weiskrantz, 1956). Monkeys with amygdaloid lesions show a marked lack of fear. They may play with objects, such as snakes, that would otherwise frighten them.

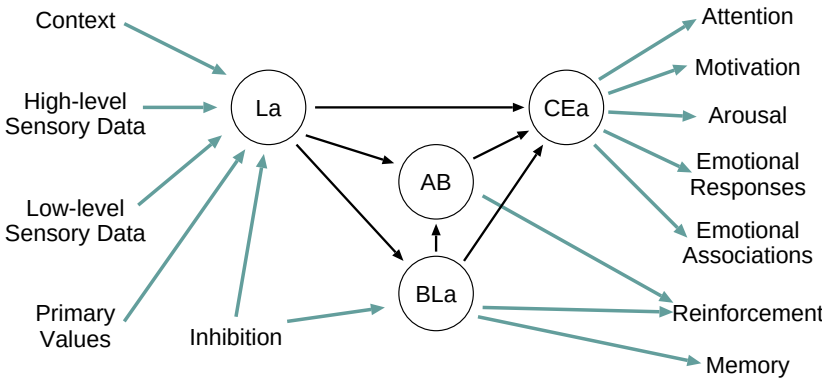

Figure 2.3: A schematic illustration of the amygdaloid areas we are concerned with, and their functional connections. La: lateral amygdala; BLa: basal amygdala; AB: accessory basal amygdala; CEa: central amygdala.

They also increase their oral behavior and have learning problems. Other problems are loss of social dominance, inappropriate social behavior, change in social and sexual preferences, together with less facial expressions and vocalization (Kolb and Whishaw, 1990).

Human lesions of the amygdala appear to contribute to a large portion of the so called Klüver-Bucy Syndrome which may result from damage to the temporal cortex (Klüver and Bucy, 1939). This syndrome consists of tameness, loss of fear, indiscriminate dietary behavior, increased sexual behavior, often with inappropriate object choice, hypermetamorphosis, a tendency to examine all objects with the mouth and visual agnosia (Kolb and Whishaw, 1990). The last effect is probably due to damage to the inferior temporal gyrus close to the amygdala. Also, patients with damage to the amygdala are reported to be unable to recognize facial expressions of fear, while the ability to recognize other emotional expressions were intact (Adolphs et al., 1996; Broks et al., 1998).

In animals, similar damage have resulted in loss of social dominance, inappropriate social behavior, change in social and sexual preferences, less facial expressions and vocalization (Kolb and Whishaw, 1990).

2.2 A System Level Description

Now we take a look at how these areas seem to interact at a system level. The focus is of course still on the amygdala, and we will present this data in

relation to its functioning as an emotional system. As figure 2.2 shows, there are a number of functional elements connected to the amygdala. Rather than just iterate over all the areas that are relevant, we will discuss them in relation to each of these functional elements.

2.2.1 Attention

One function of the emotional system is to contribute to attention. When something emotionally significant is detected, it would often be beneficial for the animal to attend to this stimulus. There exist projections from the central amygdala to the sensory cortices that may be involved in priming of sensory stimuli with the current emotional state.

There are also projections to the sensory cortices that may be involved in the emotional control of sensory categorization and motivation (Rolls, 1989a). This includes both the facilitation of memory creation in emotional situations and ability to bias or prime cortical processing with the current emotional state (Amaral et al., 1992; LeDoux, 1996). This type of projection is especially salient in the visual system where the amygdala connects to all levels of visual processing (Rolls and Treves, 1998). This should be contrasted with the projections to the amygdala that mainly involve the inferiotemporal area with the highest level of visual analysis. Through the backprojections to sensory cortex the amygdala could potentially activate emotional memories or direct attention to stimuli that are relevant to the current emotional and motivational state (Rolls, 1992; Holland and Gallagher, 1999).

Holland and Gallagher (1999) shows that the central amygdala mediates conditioned orienting responses. However, lesions of that area does not affect unconditioned orienting behaviors, or habituation. Neither do the lesions disrupt the normal responses to affective stimuli. Holland and Gallagher identified the connection between central amygdala and substantia nigra in the basal ganglia as the probable route for conditioned orienting responses, as disconnection of this route shows similar deficits in this area as lesioning of the central amygdala itself.

2.2.2 Motivation

The outputs from the amygdala to the hypothalamus are thought to be involved in motivational control of the structures in the hypothalamus (Rosenzweig and Leiman, 1982; Thompson, 1980). For example, the cortical medial

nucleus of the amygdala appear to inhibit ventromedial hypothalamus which in turn controls satiety. The effect is to stimulate eating behavior. The basal nucleus, on the other hand, inhibits lateral hypothalamus and excites ventromedial hypothalamus and thus has an inhibitory influence on eating behavior. In this way, the amygdala is able to control motivation according to the emotional value of the attended stimulus; a piece of food associated with fear or illness will cause the animal to lose its interest in food when observing it.

2.2.3 Arousal and Reflexive Responses

Through lateral hypothalamus, the amygdala is able to control autonomic fear and anxiety responses and through the paraventricular nucleus of the hypothalamus it can control the secretion of stress hormones (Davis, 2000). The central nucleus can also influence the startle reflex controlled by reticulopontis caudalis (Davis, 2000).

One set of outputs from the central nucleus is directed toward the autonomic areas of the medulla oblongata (Rolls, 1995; Davis, 1992), including the ventrolateral medulla (Petrov et al., 1996). This output is responsible for the somatic affects that usually accompany emotional states. More specifically, the ventrolateral medulla (together with nucleus of the solitary tract) seems to regulate blood pressure as an autonomic response to external stimuli. The central nucleus also projects to the periaqueductal gray that controls the freezing reaction that is a common reaction to danger (Fanselow, 1994).

As Gallagher (2000) reviews, the influence of the central amygdala on autonomic responses is limited to emotionally affective stimuli. Lesions of the central amygdala would inhibit the conditioned expression of the startle reflex, but would not inhibit the 'normal' startle reflex in response to a novel stimulus.

The medial amygdala connects to the ventromedial nucleus of the hypothalamus, which initiates some aspects of mating behavior through the periaqueductal gray (Rolls, 1999, p. 221).

2.2.4 Memory, Learning and Reinforcement

There is a lot of evidence that the amygdala plays a role in the formation of memories in other areas of the brain, in addition to the conditioning taking place in the amygdala itself (Fanselow and LeDoux, 1999). Packard and

Teather (1998) shows that the amygdala influences both hippocampal and striatal (specifically the caudate and putamen) memory expression. As their experiments show, the hippocampus is implicated in the short term memory retention of spatial locations, while the striatum is involved in the learning of reactions to a present stimulus (i.e. instrumental conditioning). The amygdala is found to exert an influence in both memory systems. Poremba and Gabriel (1999) show that the amygdala is involved with the formation of instrumental responses, but not with its maintenance. This is consistent with the view of the amygdala as an evaluative part of a two-process system of instrumental learning (see chapter 3).

Another type of backprojections pass through the basal forebrain and may be involved in the formation of emotional memories by enhancing the learning in emotional situations (Weinberger, 1995, 1998). Experiments have shown that the formation of sensory categories in auditory cortex can be controlled by the amygdala (Weinberger, 1995). The emotional evaluation of the amygdala is also sent to the prefrontal cortex (Rolls, 1995) and to the basal ganglia (Gray, 1995). These outputs originates in the basal and accessory basal nuclei. The influence to striatum is mediated mainly from the basal nucleus in the amygdala (Everitt et al., 2000). Lesions of the basal nucleus, while impairing conditioned reinforcement, does not impair the formation of emotional associations. On the other hand, lesions of the central amygdala CEa disrupts Pavlovian conditioning, but not conditioned reinforcement.

One result of emotional reactions is the release of stress hormones (such as cortisol and adrenaline (LeDoux, 1995). These hormones are known to modulate memory consolidation. Roozendaal et al. (1998) confirms that the basolateral amygdala is implicated in the modulation of adrenaline regulation of memory consolidation.

Cahill and McGaugh (1998) reviews the role of the amygdala in memory modulation. They conclude that the basolateral, but not the central, amygdala is involved in modulation of emotionally influenced memories. They also review results that show that the hippocampus, but not the caudate nucleus, modulate the retention of spatial tasks; that the caudate nucleus, but not the hippocampus, modulate retention of cued tasks; and that the amygdala modulates both. This is entirely in line with the other results reviewed here.

Whitelaw et al. (1996), showed that lesioning the basolateral amygdala in cocaine-seeking rats impaired their ability to acquire secondary conditioning, but not simple conditioning, indicating the importance of this area in higher-order conditioning. Hitchcott and Philips (1998) showed this functional dissociation between the basolateral and central amygdala as well, with the central amygdala involved in Pavlovian approach behavior and the basolateral amygdala involved in instrumental reward-associations.

The reinforcer itself seems to enter this system in the same manner as other stimuli, at least in the case of conditioning to pain. Shi and Davis (1999) performed a set of lesioning studies on rats to find the pathways responsible. Their results show that painful somatosensory stimuli capable of inducing fear conditioning enters the lateral amygdala both from the posterior intralaminar nuclei in the thalamus, and from the insular cortex. The thalamic area receives somatic inputs from the spinal cord. The insular cortex receives inputs from the somatosensory areas, the thalamus and from the midbrain (Craig et al., 1994). Schultz et al. (1998) and Schultz (1998) reviews data showing that primary, undirected reward signals are produced by dopamine neurons in the midbrain, and are projected to many other areas of the brain. These react to mismatched predictions for primary rewards, rather than the levels themselves, and do not carry any specificity as to the nature of the reward.

Another interesting set of inputs comes from taste and olfaction areas (Rolls, 1995, 1989b). These may function as appetitive reinforcers in the learning process in the amygdala. Rolls (1999, pp. 31-37) shows that activity in the secondary taste areas in the orbitofrontal cortex are intrinsically rewarding.

Chapter 3

Learning

In this chapter, we will briefly discuss learning in general, but will concentrate on classical and instrumental conditioning. We will also discuss some computational learning models within this framework. We will not, however, discuss reinforcement learning or machine learning in any detail, as these fields are not directly relevant to the area of emotional conditioning.

The term 'learning' has many different connotations, ranging from a pedagogical process, through the large field of machine learning, to long term potentiation (LTP) in individual nerve cells. The term is so overloaded, however, that it is necessary to briefly discuss its different uses and qualify the term in the context of this work.

At the lowest level, there is adaptation of single cells (typically through LTP). Although the details are complex, the basic idea is simple. Given two cells, one connected to the other, the first cell will occasionally fire, thus inducing the second cell to fire, or suppress the second cell from firing. However, when this happens in the presence of specific chemical signals, the synapse – or connection – between the cells will alter so that a signal from the first cell now can trigger the second cell more easily. This is a learning mechanism first suggested by Hebb (1955), and subsequently confirmed as a mechanism in biological learning processes (Bliss and Lomo, 1973).

Machine learning, on the other hand, is not concerned with biological processes. The term is an umbrella for various techniques for making systems capable of behavior adapted by external events. This includes tasks such as

path finding, hill climbing, evolutionary learning and other forms of optimizing search (Kaelbling et al., 1996; Sutton and Barto, 1998). Although these methods sometimes borrow ideas from biology or psychology - and these sometimes get inspiration from machine learning - the focus is very different. Whereas biological learning concentrates on how actual organisms perform their tasks, machine learning research aims to produce methods that work in artificial systems, regardless of their biological relevance. This should not be taken as a criticism of either field, but the difference in motivation should be kept in mind so we do not prejudge results from either field erroneously.

Evolutionary learning is the idea that the process of evolution can be seen as a learning task stretched out over populations rather than individuals, and evolutionary time rather than seconds or days. Some researchers go as far as proposing that the mind indeed works as an evolutionary system where groups of neurons compete for the ability to respond, with increased connectivity as the prize (Edelman, 1989).

Learning as pedagogics is something with which all readers will be familiar. It is the science and craft of human to human, assisted, learning, especially during adolescence. While obviously important to us, the field has fairly little relevance to biological learning for several reasons. First, it is applicable almost exclusively to humans (with the possible exception of a few other primates). It is dependent on language and of imitation, something the vast majority of organisms are not capable of. Second, it deals largely (though not exclusively) with explicitly formulated linguistic knowledge, again something not applied to most animals. While this field probably can apply results from other fields of learning, it is not clear that this field can contribute back to biological learning in a meaningful way.

We thus come to the current field of interest: biological learning. This is the study of learning processes common to most animals (including humans), sometimes with a neurological perspective, sometimes with an ethological one. For animals, learning is a crucially important survival feature. While a lot of behaviors are innate, these behaviors are developed over evolutionary time, and can not by themselves assure good survival in a fast-paced, killer-eats-it-all world. As a result, for many animals the fixed behaviors are independent of outside influence (things like breathing or circulation), or fixed behavior patterns triggered by very specific outside circumstances, e.g. Greylag geese egg retrieval behavior (Lorenz and Tinbergen, 1938). At a slightly greater level of abstraction, they can be semi-general behavior fragments that can be altered, tuned and put together in various ways depending on situation and experience.

Learning is, even in this context, not a very well defined concept. It occurs when a behavior gets preferentially selected in the face of circumstance, but it can also occur in the form of imprinting, in the laying down of long-term memories and in the adaptation of motor behaviors to a changing body (this would include both the change of body size and proportions from childhood to adulthood, as well as injuries or disabilities). All of these are cases of adaptation to circumstance, though not everybody would call these processes learning.

As an example of how fixed behavior fragments and lifetime adaptation interacts, we can take a look at Clark's nutcracker (Gallistel, 1990, p. 155). This bird collects and hides caches of pine seeds in the fall to eat during winter and spring. The bird will make up to 33,000 seed caches, and must be able to recover 2,500 to 3,000 of them to avoid starvation.

While the behavior of hiding seeds for winter is instinctive and developed in evolutionary time, remembering exactly where the seed caches are is not. All evidence reviewed by Gallistel points to the conclusion that it remembers the location of each seed cache. The interaction goes deeper than that, however. The bird no doubt has innate preferences for the kind of places that are good for hiding seeds. It hides them by picking its beak into the ground, depositing the seeds, then covering up the traces, so it will try to hide the seeds in soft ground, rather than in tree hollows, under rocks or in high branches. While the precise spot must be remembered by the bird, the kind of places that it could have hidden the seeds in is restricted to those that its instincts tell it could be a decent burial place.

Among the processes being studied, none have been more so than that of classical and instrumental conditioning.

3.1 Conditioning

Conditioning was first studied in a systematic way around the beginning of the twentieth century. Although the effect had been observed earlier (and, indeed, animal trainers and others made empirical use of similar techniques), the first to systematically study conditioning was Ivan Pavlov (Pavlov, 1927). Pavlov was already established as a physiologist studying the gastro-intestinal system - and in fact received the Nobel price for this work in medicine in 1904.

In an effort to study the gastric response to food, he started to study dogs outfitted with a valve to the stomach to see how the body would regulate

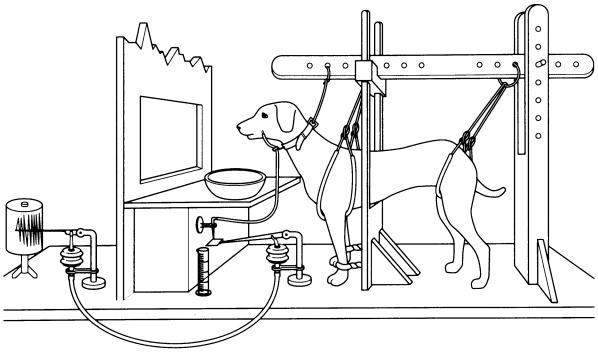

Figure 3.1: The setup used by Pavlov to measure physiological data on the gastrointestinal system. In this illustration, the setup measures the production of saliva, rather than gastric acid. (From everywhere. Used without permission.)

production of gastric acid (he also had dogs with tubes through the cheek to study the flow of saliva). The acid would drip down onto the apparatus when it saw the food, the drops would be counted, and a diagram would be produced over the rate of fluid production.

As the trials progressed, he observed that the dogs would start salivating *before* they saw the food. It was eventually discovered that the dogs reacted to the sound of the assistant that normally fed the dogs. Pavlov recognized the effect as being a new phenomenon. He gradually eliminated all extraneous stimuli which could affect the learning process, and managed to simplify the experiments to the point of being able to present a single stimulus which the animals could react to exclusively. He called this process conditioning.

Conditioning is the process of influencing behavior using rewards or punishments, generally in a psychological or neurological setting. The canonical example of this process is feeding a dog every day, and ringing a bell just before giving it the food. After a few days, the dog will start salivating at the sound of the bell, even in the absence of any food.

3.1.1 Behaviorism

The idea of conditioning as the mechanism behind all learning quickly became influential, especially in the USA, where it formed part of the basis of behaviorism (Watson, 1930; Skinner, 1974). This school was formed as a re-

action to the mentalistic approach to psychology that relied in large parts on introspection as a means of studying the human mind.

The behaviorists rejected the mentalistic approach as being unscientific. They wanted to ground psychology in the same kind of logical positivist framework as the natural sciences, where only observable phenomena were available for study. As the only observable variable for humans is their behavior, psychology should therefore study behavior only.

Pavlov's ideas about conditioning fitted perfectly into this framework. Armed with this new tool, a generation of psychologists set out to measure anything and everything that this tool afforded them. Unfortunately, the rejection of the inner mind as amenable to study soon hardened into a rejection of the internal mind itself. Behavior, according to behaviorists, is all there is. Watson (1930) went as far as saying that

> "[The behaviorist] dropped from his scientific vocabulary all subjective terms such as sensation, perception, image, desire, purpose, and even thinking and emotion as they were subjectively defined." (p. 5)

This conclusion was not received with enthusiastic praise among all psychologists. Unfortunately, when rejecting behaviorism, the idea of conditioning was often found guilty by association. Thus, for a long period, the only work in this field was done only by behaviorists.

3.2 Classical Conditioning

What Pavlov had discovered is the phenomenon known today as classical - or Pavlovian - conditioning (the term 'conditioning' actually stems from a mistranslation of Pavlov's book into English). The principle behind classical conditioning is very simple: present an unconditioned stimulus (commonly shortened as US) - the food in Pavlov's case. This will cause an unconditioned response (called UR) in the subject - salivating, for Pavlov's dogs. We want to condition the subject (that can be human as well as other animals) to make a response to some other stimulus that does not by itself have anything to do with the response. (A fuller discussion on the nature of an US is found on page 78.)

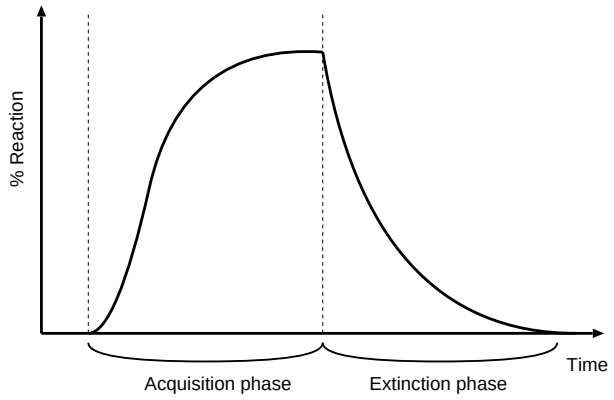

Figure 3.2: An idealized Acquisition-Extinction curve. When the animal first makes the connection between the CS and the US, the degree of response rises rapidly, and flattens out as it nears its ideal degree of response. When the connection disappears, the response decreases rapidly, but lingers at a low level for some time.

To accomplish this, we introduce a conditioned stimulus (CS) before or at the same time as the US (the food). In Pavlov's case it was the sound or sight of the assistant, but it is traditionally a simple sound or a light. The US (food) will cause the UR (salivating) to occur, but at the same time, the CS (the tone, bell, light or other stimulus) will itself become associated with the US. After a number of trials, we start giving the CS *without* the US. We will find that the CS will trigger a response just as if the US had been present; in other words, the bell will trigger salivation just as the food did. Also, if we continue to present the CS without the US, the response (called CR, to differentiate from the response given with only the US) will gradually diminish and disappear. An ideal graph of the response at the time of CS is in figure 3.2. For an excellent, and far deeper, introduction to this field, see Mackintosh (1983).

The descriptions of experiments in conditioning all follow a fairly set template. Like the description above, each phase of the experiment is addressed in turn, with a description of when each stimulus is presented during the learning phase. This lends itself to using an abbreviated notation that can describe the experiment in a short, unambiguous way. The simple acquisition-extinction experiment above would be described as:

Acquisition:
    CS + US

Test:
    CS $\rightarrow$ CR

Extinction:
CS

Test: CS → no-CR

This means the same thing as the verbose description above. Present CS (the bell) followed by US (food) repeatedly, and get a conditioned response CR (salivating). Then present CS alone for several trials, until we do not get a response. Of course, when we say there is no response, we mean that there is no response conditioned to CS. There might well be a baseline of response unrelated to any conditioning – gastric acid or saliva are produced to some extent even in the absence of food – but what we are looking for is a response that is in addition to this baseline. Also, the animal might give any number of other responses to the CS, but they are not the one we specifically have been conditioning it on, and are thus ignored.

The canonical conditioning experiment above has two phases: a first, where the association is acquired (called the acquisition phase) and a second, where the response is extinguished (appropriately called the extinction phase). That is the core of conditioning. It certainly seems simple, so why is this simple mechanism still being studied almost a hundred years after its discovery? It turns out that this simple mechanism is not simple at all.

3.2.1 Learning Protocols

There is more than meets the eye even for the seemingly simple type of learning that is classical conditioning. Let us revisit the simple acquisition/extinction experiment we discussed above. As simple as it seems, even this protocol has a few surprises.

First, what really happens during extinction? A simple idea would be that the connection between the CS and the CR that was formed during acquisition would now simply disappear again. An easy test shows us that this is not the case.

Consider again the protocol where the subject learns the connection between the CS and US, and then extinguishes the response again by presenting the CS without the US:

Condition:
CS + US

Extinguish:
CS → no-CR

At this point we cease all training for a time. After some time has passed (a couple of days, in the case of dogs conditioned with food) we try the CS again:

Test:
CS → CR(0.5)

The CR(0.5) notation signifies that the response CR is only produced at partial strength, or only part of the time.

It turns out that the connection is partially restored, albeit weaker than at its peak. We can often do this several times, with the response returning (but weaker) every time. This phenomenon is called spontaneous recovery (studied by Wagner et al. (1964) among others), and it is an important factor to keep in mind when designing learning experiments. Clearly, there is more going on than a removal of a connection. What would happen if we condition, extinguish, then recondition the connection as follows?

Condition:
CS + US

Extinguish:
CS

Recondition:
CS + US → CR

The subject has no problem relearning this connection; in fact the acquisition the second time is perceivably faster than the first time. If we do this for a few more iterations, we will see the reacquisition become more and more rapid. This reacquisition effect is also called a savings effect. We will see various explanations for this and other phenomena when we discuss proposed models below.

3.2.2 Classical experiments

There exists a collection of classical experiments in the field of classical conditioning that isolate – and exemplify – various features of this type of learning. These experiments are also often used to test learning models, isolating their strengths and shortcomings. Also, some of these experimental protocols will be used in later chapters to test our model. It is thus of some use to briefly describe these.

Where needed for proper understanding, these experimental protocols are described fully, with the experimental setup and control groups. For a purely descriptive account this is often not necessary however, and they are omitted in these cases in the interest of brevity.

Blocking

Conditioning exhibits a principle of parsimony: do not learn more than you have to. This is seen in a blocking experiment (Kamin, 1969). In blocking, we first condition a stimulus CS1 to an US. Then we try to condition CS1 together with a new stimulus CS2 to the same US.

Phase 1:
CS1 + US

Phase 2:
CS1 & CS2 + US

Test:
CS2 → no-CR

As we see, no conditioning occurs; when we test CS2, there is no evidence of learning. In other words, when one stimulus has already been associated with a certain event and elicits a response, trying to associate a second stimulus with the same event will be blocked by the first one.

The reason for this behavior is that since the first stimulus (CS1) already perfectly predicts the upcoming CS, adding CS2 does not give the animal any new information, and is thus redundant. That the system somehow evaluates (or behaves as if it evaluates) the information content can be clearly seen when we change the experiment slightly:

Phase 1:
CS1 + US(0.5)

Phase 2:
CS1 & CS2 + US

Test:
CS1 → CR(0.5)
CS2 → CR(0.5)

As CS1 predicts a weak US (again, the 'US(0.5)' indicates half strength of the stimulus), the response is also comparatively weak (in reality, the response might only show on half the trials rather than being half as strong; the distinction is not important at this level of abstraction). When presented together with CS2, it predicts a strong US, and the response is also proportionally stronger. Lastly, when we test CS2, it will elicit a response in proportion to the difference between the strong and the weak US. The information content needed to predict US has thus been split between CS1 and CS2. Gallistel (1990, p. 409) calls this explanation the *multidimensionality principle*: the change in stimulus strength depends on the strength of the associations to all the relevant stimuli.

Conditioned Inhibition

As we saw earlier with the phenomenon of spontaneous recovery, extinction does not seem to be a simple weakening of a previously learned association. Pavlov believed that extinction rather was the active learning of an inhibitory association that counteracts the excitatory association previously learned (Pavlov, 1927). Seen this way, extinction is thus an active learning process in its own right. The prototypical way of showing the existence of this independent process is by an experiment called conditioned inhibition, and was done first by Pavlov.

If the extinction is an active learning process, we should be able to show its existence independently of the acquisition process that previously spawned it. We do this by first associating a stimulus CS2 with an CS, then associating another stimulus CS1 with the absence of the CS when it would otherwise be expected. These two associations are alternated until no further change in conditioning occurs. We can then test the inhibitory stimulus CS1 with a third, already associated stimulus to see if it also becomes inhibited:

Prepare CS0:
CS0 + US → CR

Condition:
CS2 + US → CR
CS1 & CS2 → no-CR

Test:
CS0 & CS1 → no-CR

Here we teach the subjects that while CS2 by itself is reinforced, CS2 and CS1 together is not, which is interpreted that the CS1 somehow stops the US from occurring. When we test this with CS0, CS1 is able to inhibit CS0 even though they have never been paired before. Thus inhibition is an active process, not just the removal of a pre-existing association.

Lysle and Fowler (1985) found an interesting effect. They conditioned a group of rats to regard a CS1 as inhibitory as described above. When they extinguished the CS2 that had been used for the inhibitory conditioning, the inhibition of CS1 disappeared. When CS2 was conditioned again, CS1 once again became inhibitory. From this experiment it appears that the inhibitory properties are dependent on the excitatory conditioning. Conditioned inhibition is said to be a 'slave' process.

Negative Patterning

Negative patterning is an example of configurational learning, studied by Bellingham et al. (1985), mainly using eye-blink responses in the rabbit. It is also known as the XOR problem in connectionist literature. The required pattern for the system to learn is that either one of two stimuli is reinforced, while both of them together is not. Just like conditioned inhibition, negative patterning depends on inhibition as an active process.

Condition:
CS0 + US
CS1 + US
CS0 & CS1

Test:
CS0 → CR
CS1 → CR
CS0 & CS1 → no-CR

This is quite a bit harder to learn than it might seem at first; CS0 and CS1 both predict a reinforcer, so both together should, if anything, predict an even greater one. The system thus needs to be able to treat the combination of both stimuli as different from either stimulus alone, and selectively inhibit the output when this occurs.

This problem is not linearly separable - when represented in a space, there is no line that will separate the correct outcomes from the false ones. It can thus not be solved by a simple non-layered collection of nodes with excitatory or inhibitory connections. For a thorough treatment on this subject, see for instance (Hassoun, 1995).

Positive Patterning

Positive patterning is the flip side of negative patterning. This protocol, as well as negative patterning above, has been studied by Kehoe (1986). The subjects learn that while either of two stimuli do not predict a reinforcer, both together do:

Condition:
- CS0
- CS1
- CS0 & CS1 + US

Test:
- CS0 → no-CR
- CS1 → no-CR
- CS0 & CS1 → CR

The problem here is on the surface similar to negative patterning. The system needs to distinguish between either stimulus alone, or both at once. Whereas both stimuli together were inhibited in negative patterning, here it is either stimulus alone that is inhibited. As with negative patterning above, there is a simple logical interpretation of this protocol; it is an AND function.

It is important here to distinguish positive patterning from feature discrimination. In feature discrimination, either individual stimulus is still conditioned, but the combination of both gives a stronger reaction than either alone; the reaction is added from both stimuli. For positive patterning, it is required that the reaction to the single stimulus is very weak or nonexistent for the protocol to work (we will revisit this issue in chapter 4).

In practice, animals will respond weakly to the single stimulus even after extensive training.

Secondary and Higher-Order Conditioning

Some stimuli are innately reinforcing. The taste or smell of food, the sight or smell of a potential mate, the opportunity to socialize, pain or sickness, they are all reinforcers that animals respond to at birth. By contrast, other reinforcers need to be learned by being associated with reinforcing stimuli. Phobic stimuli seem to reside in a middle ground, where the animal is well prepared to acquire such associations, but nevertheless needs to be acquired during its lifetime (Öhman and Mineka, 2001).

To test whether a stimulus CS0 has become reinforcing, we can simply use it as an US and see whether the new stimulus CS1 has become conditioned (this was first studied by Pavlov (1927)).

Phase 1:
CS0 + US

Phase 2:
CS1 + CS0

Test:
CS1 $\rightarrow$ CR

Here, we first teach that CS0 is a predictor of a reinforcer US. Then, we teach that CS1 is a predictor of CS0. And when we test CS1, we find that this stimulus also elicits a response. In this way, it is in principle possible to build long chains of reinforcing stimuli.

In practice, this experiment is quite hard to execute properly; at the same time that we train CS1 to predict CS0, CS0 will repeatedly be presented without the US, and will be extinguished. To do this, the experimenter has to present CS0 on its own with the US occasionally, and even then, the effect is weak. The US is different, of course, in that it is innate in the subjects and not susceptible to extinction (indeed, to extinguish an US, we need to actively inhibit its expression, as we saw at the start of this section).

The existence of secondary conditioning is somewhat controversial. A number of studies have shown the existence of this effect, but in many cases, it

has been hard to reproduce (see the discussion in (Klopf, 1988)). It may be the case that these divergent results could be explained by the sensitivity to the stimulus length and inter-stimulus interval for this protocol.

Generalization and Discrimination

The prototypical experiments described above do capture much of the essence of conditioning, but there are a number of other factors that influence learning in general and classical conditioning in particular. We will briefly discuss some of these factors.

Generalization is the effect that stimuli similar to the one the animal has been trained on will elicit a response in proportion to its perceived similarity, and was first formally described by Pavlov (1927). When we talk about a stimulus, we often envision a bell, a light or some other distinct signal. But neither sounds nor lights are ever completely distinct from one another, but are points of a continuum in several dimensions like pitch, loudness and tonal quality for sounds, and color and intensity for light (Gärdenfors, 2000). How is a system to distinguish between similar stimuli - and when should it see similar stimuli as variants of the same signal, and when should it see them as separate signals?

In the general case, this problem is unsolvable; the system needs more information on which to base its learning. This added information is also given in the form of rewards and punishments.

Consider a dog that has learned that it will get a treat when a 440 Hz tone sounds. What should it do if we play a 450 Hz tone? 480 Hz? 2000 Hz? 440.1 Hz? As it turns out, functional characteristics of the aural system will quite naturally guide the learning to assume a response pattern much like the idealized one depicted below.

The response pattern assumes that sounds that are close enough in the auditory space will be treated almost the same as the exact stimulus we have the system conditioned to, while sounds further away will be treated as largely irrelevant.

If we continue the experiment by rewarding the 440 Hz tone, but not rewarding the 450 Hz tone (for example), the system will learn to discriminate between these two frequencies even though they were treated as similar at the outset. A side effect of this discriminatory learning is an effect called ‘peak

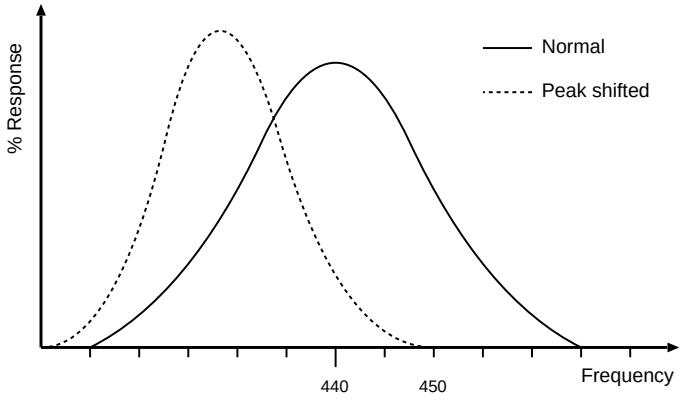

Figure 3.3: An idealized diagram of the response pattern as we vary the similarity of the presented stimulus with the learned one. Here, the subjects would have learned to respond to a sound with a frequency of 440 Hz. Playing sounds around this frequency will give a response pattern like the solid line in the figure. If we then inhibit responses for a sound with a frequency of 450 Hz, we get a response pattern similar to the dotted line.

shifting´ by Hanson (1959), where the point giving the maximum response will shift slightly away from the inhibitory stimulus. In the same way, starting to reward 480 Hz tones as well as 440 Hz will broaden the acceptance range of tones that will elicit a response.

Delay and Trace Conditioning

There is one further empirical aspect of learning that has an impact, and that is the timing of the conditioned and unconditioned stimulus (Gallistel, 1990). Unlike the previous protocols, this explicitly acknowledges the real time aspects of conditioning, such as inter-stimulus interval (ISI) effects and the modeling of timing of the conditioned response. The rate of conditioning is optimal at a certain ISI. In classical eye-lid conditioning in the rabbit, the optimal ISI is approximately 250 ms and decays exponentially with increased time (Smith et al., 1969; Schneiderman and Gormezano, 1964). Also, independently of the ISI, the CR tends to appear slightly before the CS (Desmond, 1990).

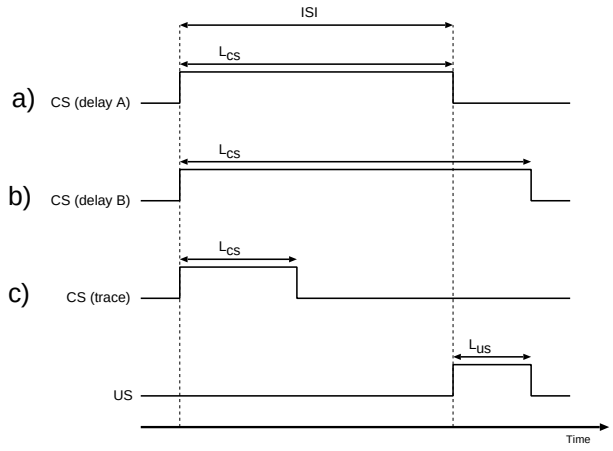

Figure 3.4: Delay and trace conditioning. L_{CS} is the length of the CS. ISI is the interstimulus interval, or the length of time between the onset of the CS and the onset of the US. In this illustration, the ISI is the same for all three cases. In a), we see so called delay A conditioning. The length of the stimulus (L_{CS}) is identical to the ISI, so the stimulus CS ends at the same time the US starts. In b) we have delay B conditioning. Here the stimulus length is the ISI plus the length of the US, making the CS and US overlap and making them end simultaneously. In c) we see trace conditioning. Here the length of the CS is shorter than the ISI, making the CS stop before the onset of the US. Unlike the other forms, this form of conditioning has a span of time between the CS and the US, making it necessary for the animal to somehow retain a memory trace of the CS for conditioning to occur.

Inter-stimulus interval effects

The effect of the inter-stimulus interval on the response level is one of the primary learning effects described in the model by Sutton and Barto (1990). Empirical studies have been done with both trace conditioning, where the CS terminates before the US is presented; and with delay conditioning, where the CS offset occurs at the onset of the US, or later (see figure 3.4). Note that when we talk about varying the ISI, this is under the assumption that other timing factors are held constant. Gallistel and Gibbon (2000) reviews data that show that this effect is time scale invariant. If the intertrial interval (ITI) is increased in proportion to the ISI, there will be no net effect on the response level.

In this section we identify three types of timing; two types of delay conditioning and one type of trace conditioning. In the first type of delay conditioning, that we will call delay A, the CS terminates exactly at the onset of the US. In the second type, delay B, the CS is present until the termination of the US.

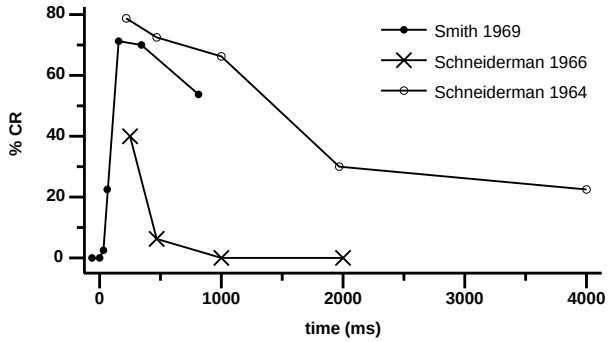

Figure 3.5: The effect of the inter-stimulus interval in classical nictitating membrane conditioning. Delay B conditioning after Schneiderman and Gormezano (1964). Trace conditioning after Smith et al. (1969) and Schneiderman (1966).

Finally, in trace conditioning, the CS and US have fixed lengths and only the ISI changes. In trace conditioning, the ISI can be both positive and negative. When varying the ISI in delay A and B, the length of the CS is varied as well, while in trace conditioning the length of the CS and of the ISI are independent.

The empirically determined profile for delay B and trace conditioning, according to Schneiderman (1966) and Smith et al. (1969) can be found in figure 3.5. The desirable behavior is for the response level to have a single peak at small positive ISI:s, no response at all for negative ISI:s, and asymptotically declining values as the ISI grows large.

Facilitation

As we have seen above, the response strength depends in part on the inter-stimulus interval. When the interval is long, the response will be low. If another stimulus is introduced between the CS and the US, the response can be made stronger:

Without facilitation:

CS1 + US

Test:

CS1 $\rightarrow$ CR(weak)

With facilitation:

CS1 + CS2 + US $\rightarrow$ CR(strong)

Test:

CS1 $\rightarrow$ CR(strong)

This is called facilitation with intermittent stimulus (the intermittent stimulus being CS2 above). The ISI between CS1 and the CS is the same in both cases; CS2 is introduced in between to act as a 'bridge' between the two other stimuli (Kehoe, 1982).

Feature Discrimination

A comparatively recent set of protocols for associative learning is serial and simultaneous feature discrimination. Serial feature discrimination is also known as *occasion setting* (Bouton and Nelson, 1998). Serial feature-negative discrimination (or negative occasion setting) aims to establish a stimulus CS1 as an occasion setter that indicates that no reinforcement will be forthcoming on the next conditioning trial.

Condition:
CS1 + CS2
CS2 + US

Test:
CS1 → no-CR
CS2 → CR
CS1 + CS2 → no-CR

The characteristic of an occasion setter is that it does not itself become conditioned to the US; compare this protocol to second-order conditioning above. Instead, it acts as a contextual signal to the learning system, indicating the applicability of a learned association.

Serial feature-positive discrimination is similar:

Condition:
CS1 + CS2 + US
CS2

Test:
CS1 → no-CR
CS2 → no-CR
CS1 + CS2 → CR

Here, CS1 signals that the forthcoming CS2 will be followed by a US, while its absence indicates that it will not. The test shows that CS1 does not become conditioned itself.

Davidson (1998) suggests that a role of positive occasion setting is to signal motivational states. Internal bodily stimuli (feelings of hunger or thirst, for example) would be positive occasion setters for associations between food and post-ingestive US. Similarly, negative occasion setters could be stimuli such as satiation. The occasion setters would function as an internal context.

Simultaneous feature-positive discrimination and simultaneous feature-negative discrimination are the simultaneous versions of positive and negative occasion setting, respectively (Schmajuk et al., 1998). These protocols are normally not considered occasion setting.

Simultaneous feature-positive discrimination has a similar protocol to the serial variant. Alternate between conditioning two stimuli, and non-conditioning of one of them - note the use of '&' rather than '+' to signify simultaneous presentation:

Condition:
CS1 & CS2 + US
CS1

Test:
CS1 → no-CR
CS2 → CR
CS1 & CS2 → CR

The animal will learn to react to CS2 but not to CS1. This is similar to an overshadowing protocol. The protocol for simultaneous feature-negative discrimination is identical to that of conditioned inhibition.

Disinhibition

Contextual disinhibition occurs when a stimulus that is inhibited in one context loses its inhibition elsewhere (Schmajuk et al., 1998). It is a basic test of context-dependent inhibition. In the protocol below, the CX notation indicates the context in which the experiment takes place.

Condition CS0:
CX1: CS0 + US

Followed by extinction:
CX1: CS0

Test:
CX1: CS0 → no-CR

Move CS0 to a different location:
CX2: CS0 → CR

Thus, the inhibition of CS0 is local to the context in which it is inhibited. The conditioning, on the other hand, is intact over any context. To get context-dependent conditioning, a feature discrimination protocol must be used.

Disinhibition by novel stimulus is another experiment where inhibition is suppressed. This effect is achieved by presenting a novel stimulus for the animal to react to (Gray, 1975).

Condition:
CS0 + US

Extinction:
CS0

Test:
CS0 → no-CR

New stimulus CS1:
CS0 + CS1 → CR

In animal experiment terms, contextual disinhibition would be moving the animal to a different experimental cage, while disinhibition by a novel stimulus would be presenting a novel stimulus in the original cage. Also, in contrast to contextual disinhibition, disinhibition by novel stimulus is a transient phenomenon; the animal will quickly generalize the inhibition to the changed surroundings.

Statistical Effects

Conditioning is sensitive to the statistical dependency between events (Rescorla, 1968; Gallistel, 1990). While it was once believed that the pairing of a CS and a

US were sufficient for conditioning, it has been shown that this is not the case. In an important experiment, Rescorla (1968) presented subjects with random pairings of CS and US. On some trials the CS was presented alone, on some it was followed by the US, and on other trials the US occurred without the CS. In one experimental group, the CS did not have any statistical correlation with the US, while in the other the CS correctly predicted the US at half of the trials.

In the first group no conditioning occurred, while in the second group the animals produced the CR at a rate proportional to the predictive level of the CS. This clearly shows that the animals are able to discern statistical correlations between stimuli, and are not just blindly pairing co-occurring stimuli together. Of course, this does *not* imply that they have any kind of concept of statistics, but merely that the learning system is able to behave in such a way that it can discern statistical correlations from non-correlations.

3.2.3 Models of Conditioning

A number of models have been proposed to explain classical conditioning. Many of these are psychologically based models that only deals with a very restricted subset of available data. Some models, however, are intended to have fairly broad explanatory powers, and are detailed enough to give specific predictions in various cases. In this section, we will take a look at some proposed models of classical conditioning. This will set the stage for the introduction of our model in the next chapter.

The models in this chapter (as well as the model we will introduce in the next) implicitly assume that there are processes outside their purview that preselects the stimuli and rewards before they become subject to learning. It is assumed that various attentional processes select and sift through the available stimuli to give salience only to those that are of significance to the system; the stimuli thus have a fixed salience, while the importance of the reward signal is varied by the model according to how well the stimuli represent the reward. This is in contrast to another class of models discussed by Pearce and Hall (1980) where instead the stimulus salience is varied during conditioning. This brings some aspects of attention into the learning model itself, and gives a natural explanation to such phenomena as latent inhibition. None of the models tested here belong to this class of models, however.

To test the models, we have implemented all models – with the exception of the Rescorla-Wagner model – in a simulator. They have been run with some of the classical learning experiments described in previous sections, and then compared with each other and with experimental results involving animal

subjects. Each model has as far as possible been given reasonable parameter values (taken from the model descriptions from the authors) and the simulations have taken into account differences in temporal resolution and numerical levels.

Rescorla-Wagner

The Rescorla-Wagner model (Rescorla and Wagner, 1972) is possibly the earliest computational model of conditioning that is still discussed today. Many later models have implicitly or explicitly made use of the ideas espoused by this model, including the Sutton-Barto model presented in the next section and our own model in chapters 4 and 5. As it does not work in real time, but only on a per-trial basis, it is not tested as the other ones described here.

Their model tries to minimize the discrepancy between the reward predicted by the sum of CS:s and the actual reward given by the US. This is again the multidimensionality principle discussed by Gallistel (1990). When there is a discrepancy, all the CS:s are changed proportionally, scaled by a learning constant:

$$\Delta w_i = \alpha(US - \sum_j CS_j w_j) \tag{3.1}$$

This model turns out to handle a number of classical conditioning experiments, including blocking and conditioned inhibition (Schmajuk, 1997, page 38). As it is not real-time, however, many other effects relating to timing can not be reproduced.

The Sutton-Barto Model

The Sutton-Barto (SB) model (Sutton and Barto, 1981), is an early time-derivative model of reinforcement in conditioning. It works in real-time in contrast to the Rescorla-Wagner model (Rescorla and Wagner, 1972), from which it is descended. This model is a precursor to many later computational models, including the TD model (Sutton and Barto, 1987), the SBD model (Blazis et al., 1986) and the Klopf model (Klopf, 1988).

The reinforcement in the SB model is the time derivative of the sum of the stimulus strengths:

$$\begin{aligned} \dot{Y}_t &= Y_t - Y_{t-\Delta t}, \\ Y &= \sum_i CS_i V_i + V_{US} US \end{aligned} \tag{3.2}$$

where the V_i is the stimulus strength of CS_i, and V_{US} is the strength of US. With suitably chosen constants, these equations will account for all predictions of the Rescorla-Wagner model.

An eligibility trace mechanism is added to account for temporal relationships:

$$\begin{aligned} \bar{X}_{i,t+1} &= (1-\delta)\bar{X}_{i,t} + \delta X_{i,t}, \\ X_{i,t} &= \begin{cases} 1, & CS_i \text{ present} \\ 0, & CS_i \text{ absent} \end{cases} \end{aligned} \tag{3.3}$$

The rate of decay δ is in the range $0 < \delta < 1$. The same trace is also used in the TD model. The reinforcement signal and the eligibility trace are combined as:

$$\Delta V_i = \beta \dot{Y} \times \alpha_i X_i \tag{3.4}$$

with α_i and β positive learning rate constants.

We have run a series of simulations covering some of the learning experiments described in the previous section. During these simulations, the constants chosen were $\alpha_i = 0.1, \beta = 1, \delta = 0.2$; these are the values used by Sutton and Barto.

As the model is an early one, it is not surprising that it has a number of problems in these simulations.

Like the other models, the SB model has no problems with acquisition and extinction. Unlike the Klopf and Balkenius models (see pages 67 and 70), SB works over a wide range of CS-US intervals due to the nature of the eligibility trace function used. However, SB does not model the S-shaped acquisition curve that Klopf and Balkenius does.

The inter-stimulus interval effects are poorly modeled by this method. For example, the model shows inhibitory conditioning with ISI close to zero during

trace conditioning (figure 3.6). This seems to be inconsistent with empirical data (Smith et al., 1969). The model also fails to account for ISI effects during delay conditioning, with no decay of asymptotic learning with longer ISI:s, and with overall learning level dependent on proximity between CS_{off} and US_{off}, as seen in figure 3.1. As Sutton and Barto (1990) notes, this model does not, and is not intended to, account for these effects.

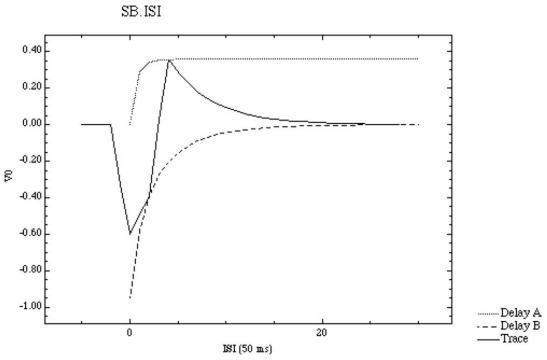

Figure 3.6: The behavior of the Sutton-Barto model during ISI trials.

No reacquisition effects are modeled by SB. Blocking effects are handled satisfactory, which is not surprising, as the capabilities are a superset of Rescorla-Wagner, and blocking is one of the effects modeled well by that model.

Secondary conditioning with no overlap between CS0-US or CS1-CS0 works as expected; CS1 is reinforced by CS0, albeit weakly, as CS0 is simultaneously extinguished. However, when the inter-stimulus intervals are short enough to make the stimuli overlap, an erroneous inhibitory conditioning is produced. Briefly, the reason is that when CS0 and CS1 overlap, V_0 is initially small, making $\dot{Y}$ small. Thus, change in X is the primary factor in changing V_0. The greatest change in X is for CS_{off}, and when overlapped with US, this negative change will influence V_0, thus producing the effect. The analogous effect is achieved for phase two of secondary conditioning, as V_0 is large and V_1 is small.

Facilitation effects are modeled fairly well; the stimulus trace model used directly facilitates this effect.

The Temporal Difference Model

The TD (temporal difference) model (Sutton and Barto, 1987) was introduced as an extension of the earlier Sutton-Barto model (Sutton and Barto, 1981). According to the TD model, the goal of conditioning is to predict the temporally discounted value of all future rewards. At time t, this prediction is called and is calculated as the sum of all weights V_i for the CS_i at time t:

$$\bar{V} = \max(\sum_i V_i CS_i, 0) \tag{3.5}$$

Here, the weight V_i represents the contribution from CS_i to the total prediction. During learning, these weights are updated according to the following equation,

$$\Delta V_i = \beta(\delta_{t+1} + \gamma \bar{V}_{t+1} - \bar{V}_t) \times \alpha_i \bar{X}_i \tag{3.6}$$

where δ_{t+1} represents the strength of the CS at time $t+1$, γ is the discount factor, α_i and β are learning rate constants, and $\bar{X}_i$ is a trace of stimulus CS_i. Thus, the first term of equation 3.5 can be seen as the expected change in US from t to $t+1$ (called reinforcement by Sutton), while the second term is a function of temporal proximity to the CS. The definition of $\bar{X}_i$ is identical to equation 3.3:

$$\begin{aligned} \bar{X}_{i,t+1} &= (1-\delta)\bar{X}_{i,t} + \delta X_{i,t}, \\ X_{i,t} &= \begin{cases} 1, & CS_i \text{ present} \\ 0, & CS_i \text{ absent} \end{cases} \end{aligned} \tag{3.7}$$

This section describes a number of simulations of the TD model. The actual code was downloaded from R. Sutton's web-site and interfaced with our simulator. The model itself is, thus, identical to Sutton's implementation. Using this implementation, we have successfully reproduced the simulation data as presented in Sutton and Barto (1990).

The TD model is able to reproduce several aspects of classical conditioning. Most importantly, it models the ISI-dependency shown in empirical studies.

Figure 3.7 shows the asymptotic value of the weights as a function of the ISI for trace conditioning.

Unlike the other models discussed in this chapter, the ISI-curve is an emergent effect of the learning equations and handles a variety of different ISI:s and stimulus lengths.

The TD-model also handles blocking effects in a satisfactory manner, but the presence of secondary conditioning depends on the stimulus length. This appears as a shortcoming of the model.

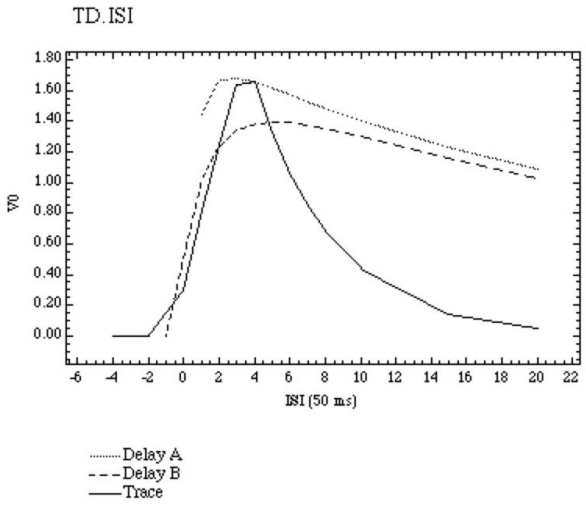

Figure 3.7: The ISI curve for the TD model.

Interestingly, the TD model suffers from a similar problem as the Sutton-Barto model for secondary conditioning: overlapping CS:s produce inhibition rather than excitation, as shown in figure 3.8. The reason this effect is not apparent for first-order acquisition, is that US level is not taken into account when computing $\bar{V}_i$, in contrast to the Sutton-Barto model.

The results for blocking, inhibitory conditioning and facilitation are the same as for the Sutton-Barto model; the model performs reasonably well. Like the SB model, the TD model does not show the S-curve for acquisition. A related shortcoming is that the TD model does not show any reacquisition effects or spontaneous recovery.

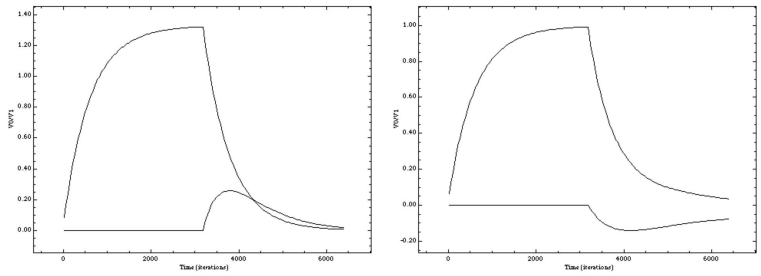

Figure 3.8: Secondary conditioning in the TD model with and without overlapping stimuli. The upper curves show V1 and the lower curves show V2. (Left) With non-overlapping CS1 and CS2, the secondary conditioning effect is clearly visible. (Right) With overlapping conditioned stimuli, V2 shows inhibition rather than excitation.

The Klopf Model

The Klopf model was introduced in 1982 as the Drive-Reinforcement (DR) model (Klopf, 1982), but we present the model as described in (Klopf, 1988). In this model, the output at time t – the UR or the CR – is given by the value $y(t)$. This is calculated as the sum of all CS representations x_i multiplied with their corresponding weights w_i. The value θ is a threshold which was set to 0 in the simulations reported by Klopf.

$$y(t) = \sum_{i=1}^{n} w_i(t) x_i(t - \theta,) \tag{3.8}$$

During conditioning the weights change according to the equation,

$$\Delta w_t(t) = \Delta y(t) \sum_{j=1}^{t} c_j |w_i(t-j)| \Delta x_i(t-j), \tag{3.9}$$

where c_j are learning constants, $|w_j|$ are the magnitude of the individual weights, and Δx_i the change in the CS representation. It is stated in (Klopf, 1988) that only positive changes should be considered in the equation above, that is, if Δx_i is less than 0 it is set to 0 for the purposes of equation 3.9. It should also be noted that there are both excitatory and inhibitory weights in the model; these are treated separately and are constrained to stay on the positive and negative side respectively. The model also requires that all weights

must be larger than 0. In the simulations in (Klopf, 1988), a minimum value of 0.1 is used.

Equation 3.9 illustrates a fundamental assumption of the Klopf model, namely that changes in the output should be correlated with changes in the inputs to determine whether learning should occur. When both the input and the output changes, the weights should increase. To allow non-zero ISI:s, each CS is assumed to leave an eligibility trace in the system. This is represented by the sum in equation 3.9 which is used as a memory that extends backwards in time.

The model is able to mirror a complex ISI curve since an array of learning constants, c_i, are used to determine the shape of the curve. The constants are explicitly chosen to reflect the ISI-curve for classical delay conditioning (Klopf, 1988). Figure 3.9 shows the simulation result for the same experiment as the TD model described above.

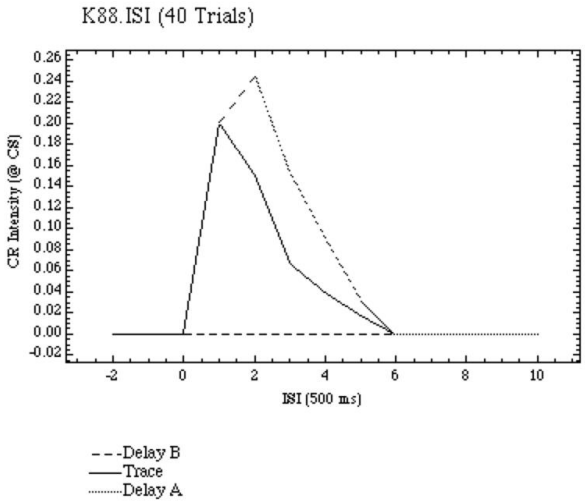

Figure 3.9: The ISI curve for the Klopf model.

Although the Klopf model has the ability to produce a range of ISI-curves depending on the learning rate constants, it has almost no predictive value in this area since the curve is essentially rigged by the constants c_i. A further limitation is the fixed length memory buffer used for the eligibility trace. The use of a memory buffer limits the temporal resolution of the model as well as the length of the ISI.

The model also shows the initially accelerated S-shaped learning curve that is observed in animals. This is the result of the multiplication of w_i in the eligibility trace. This makes learning faster when the weights are larger.

This unusual aspect of the learning equation also results in reacquisition effects. Since the weights are larger after extinction than before any learning has taken place, learning will be faster in the second acquisition phase. Although Klopf presents a simulation where the reacquisition effect is clearly visible, this effect only occurs for the *first* reacquisition. All reacquisitions after the first one are identical for the parameters used in (Klopf, 1988); see figure 3.10. This was not illustrated in Klopf's original article. Also, the S-shaped learning curve disappears for the reacquisitions.

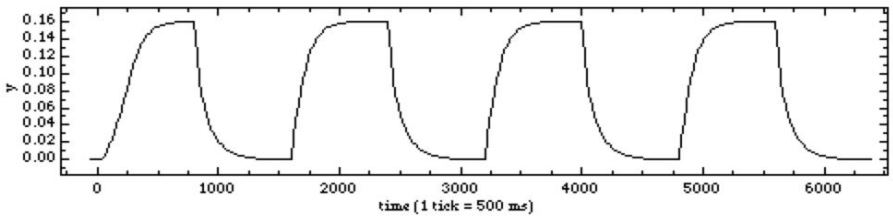

Figure 3.10: The reacquisition effect in the Klopf model. The savings effect only appears the first time.

The reacquisition effect depends on the learning rule where the weight change depends on the magnitude of the weight. A larger weight will change more than a smaller one for the same reinforcement signal. This aspect of the model is also the reason for the initially accelerated S-curve in acquisition.

Klopf (1988) reports simulation of secondary conditioning which were reproduced by our simulator. It is interesting to note that this experiment results in the inhibitory effect described on page 66 as when run with the TD model. The Klopf model handles blocking in a satisfactory manner.

Facilitation by an intermittent stimulus is also modeled correctly. This effect is both visible at an ISI smaller than the length of the eligibility trace, and at larger ISI where the mechanism in effect extends the length of the fixed memory buffer.

The Balkenius Model

Like the Klopf model, the model presented in Balkenius (1995) and Balkenius and Morén (1998b) is based on a neural interpretation of the conditioning mechanism. But contrary to the Klopf model, the Balkenius model uses a network rather than a single node. Like the Klopf model, the Balkenius model separates inhibitory and excitatory learning, but this is made explicit in the formulation of the model. Here, only the excitatory side of the network is described. The equations for the inhibitory side are identical. The output of the model is given by:

$$CR(t) = x^+(t) - x^-(t), \tag{3.10}$$

$$x^+(t) = \sum_{i=1}^{n} \sum_{j=0}^{\tau} w_{ij}^+(t) CS_i(t-j). \tag{3.11}$$

The extra index j in equation 3.11 corresponds to a tapped delay-line for each CS with length τ. This approach differs from the eligibility traces used in the other models but has a similar role. The representation is usually called a multiple-element stimulus trace and has the advantage that it can support more complex associations than a single eligibility trace. On the other hand, it requires many more variables since the number of weights must equal the length of the stimulus trace.

The reinforcement – that is, the weight change – is calculated as:

$$R^+ = [US + \delta(\Delta x_{t+1}^+ - \Delta x_{t+1}^-) - (\Delta x_t^+ - \Delta x_t^-)]^+ \tag{3.12}$$

In this equation, δ is the discount factor which is responsible for secondary conditioning. This equation can be compared to equation 3.5 for the temporal difference model. In that model, the absolute values are used in the learning equation. In this model, it is instead the changes that contribute to the reinforcement. This is consistent with the idea that the model tries to predict the level of the CS rather than the integral over it as the TD model does.

Like in the Klopf model, changes in the US level are correlated with changes in the CS levels and during conditioning, the weights change according to the equation:

$$\Delta w_{ij}^+ = \gamma^+ R^+ [\Delta CS_i(t-j)]^+, \tag{3.13}$$

where γ^+ is the learning rate.

The Balkenius model differs from the other models in the way the ISI effect is modeled. In the TD and Klopf model, it is the result of the eligibility trace. In this model, it results from secondary conditioning within the stimulus trace. Figure 3.11 shows the ISI curve for the model with the discount factor of $\delta = 0.90$. The curve directly reflects the discount factor since the asymptotic value of the weights converges to $\delta^{(i-1)}$ for an inter-stimulus interval $i \leq 0$ (Balkenius and Morén, 1999).

A somewhat surprising effect of the stimulus representation is that the model predicts the S-shaped learning curves for ISI:s larger than 1 time unit. For an ISI of 1 time unit, the learning curve is similar to that of the TD model, but for a larger ISI, the curve is initially accelerating, giving rise to the characteristic S-shape.

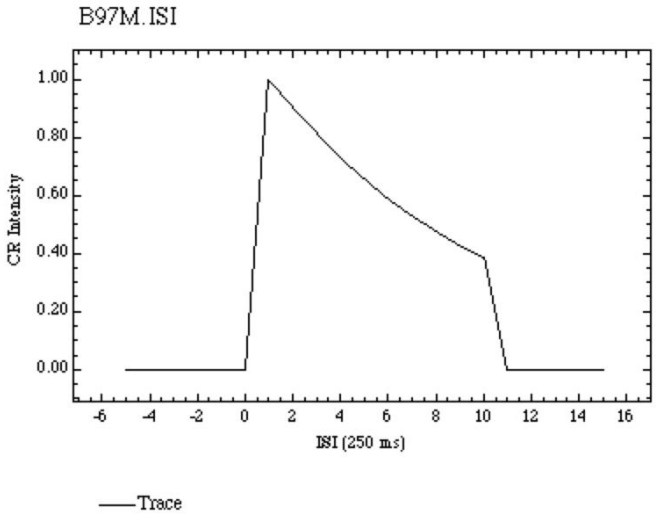

Figure 3.11: The ISI curve for the Balkenius model with $\delta = 0.9$.

This model's explanation of the S-shaped curve is radically different from the one offered by the Klopf model. In the Klopf model, the learning equation itself is constructed to give the S-shaped acquisition curve, while in the Balkenius model, it is the result of the mechanism for secondary conditioning given by equation 3.12. For an ISI of 2 ticks, the initial acquisition curve is quadratic, for an ISI of 3 ticks, it is cubic, and so forth.

Secondary conditioning is also handled by the model. In fact, the Balkenius model gives robust secondary conditioning on both the experiment described

in (Klopf, 1988) and the experiment in (Sutton and Barto, 1990). The TD model did not do well on the Klopf experiment, and the Klopf model was not able to handle the experiment used for the TD model.

This robust capability for secondary conditioning is available at the cost of ignoring the difference between trace and delay conditioning. Since only positive changes in the CS is used in the model, trace and delay conditioning will appear identical to the learning mechanism. Empirical data suggests that this should not be the case since delay conditioning usually results in faster and stronger learning than trace conditioning.

Like the other models, blocking is handled without problems. It appears that since Rescorla and Wagner (1972), blocking is the first experiment to be tested for any model of conditioning.

Another shortcoming of the model is that it is not able to model reacquisition effects. Like in the Klopf model, the weights reflect the fact that earlier conditioning has occurred, but this is not utilized in the learning equation.

Finally, it is possible to raise the same objection to the multiple element stimulus trace in this model as to the eligibility trace in the Klopf model: it limits the temporal resolution and sets a fixed length on the memory for passed events.

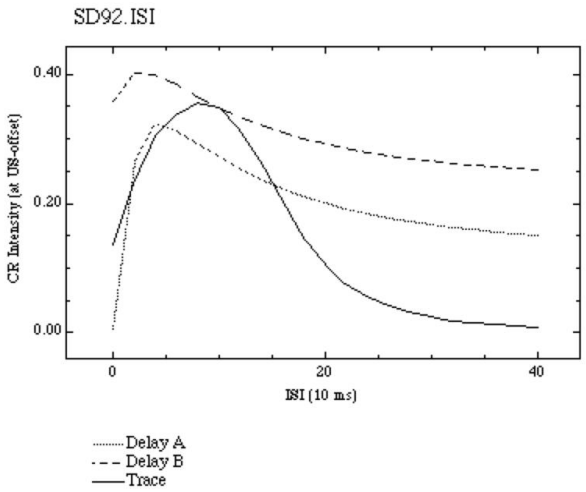

Figure 3.12: The ISI curves for the Schmajuk-DiCarlo model

The Schmajuk-DiCarlo Model

The Schmajuk-DiCarlo (SD) model was introduced in 1992 and was shown to model a number of classical conditioning phenomena (Schmajuk and DiCarlo, 1992). Especially interesting is its ability to model the effect of various types of configurational stimuli and the effects of hippocampal lesions on conditioning. Here, however, we will only investigate the more fundamental abilities of the model. The complete characterization of the model can be found in Schmajuk and DiCarlo (1992) and we will only describe the equations responsible for the dynamics found in the simulations. A stimulus, S_i, gives rise to a short-term memory trace X_i that is described by:

$$\frac{dX_i}{dt} = -K_1 X i + K_2(K_3 - X_i)CS_i, \tag{3.14}$$

K_1, K_2 and K_3 are constants that determine the passive decay of the trace, the rate of increase, and the maximum level of the trace, respectively. The associative strength, VS_i changes according to the equation:

$$\frac{dVS_i}{dt} = K_5 K_6(1 - |VS_i|)EO, \tag{3.15}$$

where K_5 is an output level constant and K_6 is the learning rate constant. EO describes the error in the prediction of the model. This error is calculated as the difference between the US level and the sum of all stimulus traces multiplied with their respective weights:

$$EO = US - \sum_i K_5 X i VS_i \tag{3.16}$$

In the simulations, we followed the complete description of the model given in Schmajuk and DiCarlo (1992), and the equations above should serve only as an indication of the dynamics of the model and not as a complete description.

The ISI curves for the model are shown in figure 3.12. Together with the TD model, the SD model is the only one to model the difference between the trace and delay conditioning in a qualitatively correct way. The SD model also correctly models blocking and conditioned inhibition. Of the models we tested, the SD model is the only one to show a reacquisition effect that increases with

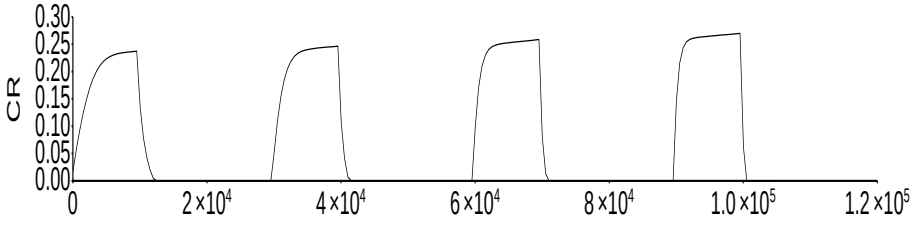

Figure 3.13: The reacquisition effects in the SD model. This is the only model discussed here where each subsequent acquisition is faster than the previous one.

each repeated relearning. This is shown in figure 3.13 where four acquisition and extinction phases are presented.

A serious difficulty for the SD model is that it does not allow for secondary conditioning as it is usually formulated. This is a direct consequence of equation 3.16. Since the change in associative strength depends on the difference between the CS and the aggregate prediction given by the sum in equation 3.16, there is no room for secondary conditioning. It can handle the special case when CS1 and CS2 are paired simultaneously, but not the more common case when they are sequential (Schmajuk, 1997).

It is interesting to note that the SD model is able to model facilitation but not secondary conditioning. This implies that secondary conditioning is not necessarily required for facilitation.

It is worth noting that Schmajuk has developed several models, including developments of the SD model described in (Schmajuk, 1997). Especially interesting is his escape-avoidance model of operant conditioning. We have not yet tested these in simulation.

3.3 Instrumental Conditioning

In classical conditioning, the subject can not influence the occurrence of a reinforcer. They get rewarded or punished no matter what they do; all they can do is learn to predict when it is coming, and prepare for the inevitable. Instrumental conditioning, on the other hand, is a process where the subject is an active participant. The actions of the subject in response to the stimulus influences whether or how much the system will be reinforced. Schmajuk (1997) presents a very good introduction to this subject; Sutton and Barto (1998) is another good text with a machine learning approach.

It was Thorndike that started the study of instrumental conditioning in the 1890:s. His experimental setup was a 'puzzle box', where some action, such as pulling a string, would release a catch and release the animal within. In these boxes he would put a hungry cat with food waiting outside, and then study its behavior and the time it took for the animal to release itself from the cage. The same animal would be set into the box repeatedly to see how and what it learned as it encountered the same situation multiple times.

When put in the puzzle boxes the cats would start to explore their surroundings in order to find a way out. Eventually, they would hit upon the action that opened the box. After a period of time, the cats where put into the box again, to find their way out. In contrast to earlier expectations, Thorndike found that the cats never had a moment of epiphany where they realized how the puzzle box was opened. Instead, the time it took for the animals to open the box gradually decreased, as the action that opened it became more likely for them to perform.

Based on these experiments, Thorndike formulated his *law of effect*: if an action is followed by a pleasant experience, the probability of performing the action increases, and if the action is followed by an unpleasant experience, the probability decreases.

Instrumental conditioning can be seen in two related, but not identical, ways. The question is if the system learns that a certain stimulus-response association leads to another state (that may or may not be reinforced), or if the system learns that the stimulus-response association leads to a reinforcer (that also happens to be another state). This can be written as an S-R-S' association, or as an S-R-S* association, respectively, where the first 'S' is called a discriminative stimulus or state.

Which interpretation is correct? Tolman and Honzik (1930) performed an experiment where a number of rats were allowed to run in a fairly complex maze. One group received a reinforcer (food) at the goal box, while the other did not. The reinforced group gradually learned to run to the goal box, while the second group ran freely, showing no particular interest in that particular place in the maze. After ten days, the second group were also reinforced in the goal box. Rather than exhibiting the gradual increase in performance of the first group, they immediately did as well in the task as the first group of rats. Evidently, they had learned the layout of the maze - including the position of the goal box - in the absence of a reward, and immediately utilized that knowledge when they got reinforced and there was a reason to do so. This implies that the rats learned to associate a state (or place, in this case) and an action with the resulting state. Had the rats used S-R-S* associations exclusively, they would not have learned anything about the maze - and certainly not as much as the reinforced rats - until they also received a reinforcer.

The perhaps most important aspect of instrumental conditioning is that the system learns associations where it can influence the outcome with its own actions. The system might have learned several associations grounded in the same state (S0-R1-S1, S0-R2-S2, ...), and is free to choose between actions depending on the perceived reward of each one. Another consequence is that the system can learn chains of actions: (S1-R1-S2, S2-R2-S3, ..., Sn-Rn-Rew), that can shape sequences of actions. Those who are familiar with computational reinforcement learning will find this very familiar, and, indeed, the difference between the field of instrumental conditioning and reinforcement learning is more in the object of study than in the actual mechanisms studied (indeed, reinforcement learning started out as implementations of instrumental conditioning for use in AI and robotics).

3.3.1 Reinforcement schedules

In instrumental conditioning, there are several *schedules* of reinforcement that can be used to reinforce behavior; these schedules influence how well the system learns its associations, as well as how long these associations are retained after reinforcement has ceased (Ferster and Skinner, 1957).

The canonical schedule is continuous reinforcement. Every time the animal performs the desired action, the researcher gives the animal a reward. There are other possibilities, however.

First, the animal can be reinforced not every time, but every n:th time the behavior is performed. It can be as often as every second time, or as rare as every hundred. This is called ratio scheduling. Instead, the animal can be reinforced after a given amount of time, no matter how much it has responded during that time (as long as it responded after the time interval). This is called interval scheduling.

Second, reinforcement can occur after a fixed time or after a fixed number of responses; or randomly, spread around an average time or average number of responses. This can be combined with ratio or interval scheduling to create four common schedules: Fixed Interval, Variable Interval, Fixed Ratio and Variable Ratio.

The choice of schedule affects both the learning and relearning for the animal (Mackintosh, 1983, pp. 127-128). The fixed schedules (Fixed Interval and Fixed Ratio) lead the animal to start to respond only after some time has passed. With an FI schedule, the rate of responding gradually increases until it reaches

a maximum around the time when the reinforcer is due to be presented, while a FR schedule will start abruptly after a resting period, and then stay at a constant level until the reinforcer appears. Variable schedules elicit a steady response frequency, dependent on the expected time or ratio for the reinforcer.

One effect of these schedules is that while learning is slower than in a continuous reinforcement schedule, the learned response is also more resistant to extinction. An intuitive way to understand this would be to consider the event from the animal's point of view: if a reinforcer normally shows up after an average number of responses, the failure of the reinforcer to do so may simply be a case of it taking longer than average.

3.3.2 Two-Process models

In Mowrer's influential two-process theory of learning (Mowrer, 1973), instrumental conditioning was assumed to proceed in two steps. First, the positive emotional aspects of the rewarded situation would be learned by classical conditioning. In the second step, the rewarding properties of the situation were used to reinforce the appropriate behavior; Klopf et al. (1993) presented a neural network architecture that can transform a model of classical conditioning into one of stimulus-response (S-R) learning in this way.

In effect, instrumental learning can be seen as secondary conditioning of a response to a specific set of stimuli. Since classical conditioning is seen as an important part of instrumental conditioning, all the properties of classical conditioning carry over to the instrumental case. This is again a simplification compared to real animals where the dynamics of the two types of learning are not entirely identical. For example, responses learned by classical conditioning are influenced directly by change in motivation while this is not always the case with instrumental or S-R learning (Balleine, 1992; Balleine et al., 1995). The effect also appears to depend on the modalities involved and the particular species examined (Gaffan, 1992). The relation between the model and motivational and emotional processes is described at length in (Balkenius, 1995).

A major benefit of this architectural organization is that it cleanly separates the issue of stimulus evaluation and action selection. First, the evaluation can be reused in different situations without change. Had the stimulus evaluation directly triggered a given response, this would only be applicable to the learned situation, but by having a second - instrumental - step, the system can learn different actions for the same stimulus evaluation. Conversely, it allows a given response pattern to be learned and reused for different stimuli,

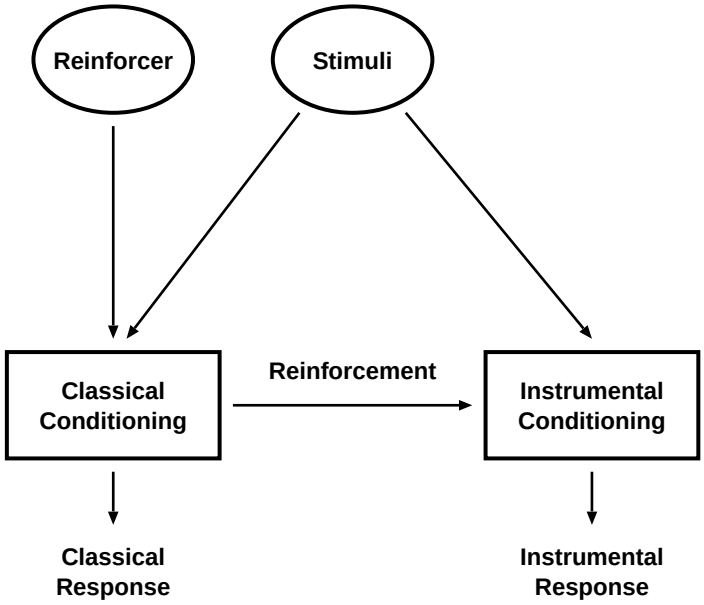

Figure 3.14: A schematic illustration of a two-process model. The classical conditioning system receives stimuli and a reinforcer, and conditions on them, giving rise to a classical response. The instrumental learning system also receives the same stimuli, but gets its reinforcement not directly, but through the classical conditioning system.

as it is not triggered by the stimulus itself, but by the evaluation which can be common over a whole class of stimuli.

Classical conditioning is here seen as a mechanism that assigns values (expected rewards and punishments) to each stimulus or stimulus combination. These valuations of stimuli or states lie at the heart of reinforcement learning as it is used in robotics and control. The classical conditioning system essentially plays the role of an adaptive critic. In this respect, the behavior of the network is very similar to TD(l) (Sutton, 1995; Sutton and Barto, 1998) as well as the Q-learning algorithm (Watkins, 1992; Morén, 1998).

3.3.3 The Nature of Reinforcement

As hinted at previously, the concept of a reward or a punishment can be surprisingly difficult to pin down. The common usage of these terms do not accurately correspond to their function in animal learning.

Early learning theorists discussed conditioning in terms of rewards and punishments. Soon it became clear, however, that the situation was more complex than anticipated. Sheffield et al. (1951) performed a simple experiment, where male rats were placed in a runway maze with a female rat in heat in a goal box at the other end. The experimenters measured the time it took for the male rats to run along the runway. When a rat arrived at the goal box, it would be removed before having a chance to interact with the female, and placed back at the start of the runway. The male rats rapidly increased their running speed. Now, the problem is that they never receive a reward. They are removed from the goal box without ever getting to mate with the female, and yet, they increase their running speed as the trials progress.

The solution is to talk about *reinforcement*. In the spirit of Mackintosh, it can be loosely defined as "..an event whose occurrence in relation to other events supports some change in behavior." (Mackintosh, 1983, p. 20). This definition allows the use of the term in classical as well as instrumental conditioning. In the case of instrumental conditioning, it means simply something that, by its presence or absence, changes the frequency of responding to a stimulus.

Reinforcers can be positive or negative. A positive instrumental reinforcer acts to increase the frequency or intensity of the response to make the reinforcer more probable, while a negative reinforcer increases the frequency of responding in order to lessen the probability of the reinforcer.

Note that a negative reinforcer is different from a punishment, where responding increases the probability of a negative event. For a negative reinforcer, the animal will respond in order to avoid a negative stimulus, thereby actively working to avoid it. In punishment, the response it what will elicit the negative stimulus, so the animal will carefully avoid responding at all (at least in the manner the experiment is set up).

Primary and Secondary Reinforcement

Not all reinforcers are created equal. A primary reinforcer is a stimulus – internal or external – whose reinforcing properties are innate to the animal, such as food, sex, pain or sickness. These reinforcers are generally strong, difficult to impossible to extinguish, and tend to be closely tied to fundamental aspects of survival for the animal.

A secondary reinforcer, on the other hand, is a reinforcer that has been learned during the animal's lifetime. The protocol of secondary conditioning (page 53) is an artificial way of creating a secondary reinforcer.

What kind of reinforcers are there? Apart from the expected - and common - ones of food or drink, pain (electric shock), and sickness, there are a number of other stimuli that act as reinforcers even without any previous conditioning. Rolls (1999, pp. 272-273) has an extensive list of primary reinforcers that includes stimuli such as grooming, pheromones, snakes, conspecific vocalizations and novelty. Butler and Harlow (1954) found that monkeys confined to an opaque cage would persistently act to open a hatch to get a view into the rest of the laboratory. Siqueland and DeLucia (1969) found that infants found sensory stimulation reinforcing.

The typical way of applying a reinforcer is by actively present it to the animal as a way of increasing the frequency of the desired behavior. While this certainly is one way of doing this, there are others. First, of course, a reinforcer can be used to lessen the frequency - it is in effect a punisher. Punishment is not a mirror effect to a reward. A rewarding reinforcer will increase the frequency of one specific behavior. A punisher, on the other hand, will lessen the frequency of a given behavior, but gives no indication as to what the animal should do instead.

Gray (1975) classified three different contingencies for reinforcers: presentation, which is simply presenting the reinforcer; termination, which is the opposite of ceasing to present the reinforcer; and omission, when an expected reinforcer fails to appear. This is orthogonal to reinforcers being rewarding or punishing. This is summarized in the following table:

	P(R)↑	P(R)↓
Presentation	Rew	Pun
Termination	Pun!	Rew!
Omission	$\overline{\text{Pun}}$	$\overline{\text{Rew}}$

Table 3.1: Reinforcement contingencies. After Gray (1975)

As table 3.1 shows, there are six different contingencies possible according to Gray (1975). Presentation of a rewarding reinforcer (Rew) is of course perceived as a reward, and presentation of a punishing reinforcer (Pun) is perceived as a punishment.

Termination occurs when a continuously presented reinforcer is terminated; this could be the termination of an annoying sound, or the closing of a window with a view. Of course, one can debate when an event is a presentation of a stimuli (sudden darkness) and when it is termination of a different one (termination of light). This discussion is, however, outside the scope of this text.

Omission is somewhat different than the other contingencies. It is the only contingency that explicitly acknowledges the role of the animal's expectations in learning. On the surface, there has been no event. What does happen is that the animal *expects* a reinforcer, and it fails to materialize. Extinction is the prototypical protocol for this.

The Crespi Effect

A vivid example of how complex the issue of reinforcement really is, is the so called Crespi-effect (Crespi, 1942), as described by Lieberman (1990, pp. 192–195). This experiment shows that the valuation of a reinforcer is not a simple function of its objective value to the animal.

Crespi had three groups of rats, being reinforced in a runway task, where the speed of moving through the runway is the indication of motivation for the subjects. First, one group would find 256 food pellets at the end of the runway, another group would find 16 pellets, and the last group would find just one pellet. Predictably, the 256-group would run the fastest, followed by the 16-group, and trailed by the 1-group that hardly ran at all.

In phase two, all the groups received the same number of pellets (16) when they arrived at the goal. The results were interesting: The 16-group stayed at the same level of running speed they were before - for them, nothing had happened after all.

The 1-group very rapidly increased their speed, and not only up to the level of the 16-group, but to a level *higher* than the 16-group. This is called an **elation** effect or positive contrast effect. Conversely, the 256-group decreased their running speed to a level below that of the 16-group; this is termed **depression**, or a negative contrast effect.

Note that Spence (1956) found a negative contrast effect in variations of Crespi's experiment, but did not find a corresponding positive contrast effect. Spence suggested that the reason was that Crespi had failed to fully condition his subjects prior to the shift.

What do we make of this? It shows us again that expectations play a large role for the subjective valuation of a reinforcer. A subject that expected 256 pellets will value 16 pellets far less than a subject expecting only one.

The Premack Principle

What turns out to be reinforcing at any time? Animals are not always hungry or thirsty, and sometimes, being social is felt as an arduous obligation rather than a pleasure.

Premack suggested his **Premack Principle**:

> "Given two responses, the opportunity to perform the higher probability response after the lower probability response will reinforce the lower probability response. However, the opportunity to perform the lower probability response after the higher probability response will not reinforce the higher probability response." (Premack, 1965)

He tested this principle in a celebrated experiment. He assembled a group of children that were allowed to eat candy and play pinball. Based on observational data, the children were divided into one group that preferred pinball over candy and one that preferred candy over pinball. For both groups, the candy was withheld unless they played pinball to 'earn' it. The group that preferred candy increased their pinball playing, while the group that preferred pinball did not.

In a similar experiment, Premack had two groups of rats, one water deprived, and the other not. The rats also had access to a running wheel, which in normal circumstances was preferred over drinking water. For the group with access to water, drinking water could be conditioned with running in the wheel as a reinforcer, while running in the wheel could not be associated with drinking water. For the water-deprived group, however, the result was the reverse.

This is somewhat surprising; candy would be expected to be a good reinforcer for other responses, and access to wheel running for rats is normally a good reinforcer as well. What Premack showed is that the incentive value of a reinforcer is relative to the response being conditioned, rather than having an absolute value for the animal.

Chapter 4

A Model of the Amygdala

The two-process model of learning described by Mowrer (1973) separates learning into first a stimulus-emotional system that evaluates incoming stimuli, and a second learning system that uses this evaluation as a reinforcer for stimulus-response learning. Among the advantages of this approach is that the motivation to respond and the response itself are cleanly separated (Rolls, 1986). We believe that the amygdalo-orbitofrontal system implements the evaluative functionality of such a system.

We have attempted to capture these features in a computational model suitable for comparisons between neurophysiological data and simulations. We hope that this approach will enable us to attain a clearer understanding both of the functions of the amygdala and of the limitations of the model; this would have been difficult to accomplish with a model that is not testable in simulation. The model described below is of course only a part of a complete emotional system; specifically, this model does not fully address context, configurational stimuli or higher-order conditioning. An extension with an accompanying hippocampal model and a revised orbitofrontal model will be described in the next chapter.

As discussed in chapter 1, a motivation for having an emotional system instead of only a goal generating system is that emotions are a more flexible way to generate motivations. Whereas a goal-directed system tends to specify what to be done, and in some cases how to do it, an emotional system only points out good or bad features, leaving the system to devise ways of coping with the situation.

As part of this investigation, we have implemented a computational model of the amygdala [1] and the orbitofrontal cortex, and tested it in simulation. This is not a detailed physiological model, even though it shares its larger-scale structure with that of the real amygdalo-orbitofrontal system; instead, the aim is to make use of neurophysiological data to construct a *functional* model of emotional processing as part of a general learning system. This system can in turn be used as a learning component in autonomous systems.

As we will see, this model does not in itself handle all the common learning protocols. It lacks both a trace mechanism as well as any notion of context. Nevertheless, it does show that a learning architecture on coarse neurophysiological structures will perform the basic functions needed for the structures in question. A more complete version of this model will be presented in the next chapter.

As we've seen in chapter 2, the amygdala is, despite its small size, a reasonably complex structure; multiple layers of interconnected areas, fed by data from large parts of the sensory cortices as well as subcortical areas, with multiple output areas and extensive connections with several other regulatory areas such as the hippocampus, hypothalamus and orbitofrontal areas. Add to that semi-independent areas for smell, and somewhat poorly understood interconnections between the areas across the hemispheres and it is easy to see that a comprehensive model of this area is no small undertaking. We do not attempt such a comprehensive model in this thesis.

While it would be very interesting to be able to understand and model every area, it is not feasible at this time; too little is known about the functionality of this structure to model it at great detail. Instead we attempt to extract the intended functionality of the amygdala at a higher level and implementing this in a computational model designed to replicate the functional role of this structure, rather than the precise means by which it accomplishes this. This work is thus situated in between efforts to understand neurological structures at a detailed level, and large-scale psychological models and artificial intelligence efforts aimed at replicating high-level aspects of behavior.

This means that, on one hand, no effort is made to model actual neuronal activity, nor do we take into consideration the undeniably important complexities of transmitter substance signaling. On the other hand, we do not try to model 'real world' phenomena either. Rather, we try to make a functional model of a few components involved in emotional learning. This technique can be used both to confirm or disprove hypotheses on the function of the rel-

[1] when we are talking about the 'amygdala' or other brain areas in regards to the model, we of course use it only as a label for a part of the model, and I am not implying any functional connection with the real amygdala other than when stated explicitly

evant areas, as well as give a clearer understanding of the mechanisms that are simulated.

As we have seen in chapter 2, the primary input area for the amygdala is the lateral nucleus. Data from this area is projected to the basal and basolateral nuclei, which seems responsible for generating emotional reinforcing signals for instrumental conditioning in other areas. It is also projected to the central amygdaloid nucleus, which is able to selectively generate responses to emotionally significant stimuli, as well as act as a reinforcer for instrumental conditioning in the basal ganglia.

4.1 The Model

The model is divided into two parts, conceptually corresponding to the amygdala and the orbital frontal cortex, respectively. Of course, these areas are complex, and we have not in any way attempted to capture all of their functionality. The amygdaloid part receives inputs from the thalamus and from cortical areas, while the orbital part receives inputs from the cortical areas and the amygdala only.

4.1.1 Amygdala

This part of the model attempts to capture a few aspects of learning in the amygdala. It is very much like the Rescorla-Wagner model discussed in chapter 2, with nodes learning associations controlled by a reinforcement signal defined as a difference between expected and actual reinforcement.

There are studies that imply that conditioning in the Amygdala is permanent, or at least very hard to reverse (Sanghera et al., 1979; Rolls, 1992; Wilson and Rolls, 1993). The rationale behind this is that once learned, a reaction – especially a negative one – is so expensive to retest that it pays to assume this negative association is valid everywhere unless definite proof otherwise has been established. Thus, the current model is also one-way; the model learns emotional relations, but does not unlearn them.

This model (the lower part of figure 4.1) includes the amygdala proper and connections from the sensory cortices directly and the thalamus – we assume both connections are represented as the input stimuli S in the model. In the

next section, we will add an orbitofrontal cortex to this model. There is one A node for every stimulus S. There is one output node in common for all outputs of the model, called E above. The E node simply sums the outputs from the A nodes. The result is the output from the model.

For each A node, there is a plastic connection weight V. Any input is multiplied with this weight to become the output from the node.

$$E = \sum_i A_i,$$

$$A_i = S_i V_i$$

The connection weights V_i are adjusted proportionally to the difference between the reinforcer and the activation of the A nodes. The α parameter is a standard learning rate parameter, settable between 0 (no learning) and 1 (instant adaptation). In practice, it is usually set at a fairly low value; here we keep it at $\alpha = 0.2$.[2]

$$\Delta V_i = \alpha S_i [R - \sum_j A_j]^+$$

This is an instance of a simple associative learning system, very much like the Rescorla-Wagner model of learning (Rescorla and Wagner, 1972). The real difference is in the fact that this weight-adjusting rule is monotonic, i.e. the weights V can not decrease. This may at first seem like a fairly substantial drawback; however, as seen previously, there are good reasons for this design choice. Once an emotional reaction is learned, it should be permanent. It is the task of the orbitofrontal cortex to inhibit this reaction when it is inappropriate. As seen in chapter 2, there is experimental evidence which supports that this is how the amygdala works in this respect.

The learning rate for the system is a function of the α parameter; the difference between the strength of the reinforcer R and the current output of the A nodes; and the strength of the stimulus signal S_i. The stronger the stimulus, and the larger a difference there is between reinforcer and output, the faster the learning is. This is modulated with the α parameter.

The reward signal is 'generic', in the sense that it represents the various reinforcing inputs from different areas; as we saw in chapter 2, this signal can

[2] the notation $[x]^+$ in the equation is to be read as: $\max(0, x)$, or that x is to remain nonnegative.

originate in the thalamus, the hypothalamus, or parts of the basal ganglia. Here, all those various sources have been abstracted into one input.

The existence of only one output node is of course not a realistic assumption for this model. A more realistic version of this model would have a number of output nodes which enables the model to selectively evaluate the individual stimuli. Also, these nodes would be capable of being individually inhibited by the orbitofrontal part. With this organization, the system would be able to selectively influence attention as well as allowing for a finer grained learning of actions. We are here not interested in such specifics; just as with the reinforcing signal, we we abstract the outputs into one simple indicator.

What can this model do? By itself it can learn simple associations between a stimulus and a reinforcer. It can not extinguish any learned associations, as we deliberately built this to make it impossible. It is still able to handle a couple of common learning protocols, such as acquisition and blocking.

4.1.2 The Orbitofrontal Cortex

The second part of the system is a model of the orbitofrontal cortex. This area is, as we have seen, responsible for inhibiting inappropriate reactions from other areas, including the amygdala.

The OFC needs much the same data as the amygdala: stimuli and reinforcer. In addition, this design presupposes a data path from the amygdala to the OFC indicating the current emotional evaluation of the amygdala to the present stimuli. This is so the OFC will have a target to inhibit.

The OFC model is similar to the amygdala model; it also adapts its output according to the sensory data S and the reinforcer R. An added wrinkle is that the output is used to inhibit the amygdala, so the OFC is using the output of the amygdala as another parameter to determine the level at which the inhibition should be applied. Note that we are not arguing that it is inhibitory at the synaptic level. Instead, what we are saying is that the functional *effect* of this connection is such that it acts to block the response of the amygdala.

The mechanism is the same as for the amygdala:

$$E_o = \sum_i O_i,$$

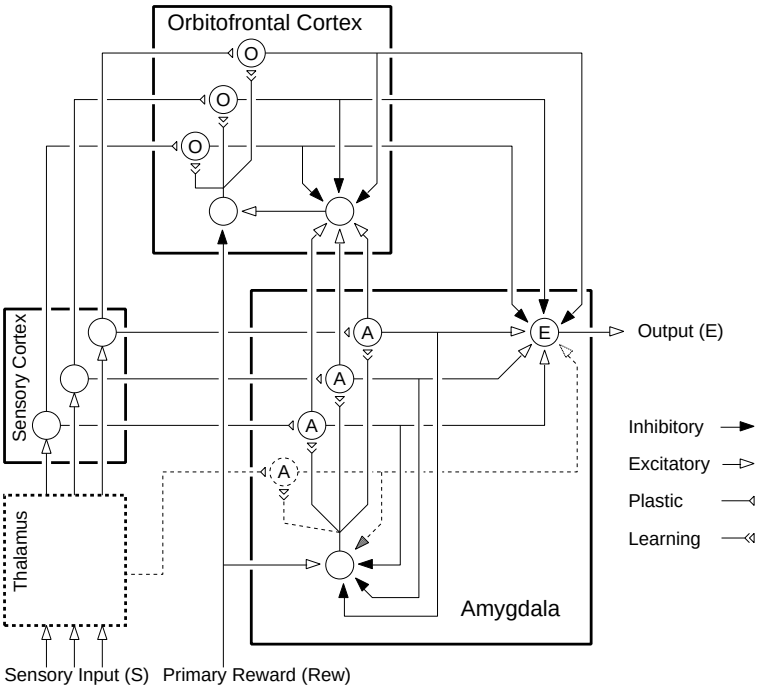

Figure 4.1: A graphical depiction of the model. At the top is the orbitofrontal part (here without an external context), at the bottom right is the amygdaloid part and at left are the thalamic and sensory-cortical modules. The thalamic and sensory-cortical parts are place-holders in this version of the model, shown to illustrate their place. The sensory inputs S enter the thalamic part, where a thalamic input to the amygdala is calculated as the maximum over all inputs. A primary reward signal Rew enters both the amygdaloid and orbitofrontal parts. The output from the model is a scalar value E that represents the emotional evaluation of the incoming stimuli.

$$O_i = S_i W_i$$

The O nodes behave analogously to the amygdala A nodes, with a connection weight W applied to the input signal to create an output.

The connection weights W_i are updated as a function of the input and the internal reinforcer for the OFC:

$$\Delta W_i = \beta S_i R_o$$

Unlike the amygdala model, this learning is not constrained to be monotonic. The β parameter is, just as the α parameter for the amygdala, a learning rate parameter, set at $\beta = 0.4$, to reflect a faster learning rate.

The internal reinforcer R_o is however more complex than the simple function of the reinforcer and output we had for the amygdala:

$$R_o = \begin{cases} [\sum_i A_i - R]^+ - \sum_i O_i & \text{if } R \neq 0, \\ [\sum_i A_i - \sum_i O_i]^+ & \text{otherwise} \end{cases}$$

The internal reinforcer does quite a bit more work for us than the one in the amygdala. What this expression says, is this: If there is a reinforcing signal ($R \neq 0$), then the reinforcer R_o is the discrepancy between the amygdala output and the reinforcer R (that is, the error between expected and real reinforcement) minus the OFC output. So, the OFC output is adjusted to minimize the discrepancy of the amygdala output and the reinforcer, which is exactly what we want.

When the reinforcer R is not present, on the other hand, the system works a little differently. The internal reinforcer R_o then becomes the discrepancy between the amygdala output and the OFC output – if it is positive, in other words – if the amygdala is reacting too strongly. If the OFC inhibition is stronger than the amygdala output, on the other hand, there is nothing wrong with that, as the inhibition almost certainly is proportioned for a higher reaction level of the amygdala for the present stimuli.

4.1.3 The Combined System

The combined system works at two levels: the amygdala learns to predict and react to a given reinforcer. This subsystem can never unlearn a connection; once learned, it is permanent, giving the system the ability to retain emotional connections for as long as necessary. The orbitofrontal system tracks mismatches between the base systems predictions and the actual received reinforcer and learns to inhibit the system output in proportion to the mismatch.

These subsystems receive partially different inputs. The amygdala receives stimuli (that conceptually can be seen as coming from both the thalamus and the sensory cortices) and an abstract reinforcer. The orbitofrontal cortex receives much the same inputs; but in addition, it receives a projection from the

amygdala with its evaluation of the stimuli. The amygdala is then inhibited by the OFC as appropriate in the current situation.

The orbitofrontal system currently receives the same input as does the amygdaloid system (this will be changed in chapter 5). These inputs work as a simple substitute for a proper context representation. In the next chapter, however, this system will receive a context representation that will enable the model to handle contextual cues properly.

Each subsystem is very simple. The power comes not from the learning mechanisms within each component, but from the way they are interconnected.

Opponent Process Theory

One alternative model of emotional conditioning is the opponent process theory of Solomon (1980). This views emotional learning as consisting of two opposed processes: a quick reaction to the event (called the primary affective response), and a slower, opposite reaction (the affective after-reaction) that is triggered by the primary response and is gradually strengthened to cancel out the immediate reaction. Since both onset and offset of the canceling reaction is slower than the immediate reaction, this will lead to a quick surge of emotional reaction at the onset of the stimulus, followed by depression at its offset. This model has had some success in describing emotional and motivational states.

Superficially, our model resembles this idea: we have a excitatory system that learns to react to stimuli, and a second, inhibitory system that suppresses this excitation as needed. This resemblance is illusory, however. The inhibitory system in the OFC is tightly coupled to the level of the excitatory reaction in the amygdala, and is as fast-acting as the excitatory system. That said, with some changes to the inhibitory conditioning (to loosen the tight temporal coupling between the systems), and with suitably chosen learning parameters, this model is likely to be able to model at least some of the effects seen with opponent-process theory as well.

4.2 Simulations

We have run a set of simulations to verify some assumptions about the workings of this model. The basic features we have tested are acquisition-

extinction-reacquisition, simple blocking and conditioned inhibition. We also bring up positive patterning and show that it can't be accurately performed by this model. These represent common features of emotional learning and are often tested directly or indirectly in animal experiments.

While these simulations feature abstract conditioning protocols, they do represent aspects of general Pavlovian conditioning and gives a good indication as to how well the model can explain the conditioning that does take place in the relevant areas.

The protocols lack any timing-related information as presented here. Our model has no timing functionality; one stimulus is a one-tick pulse. This is a shortcoming of the model that is addressed in later sections, but for the protocols it means that we can present them in this abstract manner. To simplify the presentations further, we keep the intertrial interval very short. As the model does not handle intertrial effects this does not alter the results and makes the figures clearer.

Acquisition of emotional reactions are usually much faster than other types of conditioning. Emotional conditioning has been reported in as little as eight pairings of a visual cue used as CS and electrical shock (LaBar et al., 1998). In experiments with rats using an auditory stimulus four pairings of the CS and US was sufficient for the sound to elicit freezing behavior (Morgan and LeDoux, 1999).

The time course of extinction is similar to acquisition, but depends, like acquisition, on the exact experimental procedure and the stimuli used. In general, reacquisition is faster than initial learning (Smith and Gormezano, 1965). With repeated acquisition and extinction, relearning becomes faster. This is known as a *savings effect*.

4.2.1 Acquisition

Acquisition and extinction is a basic learning experiment, where the model is expected to associate a stimulus with a reward/reinforcer, disassociate the stimulus once the reinforcer is absent, then re-associate them again. This represents a minimal functionality of any associative learning model. This is the protocol for the basic function of the amygdala, where it forms associations between emotionally significant stimuli and neutral stimuli predicting its onset.

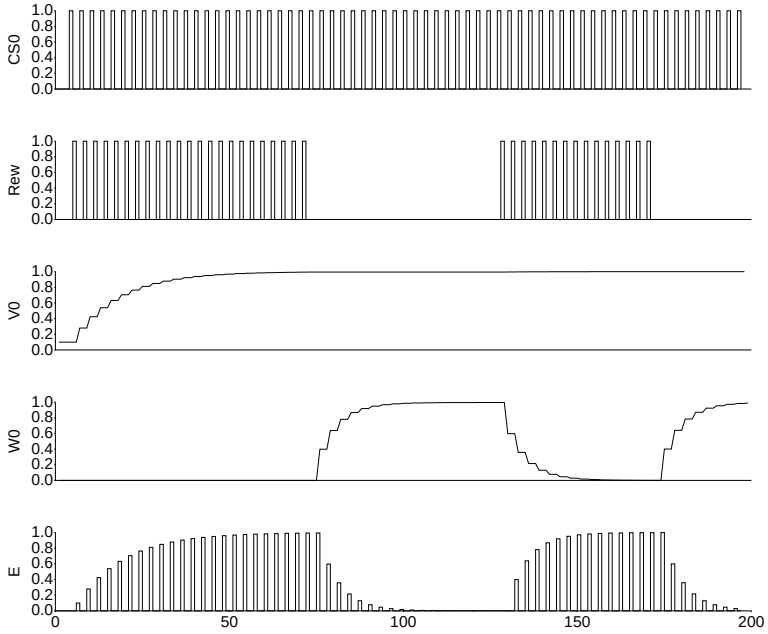

Figure 4.2: The result of acquisition, extinction and reacquisition. From top to to bottom, the graphs are: CS0, the stimulus input; Rew, the reinforcing signal; V0, the amygdaloid connection weight for the stimulus; W0, the orbitofrontal connection weight for the stimulus; and E, the output of the model (smoothed); The learning parameters are $\alpha = 0.2$ and $\beta = 0.4$. The figure shows that the amygdala does not extinguish its output; instead, the orbitofrontal part takes responsibility for inhibition of the output when no reinforcer is present. As the model ignores intertrial interval length, we can present protocols in this compressed manner without having a realistic ITI.

The protocol for this simulation is:

Phase 1:
    CS0 + US
    CS0
Phase 2:
    CS0 + US
    CS0

In figure 4.2 we see the input CS0, the reinforcement signal (Rew) and the output (E). The acquisition-extinction cycle is repeated to see how the system reacts during reacquisition. In this figure, we have included the connection weights for both the amygdaloid and orbitofrontal parts of the model. The V0

connection weight is the amygdala learning weight, while W0 is the weight for the orbitofrontal node.

As seen in the figure, the system manages to learn this simple association well; the output tracks the reinforcer without any problems. During acquisition, only the amygdala reacts; the orbitofrontal part has no reason to inhibit anything, and stays dormant. When the reinforcer disappears for extinction, the amygdaloid weight is not affected; instead the orbitofrontal weight W0 increases to inhibit the output. When the reinforcer reappears, W0 decreases again, allowing the amygdala to express the previously learned association. During this cycle, the learned response from the amygdala stays high.

The output E increases to its full level faster on subsequent trials than it did the first time. This savings effect is well established in the literature; see cf. Mackintosh (1983).

Acquisition of emotional reactions gives qualitatively similar results as animal experiments, as seen above. We do not, however try to do a *quantitative* comparison, even though it would be possible to adjust the learning constants to match the results from animal experiments.

4.2.2 Blocking

When an emotional reaction has been associated with an eliciting stimulus or context, subsequent conditioning to another stimulus is substantially impaired (Kamin, 1968; Mackintosh, 1983; McNish et al., 2000). The original learning *blocks* the acquisition of further associations. As described in chapter 3, blocking can be explained by a principle of parsimony. If a stimulus already predicts the reinforcer, other stimuli are not needed for prediction, and are thus not conditioned to the reinforcer.

In the simulation we show the ability of the model not to associate a stimulus with the reinforcer if there is already an established association that can explain the contingency. In other variants, we also show what happens when the stimulus is able to only partially account for the reinforcer.

A blocking schedule is run in three phases: first associate CS0 with the reinforcer, then present both CS0 *and* CS1 together with the reinforcement, and last, test CS1 to see whether it has been associated with the reinforcer. There should be no response to CS1. This is consistent with experimental data, and is explained by the principle of parsimony: do not associate a reinforcer with several stimuli, when one is enough to explain the association.

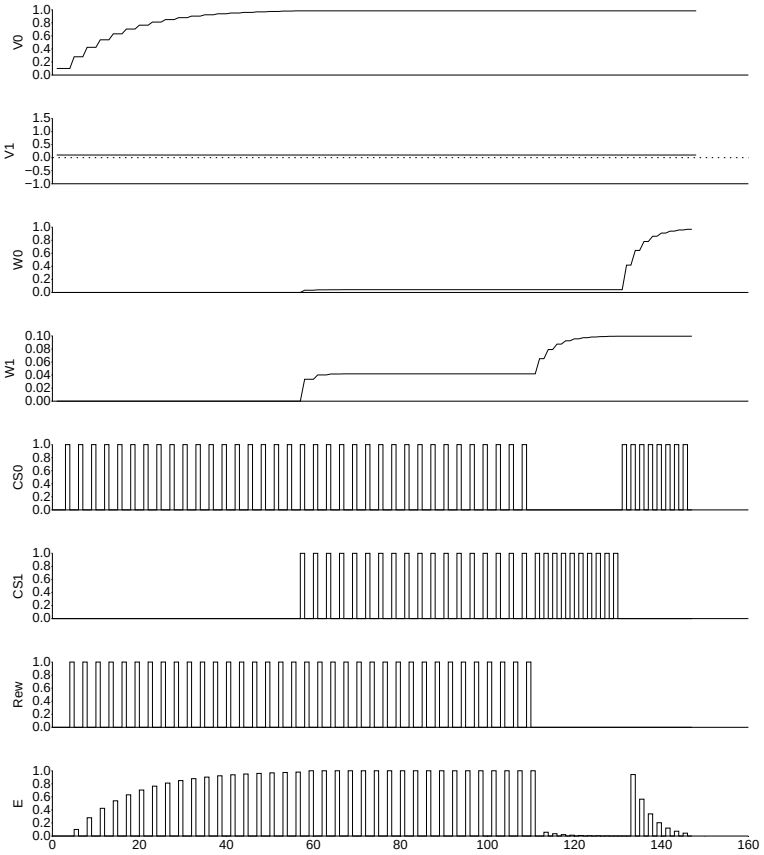

Figure 4.3: The result of a blocking experiment. CS0 and CS1 are stimulus inputs, Rew is the reward and E is the output from the model. The weights V0 and V1 are the connection weights for the input nodes, and W0 and W1 are the connection weights for the orbitofrontal inhibition. Looking at V1, it is clear that blocking is taking place. As in the acquisition simulation, α and β are set at 0.2 and 0.4, respectively.

The protocol for basic blocking looks like this:

Phase 1:
CS0 + US

Phase 2:
CS0 & CS1 + US

Test:
CS1 $\rightarrow$ no-CR

If we look at figure 4.3, we see that this is indeed what happens. CS0 takes up the sole responsibility for reacting to the stimulus. When we pair CS0 with CS1 and the reinforcer Rew, we get no learning for CS1, as seen for its V1 learning weight. We do get a small reaction from the inhibitory weights W0 and W1, as the default values for the learning weights V aren't zero, but a small positive number (0.1 in all our simulations). The evaluation for CS0 and CS1 thus add up to more than the reinforcer, and both are slightly inhibited.

When we present CS1 alone, we see (as expected) that it has not been conditioned to the reinforcer. the inhibition rises slightly again, to fully inhibit the default response that is already partly inhibited from the earlier phase.

As we show CS0 by itself, without reinforcement, we see that indeed, it is fully conditioned, though the response falls rapidly as it is extinguished. It has thus been unaffected by the testing of CS1.

Unblocking by US Intensity Change

We can vary the protocol slightly to see what would happen if the reinforcer changes with the added stimulus. According to established theories, the new stimulus should take up the 'slack' from the increased reinforcement.

Phase 1:
CS0 + US(0.7)

Phase 2:
CS0 & CS1 + US(1.0)

Test:
CS1 → CR(weak)

We see here a similar behavior to the 'pure' blocking simulation above. CS0 is conditioned to respond at a level proportional to the reinforcer. However, when CS1 is introduced, together with an increase of the reinforcer, it shares the responsibility of responding to the added reinforcement with CS0. In effect, what we see is that CS0 and Rew specify a baseline, and the stimuli react only to the differential of this baseline.

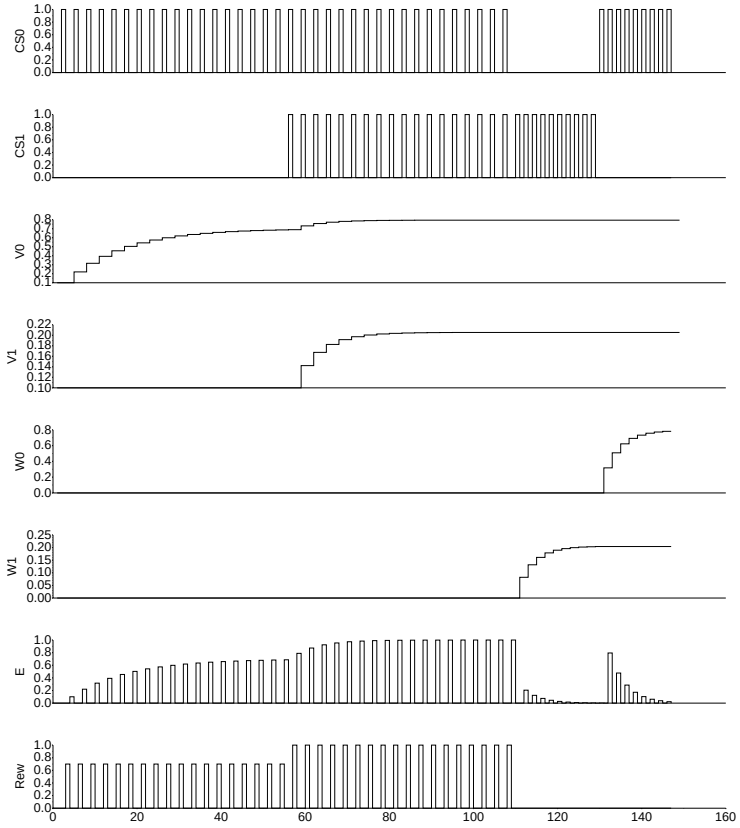

Figure 4.4: Unblocking by US intensity change. The difference here is that the reinforcer is set at a lower value (Rew=0.7) during conditioning of CS0, and raised to 1.0 simultaneously with the introduction of CS1 as a second stimulus.

Unblocking by CS Intensity Change

Instead of increasing the US strength when we add another CS, we can decrease the strength of the original CS. The results should be the same: the new CS should get conditioned slightly to the US.

Phase 1:
    CS0 + US
Phase 2:
    CS0(0.5) & CS1 + US

Test:
    CS1 → CR(weak)

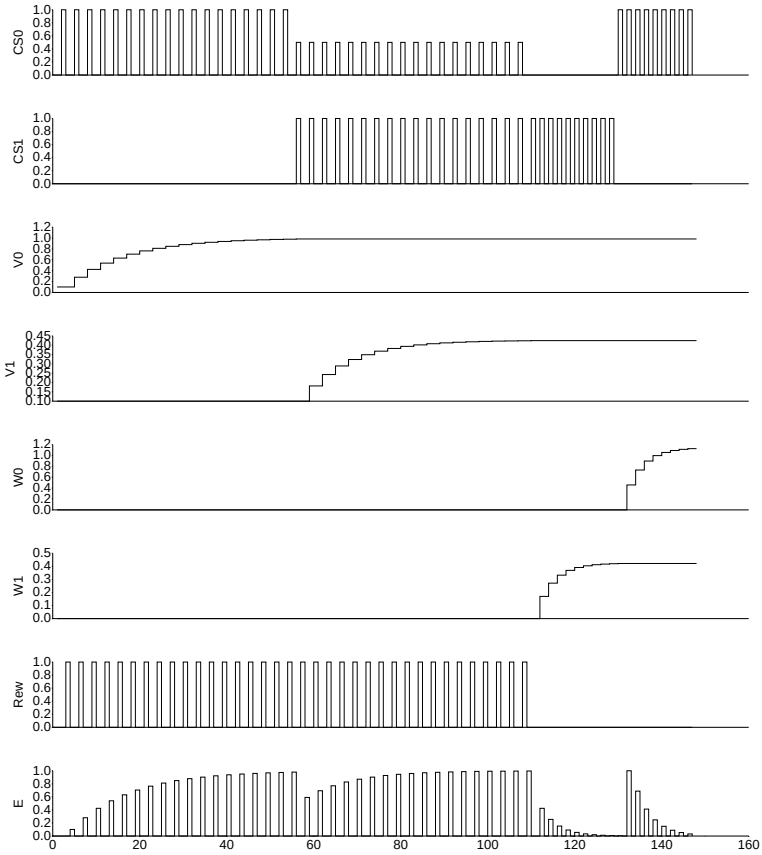

Figure 4.5: Unblocking by CS intensity change. CS0 is conditioned to the US. When CS0 is presented together with CS1, the strength of CS0 is halved. This enables CS1 to partially condition to the US.

We condition CS0 to the US. Then we pull down CS0 strength to half its original value, and introduce a second stimulus CS1 that is shown simultaneously with CS0. As we see in figure 4.5, CS1 picks up the 'lost' conditioning from the weakening of CS0. We test both CS, and indeed, CS1 has been conditioned to half it's strength, while CS0 remains at full strength; this is of course clear from the connection weights as well.

Why did not CS0 pick up some of the conditioning, the same way it did for unblocking by US intensity change? The reason here is that the connection weights are capped at 1.0, preventing CS0 from taking up any of the conditioning when its signal strength decreases. Had we run the same protocol, but started with a lower signal strength for CS0 – say 0.7 rather than 1.0 - we

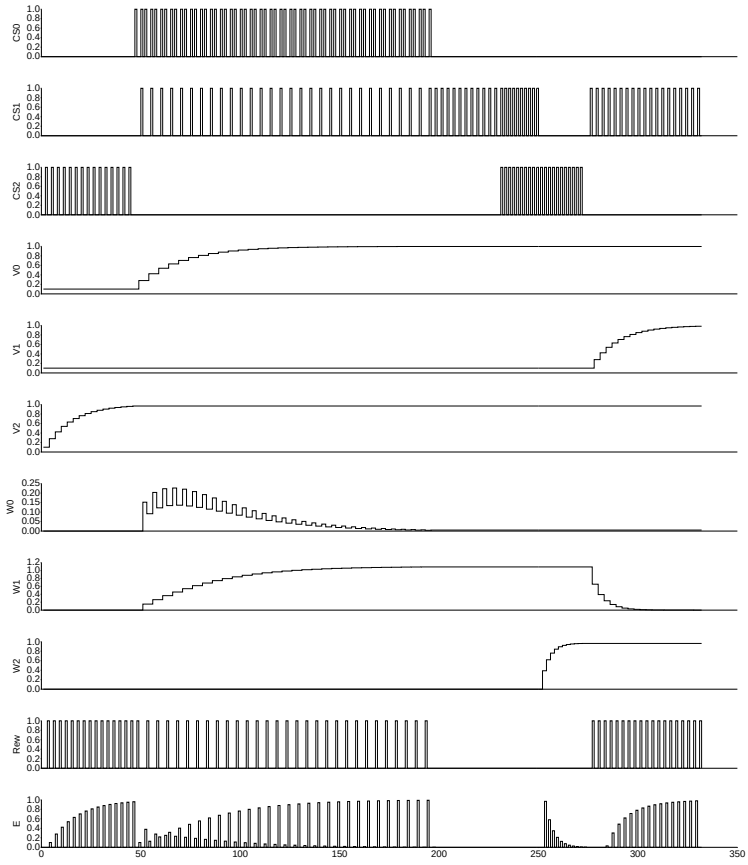

Figure 4.6: The result of conditioned inhibition on the model. CS0, CS1 and CS2 are the stimulus inputs, Rew is the US and E is the output from the model. α is 0.2 and β is 0.4 .

would have seen a behavior more like that in the previous unblocking experiment. This exposes a subtle problem with computational simulations of this type, and is discussed further at the end of this chapter.

4.2.3 Conditioned Inhibition

In a conditioned inhibition schedule, the aim is to show that inhibition is an active process, not merely a decrease in associative strength. A stimulus can be given inhibitory properties, that can actively inhibit the response of other

stimuli. Again, there is experimental evidence that this effect is present in animals (Mackintosh, 1983).

The schedule for conditioned inhibition is somewhat involved. We want to establish an inhibitory association with a stimulus, then test it with another stimulus that already has an association with the reinforcer. Creating an inhibitory association can be done by explicitly omitting the expected reinforcer whenever the stimulus is present.

Prepare CS2:
    CS2 + US

Condition:
    CS0 + US
    CS0 & CS1

Test:
    CS2 & CS1 → no-CR
    CS2 → CR

First, associate CS2 with the reinforcer; this is the stimulus that will be used for testing. Next, alternate CS0 with reinforcement and CS0 & CS1 with no reinforcement. This should give CS1 inhibitory properties as CS0 predicts the presence of the reinforcer. To test the result, CS2 (the test stimulus) and CS1 are presented together, and should give little or no response. Last, we present CS2 alone, to show that it has not been affected by the inhibitory associating stage.

The results are as expected: CS1 and CS2 give only a small, immediately decaying response, while CS2 alone gives a satisfactory response. This result is due to the fact that the orbitofrontal part actively learns to inhibit responses in the presence of CS1, rather than the amygdaloid part unlearning anything.

Conditioned Inhibition by Stimulus Change

We can run a variant of conditioned inhibition as well. Instead of varying the intensity of the US, we can vary the intensity of CS1 instead:

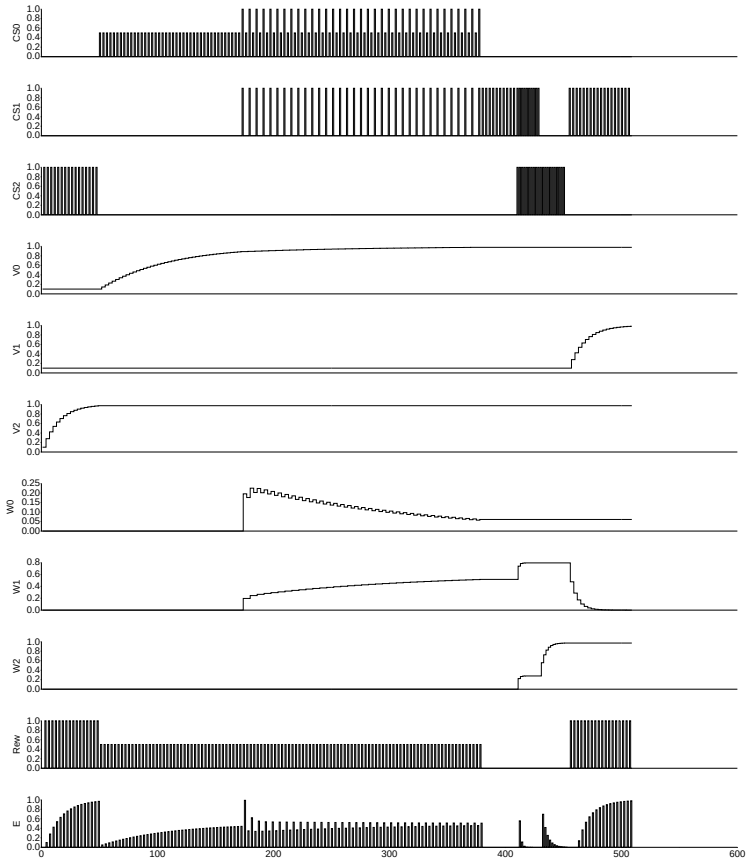

Figure 4.7: A variation of conditioned inhibition, where the CS0 stimulus strength is varied, rather than US. CS0, CS1 and CS2 are the stimulus inputs, Rew is the reward and E is the output from the model. α is 0.2 and β is 0.4 .

Prepare CS2:
CS2 + US

Prepare CS1:
CS0(0.5) + US(0.5)

Condition:
CS0(0.5) + US(0.5)
CS0 & CS1 + US(0.5)

Test:
CS2 & CS1 $\rightarrow$ CR(weak) CS2 $\rightarrow$ CR(stronger)

First, we condition the control stimulus CS2 so we later on can see if the stimulus CS1 really has become inhibitory. Then we condition CS0 at a lower strength to the US (also at a comfortably low strength), also in preparation for the attempted inhibition. We then alternate pairings of CS0 at its original strength and the US, with pairings of CS0 at full strength and CS1 at full strength and with US at its usual strength.

What happens is that the increased strength of CS0 predicts a stronger rewarding US. When that stronger US fails to appear, both CS0 and CS1 gain inhibitory properties. The inhibition of CS0 gradually disappears (see the 'W0' row in figure 4.7), while CS1 gradually becomes inhibitory enough to offset this lack of US strength. In effect, the system blames CS1 for the faulty prediction that US strength would increase with CS0 strength.

When we test this inhibition, CS1 is able to partially inhibit CS2 – it is partial, because the inhibitory strength of CS1 is only enough to inhibit CS0 properly, but not enough to fully inhibit a maximum strength CS2.

Discrimination Learning

A basic discrimination learning protocol is used to test the possibility of discrimination between stimuli sharing similar features. In its simplest form, they are presented alternately, with one stimulus reinforced and the other not (Pavlov, 1927). The animal is expected to learn to respond to one stimulus and not the other, despite the similarities between them.

An added complication is to run discrimination reversal. After the animal has learned to discriminate between the stimuli, the reinforcement is reversed, so that the first stimuli is no longer reinforced, while the other one is. Gradually, it is expected that the animal will become faster at this reversal.

Similarity for stimuli is in this model handled by having several stimuli represent the stimulus components of the actual stimuli. So our first stimulus will consist of components CS0 and CS1, and our second stimulus will be CS1 and CS2. The CS1 component is common to both.

The protocol is:

Phase 1:
    CS0 & CS1 + US
    CS1 & CS2

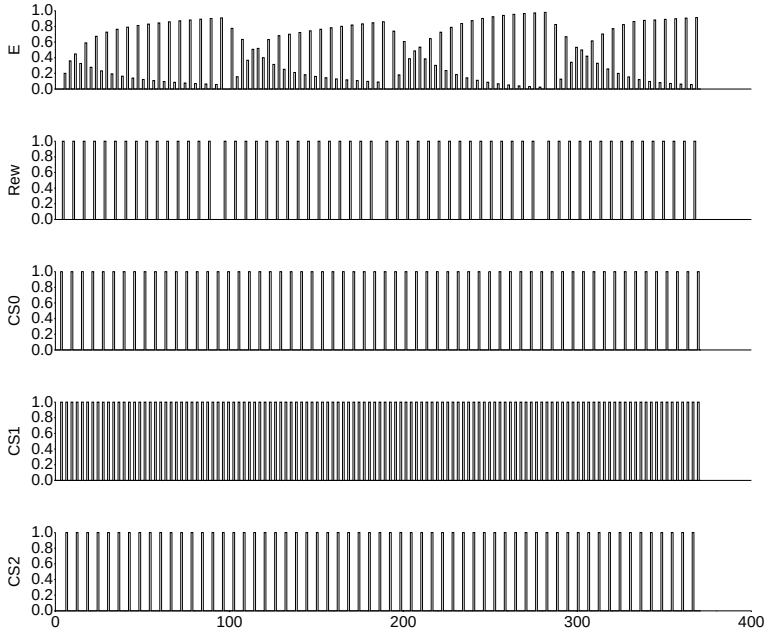

Figure 4.8: A discrimination experiment. Two stimuli, each consisting of two stimuli, one of which is common to them both, are presented. One is reinforced and the other is not. Gradually, the generalization disappears and they are discriminated. The protocol is run four times.

Phase 2:

CS0 & CS1

CS1 & CS2 + US

Each phase consists of multiple alternations. In figure 4.8, the phases are repeated once. We see that the model can handle both discrimination and discrimination reversal. Also, though the effect is subtle in the figure, the speed of reversal increases somewhat after the first reversal. This is due to the amygdala having been conditioned to all stimulus components, leaving the reversal entirely up to the OFC.

4.2.4 Positive Patterning

Positive patterning is a test to see whether the system can condition to the combination of two stimuli, while refraining from reacting to either stimuli alone (for the mathematically or logically inclined, this is an "AND" function).

We bring this protocol up, as it at first *appears* to work using this model, when in reality the slight effect we will see is an artefact, and it is a mildly interesting illustration of the function of this model. In chapter 5, we will see this protocol being simulated properly.

The protocol is:

Condition:
- CS0 + US
- CS1 + US
- CS0 & CS1

Test:
- CS0 → no-CR
- CS1 → no-CR
- CS0 & CS1 → CR

As described in chapter 3, the 'task' for the system is to react to the occurrence to both stimuli appearing simultaneously, but not to either alone. In terms of emotional reactions, this could be not reacting when seeing a moving car or a child playing in the street, but reacting strongly when both occur at the same time. Take a look at figure 4.9. At first sight, it might look like some discrimination is taking place; after all, the response is greater when both stimuli are presented together (and they are reinforced) than with either stimuli alone.

This is however not really the case. In this simulation, we have run both stimuli alone twice between each pairing of the stimuli to better show what happens. We can see that the non-reinforced stimuli gives a lower response the second time they are presented alone each sequence. That is of course due to the fact that they are inhibited bu the OFC in the absence of a reinforcer. When they appear together – and together with the reinforcer – they are disinhibited again, giving rise to an effect that is superficially similar to real positive patterning.

4.2.5 Lesions

Here, we briefly discuss the results of 'lesioning' parts of the model. There are several points at which a connection could be broken: The stimulus path to the amygdala or the OFC; the reinforcement signal to either area; the current prediction of the amygdala to the OFC; or the orbitofrontal inhibition to the amygdala.

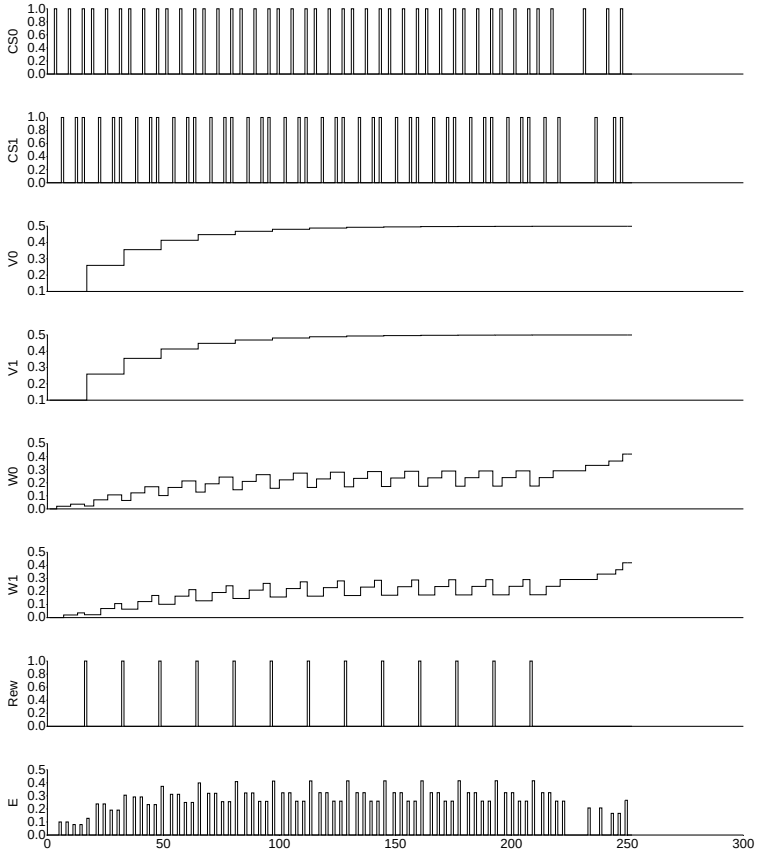

Figure 4.9: Positive patterning on the model. CS0 and CS1 are presented alone and together, with reinforcement only when presented simultaneously. The result of this discrimination protocol is shown at the bottom row. As the orbitofrontal connection weights W0 and W1 show, the apparent discrimination is only an effect of repeated extinction, rather than true discrimination.

Disabling the orbitofrontal part results in a system that can learn emotional predictions, but can not extinguish them or inhibit them depending on the current context. This is consistent with the observations of Shimamura (1995) and Kolb and Whishaw (1990) on patients with frontal lesions.

With the current version of the model, the other 'lesions' are fairly uninteresting. A more thorough review of lesioning effects on the full model is available in section 6.1.1.

4.3 Summary and Discussion

The model presented in this chapter has several attractive characteristics. It can handle some common emotional conditioning experiments and it is easy to implement and use as part of a larger system.

As it stands here, this model is not a complete learning system. As it is an emotional evaluator of stimuli, it needs several components to handle "real" learning tasks. The two most important missing parts are a context system and some form of motor system that can use the output of this model.

To use the model in an artificial system, it should be coupled with an instrumental learning system such as a two-process learning architecture, where the evaluation of a stimulus and the choice of action to be taken as a result of this evaluation is clearly separated. Once the action selection system has learned an appropriate response, the emotional evaluation is no longer needed; this seems to be in agreement with experimental data (Poremba and Gabriel, 1999). This also means that if the learned response starts to fail (i.e. it is now inappropriate), the evaluative system will still retain its learned evaluation, which immediately can be used to start to learn another, more appropriate response.

We know that the basolateral amygdala is implicated in instrumental conditioning in the basal ganglia, as reviewed in chapter 2. The outputs of this model would be the reinforcing signal sent to the premotor areas. The present system would be the classical component, while the basal ganglia would use instrumental conditioning.

If we wanted to make a motor learning system, we would use the output of this model as a reinforcing signal for learning motor sequences. There are any number of models capable of being used, such as Q-learning (Watkins, 1992), HQ (Morén, 1998) and TD(λ) (Sutton and Barto, 1998).

For an overview of the amygdalo-orbitofrontal system as a part of a larger motivational system, see Morén and Balkenius (2000b), or see chapter 1.

A recurring question is the origin of the reinforcing signal. As the model currently stands, it appears from nowhere (or rather, from the simulator) to control the behavior of the model. We believe this signal is the result of a reaction to the presentation of a primary stimulus or an emotionally charged stimulus. First order conditioning would be the association with a primary stimulus (a stimulus that has an intrinsic emotional charge), while second order conditioning would be the association with previously learned stimuli.

An interpretation of primary stimuli would be that they use the same pathways as other stimuli, but are hardwired to produce a reaction at the outset. This would imply that first-order and second-order conditioning use the exact same mechanism.

As previously stated, the current version has only one output node that sums the output from the evaluating A nodes. This node can only signal a general emotional evaluation of all the input stimuli and can thus not give more than one kind of evaluation. A more complete system would be able to give both positive and negative evaluations, and give them selectively to different stimuli. This is of course not the case in a real emotional system; even disregarding the fact that each output area projects to a number of other sites, the connections between areas are far more complex than a single signal. The ability to selectively evaluate stimuli would at least in part be solved by an attention model that would preselect the stimuli the system would attend to in the first place.

What we do here is to again bring together several complex interconnections into one symbolic output representing the reaction of the whole system. There are very good reasons for making this aspect of the model more realistic; a model designed with a finer-grained output system can more easily be meshed with models of the areas to which it would connect, and could be the basis of a system intended to use the technology to do real work. The focus here is on the model as a partial description of the operation of amygdala, and as such, as simple output indicator is sufficient. Of course, if one wants to use this model simply as one part of a two-process system of instrumental learning then having a single output is sufficient.

As we saw in the experiment on unblocking with CS intensity change (section 4.2.2) in this chapter, there is a problem with the kind of simulations done here. In that experiment, we lowered the strength of a previously conditioned CS to allow for a second CS to also attain a measure of conditioning. The problem lies in the fact that CS0 in that experiment did not pick up any further conditioning as a result of this change. Had we run this simulation with all strength levels halved, say, we would have had further conditioning to both CS0 and CS1, but in this case, the connection weight for CS0 was already at its maximum, and no further conditioning could take place.

Now, it is reasonable to have a limit on connection weights; it is quite obvious that signal levels in the real brain must be bounded (they in fact have a number of limitations; see Rieke et al. (1997) for a fascinating account). The problem really lies in the fact that we tend to simulate signal strengths at their maximum possible level – a level that would be expected to rarely be attained in a 'natural' setting. Of course, we do not know where the 'normal' level

of stimuli is placed in the internal scales of the experimental subjects, making it difficult to do more than educated guesses regarding this issue. In this thesis, this subject occurs only rarely, so we sidestep the issue by ignoring it whenever it does not matter.

4.3.1 Other Models

Armony et al. (1997) have developed a model of amygdala, thalamic and sensory cortex interaction for fear processing, focused on auditory stimuli. Their architecture is based on a similar view of the system as we have, which of course is not surprising, considering that both are based on neurophysiological data.

Their model is based on modules of homogeneous nonlinear units with lateral inhibition. The inhibitory connections enables the units to organize into receptive fields over their input. Learning within a module is done by adding the US signal as a static input, to further increase the responsiveness of the most responsive units through the inhibitory connections. Learning was enabled in one part of the thalamic network and in the amygdala. In addition, another parallel thalamic module, as well as a module modeling the auditory cortex is present. The output is the sum of the node activation in the amygdala.

They show that the model can generate generalization gradients over the 'auditory' input (a collection of ten nodes), and that it can shift the peak of the gradient in response to conditioning, in a way reminiscent of experimental data on rats. In addition, they show that lesioning of the auditory module does not increase the generalization gradient, a prediction that was subsequently confirmed in an experimental setting.

The within-module architecture is sound, and more realistic than the architecture used in the model presented in this thesis. There are two problems with the Armony model in our view, however. Armony et. al. focus on the cellular level, leaving the question of the larger scale architecture unanswered. While the individual nodes certainly are more realistic than in our model – and they are investigating more neuron-like architectures – they make the simplification that each area is just identical homogeneous collections of these nodes. Second, the model does not attempt to build a general architecture for emotional conditioning, but is interested instead in the neuronal changes and precise response patterns for cortical-thalamic sensory integration.

The features of this model are situated at a lower level than the model presented in this chapter, and it would be interesting to attempt to incorporate these features into our model.

Another model of some interest is the amygdala model described by Rolls and Treves (1998, pp. 146-151). Like the previous model, it is neurologically inspired. Unlike the Armony-LeDoux model, it concerns itself only with the function of the amygdala itself. Also unlike that model, while possibly implementable, it is not described in sufficient detail to make a rigorous implementation, and no simulation data has been presented.

In the authors' view, the amygdala can be regarded as a simple pattern associator. Primary reinforcers are non-modifiable signals that travel straight through, while secondary reinforcers have modifiable connections to these signal pathways. A Hebbian learning rule allows the secondary reinforcers to activate these pathways.

The model as presented in (Rolls and Treves, 1998) has the benefit of simplicity. The purpose of the model is to present the perceived functionality of the amygdala, rather than being testable in simulation. In fact, running the model in simulation will likely not accomplish very much unless it is integrated as a part of a larger system, much in the same way as our model.

4.3.2 Previous Versions

In earlier versions of this model (Morén and Balkenius, 2000a), there was a thalamic input, th, formed as the maximum over all stimulus inputs (the dotted parts in figure4.1). This connection is known to exist, and represents a shortcut directly from the sensory thalamus to the amygdala, bypassing the sensory cortex (Fendt and Fanselow, 1999; LeDoux, 1996). At this time, we do not have a model of the thalamus that is able to produce reasonable inputs for this pathway, so this connection is not present.

Two factors speak in favor of this interconnection: speed and fault tolerance. First, picking a rough estimate from the thalamus directly is a faster data path than going through the sensory cortex, allowing the system to react faster to broad classes of stimuli. Second, this path allows some emotional learning to proceed even if parts of the sensory cortex are damaged.

Having this input as part of the model gives rise to some interesting behaviors that in some ways mirror experimental data better than the present system.

With this input, we can simulate the impairment of the sensory cortices while still retaining the ability to react to some stimuli.

However, this input as we defined it is too coarse to really model early sensory inputs. Also, the very feature that made it work as a coarse model of these connections – the maximum over all inputs – also interfered with normal learning to such a degree that it was finally deemed unsuitable to be a part of the model until such a time when we can do the work to incorporate a good thalamic pathway properly into the model. At the time of writing, this has yet to take place.

We have previously run another set of simulations with more emphasis on the physiological aspects of the model both with and without the orbitofrontal or cortical parts and compared its performance with data from animal studies. These results are available in (Balkenius and Morén, 2000b).

In the next chapter, we will extend this model with a model of the hippocampus, and see what changes this can bring to the system.

Chapter 5

Context

In the previous chapter, we described a model of the amygdala and the orbitofrontal cortex. This model is based on the idea of the amygdala-orbitofrontal system as a Pavlovian learning system supplying emotional signals to other brain areas to be used in reinforcement learning in motor systems, to control attentional systems, influence long-term memory retention as well as directly control autonomic responses.

One component has been missing from the model, though. We know that the hippocampus is an important component in emotional conditioning, and we believe that it is responsible for supplying the system with a context. Note, however, that the hippocampal system is seen as an adjunct to the amygdala model we have been discussing; our focus is on its effects on the amygdala model rather than on its own properties.

5.1 Hippocampus

We reviewed the physiological structure of the hippocampus in chapter 2. Here we will briefly discuss the area as a computational model. In the search for a neurally based model of conditioning, it has become clear that many emotional learning phenomena depend in intricate ways on the hippocampus. This includes configurational and contextual conditioning of various types (Eichenbaum, 1999).

Our first attempt to model the processes that are thought to take place in the hippocampus were presented in (Balkenius and Morén, 2000a). This model differs from many models of the hippocampus such as those of Rolls and Treves (1998), Gluck and Myers (1993), and Schmajuk (1997) in that it does not attempt to model the underlying physiology directly. Instead, it is formulated at a more abstract level and is implemented with the overlying functionality of the area in mind, rather than as a direct mapping of the anatomical structure of this area. Of course, the model is still heavily influenced by neurological data and theories of hippocampal function on a larger scale. A version of this model has also been used to study contextual influences on action selection (Balkenius and Björne, 2001).

5.1.1 Other Models

Among the areas we are interested in, there is probably none that have been the focus of so much work in computational modeling as the hippocampus. No doubt this is at least partly because the functions attributed in whole of in part for the hippocampus – spatial map, short-term memory, context – are well-defined, easily understandable and obviously important. Also, the structure of the Hippocampus seems to lend itself well to modeling.

An early model of hippocampal function was created by Marr (1971) that modeled the hippocampus as an autoassociative network. His idea was that the hippocampus act as a short term associative memory. This idea has lived on in most later models in one form or another.

Some models view the hippocampus as a spatial map, strictly used for navigation and binding objects in space around the animal. Zipser (1985, 1986) proposed a connectionist model that maps landmarks and observer position to place field responses. A recent such model is by Redish and Touretzky (1997). Their model (called CRAWL), uses path integration on head direction to update place cells, and thus reconcile external cues with internally represented cues.

The CRAWL model is representative of this class of models. It is a largely functional model, concerned with modeling the behavioral data, rather than with neurological structure. There is thus little correspondence between the model and the neurological substrate. Also, there are no allowances for the other functions the hippocampus does seem to support.

Another class of models are those concerned mainly with neurological modeling of the area. The focus is on the dynamics, rather than system-wide behavior. The model described by Káli and Dayan (2000) and the model proposed by Rolls and Treves (1998) are both examples of such models. Though the focus is on the mapping between the model and the neurological substrate, they are nevertheless complete enough to be testable in behavioral simulations. The focus of Káli and Dayan (2000) is on hippocampus as a place-learning structure.

The model proposed by Rolls and Treves (1998), on the other hand, is focused on the role of the hippocampus as a context generator. Although it is a fairly low-level model, implemented as a set of neural networks, it still manages to be runnable in simulation. In this model, as in many other models, the CA3 area is seen as an autoassociative network used to store 'snapshots' of events during a short time frame. These events aren't only spatial, however. Instead, they are episodic memories of events originating in many other areas, including rewards and emotional states. The preceding stage, the dentate granules, form a competitive network for producing an orthogonal sparse coding of the inputs.

The final stage, CA1, implements two mechanisms. First, it uses local competition to produce new configural nodes based on commonly co-occurring representations in CA3. second, via the perforant path, it integrates the sparse recalled representations from CA3 with the full representation of the cues that elicited this representation.

Another recent context-based model has been developed by O'Reilly and Rudy (Rudy and O'Reilly, 2001; Frank et al., 2002). They look at contextual fear conditioning (where an animal associates the context to an unpleasant stimulus) and on the effects of the hippocampus on transitive inference. These observations are used to build a model based on the idea that the hippocampus encodes *conjunctive* representations, or new representations encoding for a flexible set of features representing a context. This idea is quite similar in scope to the model of Rolls above, and to our model, presented in the next section.

Their model is a neural-network incorporating a slow-learning cortical system with overlapping representations, and a faster-acting hippocampal system with separated representations O'Reilly and Rudy (2001). The hippocampal part utilizes sparse representations to achieve pattern separation and create configural representations. They also view the CA3 as an autoassociator to fill in and activate full representations of learned configurations. This model depends heavily on its cortical interconnections.

There are many other models with similar ideas of hippocampal function. Some propose a different function, however. McClelland et al. (1995) proposes that the function of the hippocampus is to aid in memory consolidation. The hippocampus stores short-term memories and acts as a teacher for supervised learning in the neocortex. This is based on the assumption that long term memory storage in the neocortex is best performed in small, incremental steps, rather than as massive change.

Another different model is the model by Schmajuk and DiCarlo, as described in (Schmajuk, 1997, p. 257). They view the hippocampus as computing an aggregate prediction of events and of error signals. This prediction is then used to modulate the the formation of CS-US associations and inhibits the US output to the cerebellar areas. This role is actually closer to the role we assign to the OFC than to the hippocampus in our model.

5.1.2 Context

Our view of the hippocampus is as a part of the system for emotional encoding. We believe that it's role is that it encodes context. Context is a feature that is often taken for granted by workers in learning theory. It is frequently defined in the negative, as any stimuli that is not directly involved in the present experiment, stimuli that somehow encode the entire situation, rather than individual features (Mackintosh, 1983), or as stimuli that are not being manipulated by the experimenter (Donahue and Palmer, 1994).

It is of course trivially true that it is the stimuli that are present that encodes context. The question becomes whether it is all stimuli (attended as well as non-attended), the stimuli that are not part of the experiment itself, or even only a (potentially) small subset of these stimuli that is used to encode contextual information. In the limiting case, only one stimulus may be involved in contextual discrimination, as seems to be the case for occasion setting (Schmajuk et al., 1998). Rolls and Treves (1998) view context as a collection of stimuli that helps activate an auto-association network that forms the core of the hippocampus.

How does the context view of the hippocampus relate to the function of the hippocampus as a memory facilitator or as a cognitive map? We believe that these views really describe the same functionality, but utilized differently in separate areas of the brain. Memory facilitation seems largely to just be the representation of contextual stimuli. Donahue and Palmer (1994) suggests that working memory and context share the same mechanisms in at least some tasks, such as matching-to-sample.

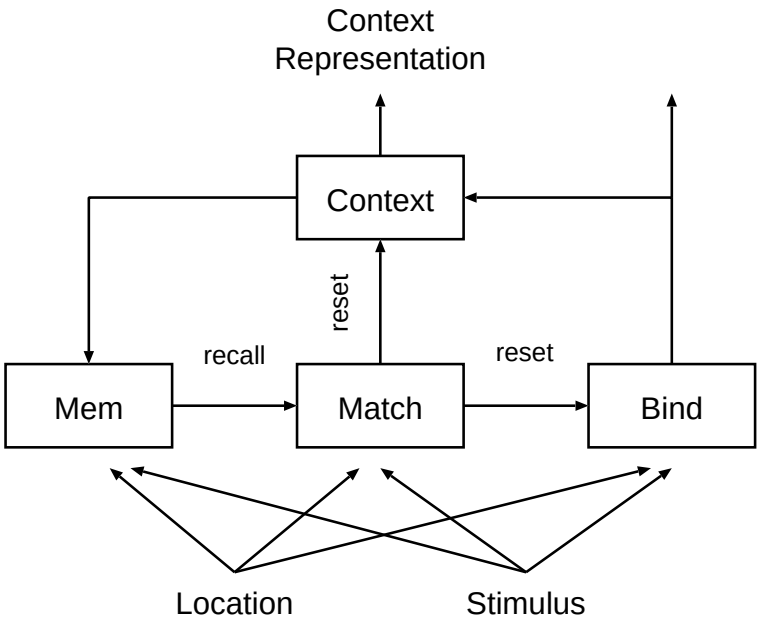

Figure 5.1: A model of contextual processing. 'Mem' is context dependent memory; 'Match' is the matching between present and recalled stimuli from 'Mem'; 'Bind' binds stimuli and locations; and 'Context' builds the contextual representation.

It is clear that the place representation of O'Keefe and Nadel (1978); Rolls and Treves (1998) does not preclude its use as context. Less obvious is that this means that the stimuli are likely to enter the hippocampus sequentially. This is due to the attentional system that is able to focus on only one thing at a time – if nothing else, when visual stimuli are spread around the area, the animal needs to move its gaze to take in each stimulus. This does bring the benefit of partially solving the binding problem; as only one object is focused at any one time, it is trivial to determine the position of each object. As only one object or feature is attended to at a time, all stimuli associated with that particular object, and only those stimuli, are received at a given time.

5.1.3 The Hippocampal Model

For this model, we have attempted to bring together some computational aspects of the hippocampus. This is not an attempt at a physiological model, but a model implementing the functionality we believe the hippocampus is responsible for.

As seen in figure 5.1, there are four main components of the hippocampus module that each relates to different aspects of hippocampal processing. The Bind subsystems is responsible for the binding of simultaneously presented stimuli. The Mem system acts as a context dependent memory that can generate predictions about stimuli that should appear in a certain context. The Match system compares actual stimuli to expectations fetched from the Mem system and signals potential mismatches. Finally, the Context system combines stimuli bound together in the Bind system to generate a contextual code that can be used by the amygdala and the orbitofrontal cortex.

The Bind is used to construct configurational stimuli that are subsequently used by the other modules. This aspect of the model is assumed to reflect the role of hippocampus in configurational conditioning (Eichenbaum and Buckingham, 1990). The Bind system is also an essential role of the context recognition system in Context. This part of the model parallels the role of the hippocampus in contextual conditioning as well as conditioning to spatial location. A context, in our view, can be either an abstract sequence of stimuli or a place defined by a number of stimuli at different locations around the animal. The system Mem stores expectations of the stimuli at different locations. This is a crude model of the role in hippocampus in episodic memory. In the current system, the prime role of these expectations are to be matched to the actual stimuli that the complete model receives, thus modeling the role of hippocampus in orienting to novel stimuli (O'Keefe and Nadel, 1978; Gray, 1995). This function includes the resetting of the contextual code and the initiation for a search for a new context representation that matches the current sequence of stimulus inputs (Carpenter and Grossberg, 1986). As input, we have a vector of stimulus inputs S and location inputs L, with the combined input vector as X.

The Bind system is essentially an ART network that forms categories of each stimulus-location combination (Carpenter and Grossberg, 1986). Unlike the ART network, however, each Bind representation stays active until it is reset by a mismatch in the Match system. As a result, a sequence of stimuli will lead to a whole set of Bind representation becoming active. The Bind module uses B, a set of binding nodes, and W, a vector of connection weights from the input to the nodes B.

The euclidean distance from the input to the existing bind nodes is calculated, and the node with the best match is selected:

$$D_i = d(W_i, X),$$

$$b = \min_i D_i$$

If $D_b \leq 0.5$, the best node is judged to be close enough, and is activated ($B_b = 1$). If, on the other hand, $D_b > 0.5$, none of the binding nodes match well, and a new node is created:

$$W_{new} = X, \text{and}$$

$$B_{new} = 1$$

The Match module matches the incoming stimuli S with the stimulus memory $M_{c,l}$ (from the Mem module) for context c at location l. This is expressed as a 'novelty' value N:

$$N = \sum_i [M_{c,l,i} - S_i]^+$$

This implies that a novel situation arises when an expected stimulus fails to be present at a particular location. Note that the mere appearance of a new stimulus is not considered novel in this sense; instead, a new binding node is created in the Bind subsystem.

If $N > 0$, the current context is wrong, so we reset all context nodes and old bind nodes and create a new context node if needed:

$$C_{0..i} = 0, \text{and}$$

$$B_i = \begin{cases} 1 & \text{if } i = new, \\ 0 & \text{otherwise} \end{cases}$$

We thus keep the the most recently activated bind node active.

If we created a new binding node in Bind, we now create a new context representation U:

$$U_{new} = B$$

U_c is the activation pattern of the active binding nodes for context c. If no expected stimulus was missing, on the other hand, we instead update the current context representation:

$$U_{c,i} = \max(B_i, U_{c,i})$$

The context representations and bind nodes are used to form the context categories (Balkenius and Morén, 2000a). The context categories in Context are asymmetric in the sense that a context can be activated completely by single stimuli, but its activity depends on how well the input from Bind matches the learned context template. In contrast to the original context model proposed in (Balkenius and Morén, 2000a), the output of the current model is constrained to be between 0 and 1 by normalizing the output from Context with regard to the maximum context:

$$C_i = (B_i U_i) / \max_i(C_i)$$

Each new combination of context, stimulus and location is stored in Mem in such a way that it can be recalled and matched against the input to the model:

$$M_{c,l} = S$$

When there is a mismatch, the Match system will reset the representations in Bind and Context which will temporarily shut off the output from the hippocampus until a new stimulus enters the model again and it begins to code for a new context.

This hippocampal model is clearly more abstract in nature than other, more neurologically oriented models, like Rolls and Treves (1998). Once the role of the Hippocampus in context generation has become clearer, this model is a good candidate for replacement with a more complete model with both the neurological underpinnings, as well as the expected functionality.

An Example

In figure 5.2 we have a small contextual experiment to demonstrate the hippocampal model in operation. In each phase either of two stimuli A or B is

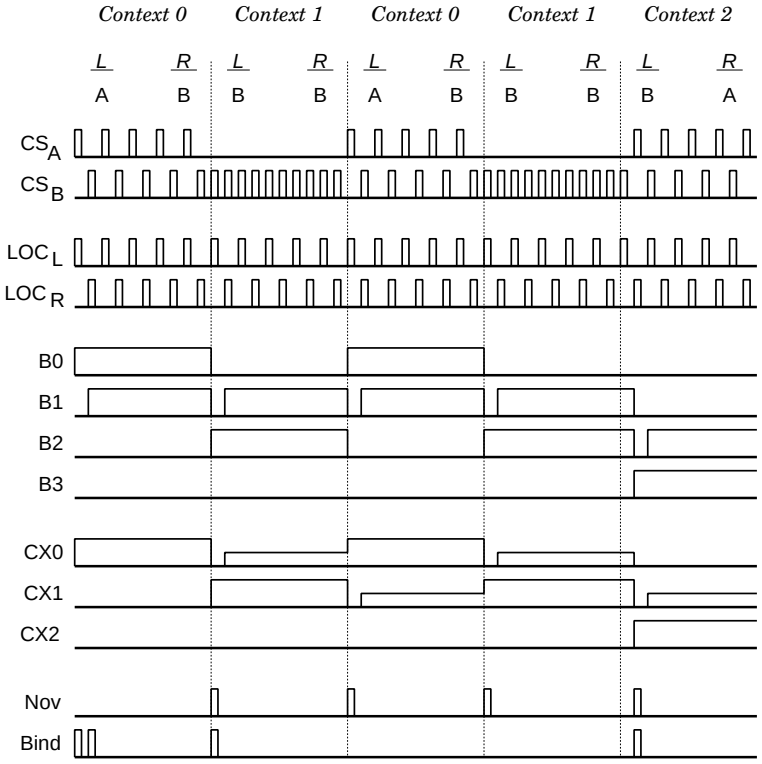

Figure 5.2: A test of hippocampal function. From top to bottom, we have two stimuli, two locations, the bind nodes, the context signals, the novelty signal and the bind signal. For each of the five phases, we can at the top see the current context and the contents of the two locations.

presented to the left or to the right, as seen at the top – of course, the terms 'left' and 'right' are used just for convenience; they are two arbitrary but different positions. Below, we have the stimulus signals, the location signals, the binding nodes and the context nodes, respectively. At the bottom it is depicted when an event is considered novel, and when a new binding node should be recruited.

This model does not model attentional processes. It is passive and relies on outside processes to direct attention within the environment. For our simulations, it means we actively signal where we are directing the attention at any one time.

This experiment is run in five phases. First, we present stimulus A to the left, then stimulus B to the right and repeat five times. Of course, this can

with equal validity be interpreted as the animal looking to the left and right repeatedly, with the stimuli staying fixed at their respective positions. When we first present stimulus A, a bind node (B0) is recruited, and a context CX0 is created and activated. When B appears at the right, another bind node is recruited, as the mismatch between the representation of B to the right and A to the left is too great. No new context is created, however, as the event is not considered novel in the sense that there is a mismatch between expectations and observations.

In the second phase, we suddenly present B at both positions, replacing stimulus A to the left. This triggers the creation of a new bind node B2 (no current node is similar enough), and will also produce a mismatch between the expectation of A to the left and finding B in that location. This results in the creation of a new context CX1. This new context is at full strength, as it matches the contents at both the left and right positions, while the old context CX0 is at half strength, as it matches B to the right (binding node B1 is activated), but not B to the left (binding node B0 is not active).

We switch back to the original configuration with A to the left and B to the right. The new binding node B2, representing B to the left, is deactivated, and the original binding node B0 is reactivated. A mismatch is triggered, as seen in the 'Nov' row, but no new binding node is needed, so no new context is created. Instead, the first context CX0 is again activated. Repeating the second phase again in the fourth phase shows the same result as for the third: the relevant binding nodes are activated, and we switch to the second context.

In the fifth and final phase, we show stimulus B to the left (which we already have had in phase two and four), and A to the right (which is a new situation). As with previous phases, we generate a novelty signal, but this time, we also generate a new bind node, as there is no existing node that encodes A being on the right. As we have both a novel situation and have created a new bind node, we also create a new context CX2, which represents this situation and becomes the most activated one. CX1, matching B on the left, but not A on the right, becomes partially activated, and CX0, which represents the reverse of the current stimulus-location pattern, does not become activated at all.

5.2 A Complete Model

Now we add the hippocampal model to the amygdala-orbitofrontal system we have been developing in chapter 4.

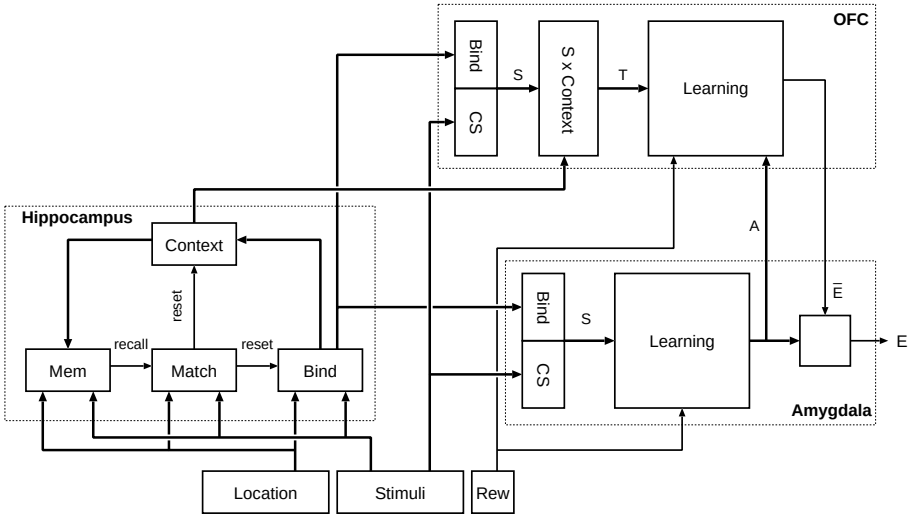

Figure 5.3: Schematic depiction of the complete model. There are three main components of the model, labeled 'hippocampus', 'amygdala' and 'OFC'. Details of each component are not depicted for clarity.

There are three major pieces to this model (figure 5.3). First, the hippocampus, that is responsible for generating location dependent stimuli and contexts. Next, the amygdala, that learns lasting emotional associations. Lastly, the orbitofrontal Cortex (OFC), that learns to inhibit emotional reactions in a context dependent way when appropriate.

5.2.1 The Amygdala

As we've been discussing in earlier chapters, the amygdala is designed to learn emotional associations. It is not used to inhibit any associations - that is the job of the OFC. This has support from neurological literature; once learned, an emotional reaction is very resistant to extinction in general. An extreme example of this is phobias, that are extinguished only with difficulty, and persists in reappearing even after treatment.

There are several inputs to the amygdala. Unlocalized stimuli are coming straight from the sensory areas. Localized, unexpected stimuli bound to a given place enter from the hippocampus. These stimuli are treated the same way in the amygdala. There is also a scalar reinforcer, used for the actual conditioning.

Outputs from the amygdala are the conditioned signals to the OFC, and the emotional conditioning. The OFC uses the conditioned signals to determine whether to inhibit the amygdaloid output, and sends an inhibitory signal to the amygdala.

The stimulus input to the amygdala consists of the stimuli CS and the bind node outputs from the hippocampus. These are concatenated into one stimulus vector S. The bind node outputs are thus treated as just another kind of stimuli. The bind input is an addition to the original model in chapter 4. Due to the nature of this model, the actual input is not the bind signals themselves, as the model is not designed to handle continuous signals. Instead it is the change in bind input strength – in practice, it signals a bind node activation with a one-step stimulus spike. An additional input signal is the scalar R representing the reinforcer.

The output E of the amygdala is calculated identically to chapter 4:

$$E = [\sum_{i=0}^{S} A_i - E_o]^+,$$

$$A_i = S_i V_i$$

where V is the connection weight vector, and E_o is the inhibitory signal from the OFC.

The connection weights are updated as:

$$\delta V_i = \alpha S_{i,t-1} [R - \sum_i A_{i,t-1}]^+$$

It is a standard associative learning rule in the manner of Rescorla-Wagner (Rescorla and Wagner, 1972), analogous to the Delta rule (Widrow and Hoff, 1988). The crucial difference is of course that the weight is allowed to only increase, never to decrease. This is quite deliberate; the amygdala should keep all learned associations indefinitely, while the OFC handles context-dependent inhibition when necessary.

5.2.2 The Orbitofrontal Cortex

The orbitofrontal cortex' role is to inhibit other areas in response to changing situations. In contrast to the amygdala, these learned inhibitions can be learned as well as extinguished. Also, again in contrast to chapter 4, these inhibitions are context dependent.

The OFC receives two new inputs as compared to the version in chapter 4, and a new subsystem to integrate contextual and stimulus information. The OFC receives the same inputs the amygdala system does – CS, Bind nodes and R – and also a context representation CON from the hippocampus and the intended output A from the amygdala.

As for the amygdala model, the CS and Bind node inputs are concatenated into one vector S. This vector is multiplied with the context vector CON to create an input matrix T:

$$T = \sum_{i=0}^{S} \sum_{j=0}^{CON} S_i CON_j$$

Thus there will be one T node for every combination of S and CON nodes.

The output E_o of the OFC is calculated the same as for the amygdala:

$$E_o = \sum_{i=0}^{T} O_i,$$

$$O_i = T_i W_i$$

with W being the connection weights for T.

The learning mechanism is almost the same as for the amygdala, the difference being that it is not constrained to only increase:

$$\delta W_i = \beta T_{i,t-1} R_o$$

The reward function is different however:

$$R_o = \begin{cases} [\sum_i A_i - R]^+ - \sum_i O_i & \text{if } R \neq 0, \\ [\sum_i A_i - \sum_i O_i]^+ & \text{otherwise} \end{cases}$$

The function looks the way it does simply because when there is no reinforcer, there is no reason to lower the inhibition for the stimulus even if the inhibition is stronger than it needs to be – it is quite probably already set at the proper level for inhibition in an already encountered situation that may well occur again.

5.2.3 Connections

As seen in picture 5.3, the model is composed of three heavily interconnected components. While each component does something minimally useful by itself, it is the interconnections that make it work like an integrated whole.

The hippocampus creates two kinds of outputs: Bound stimuli that fire whenever a given stimuli first shows up in a given place, and a context, used by the OFC for context-dependent inhibition.

The amygdala uses stimuli and the primary reinforcer to persistently learn emotional associations. It creates an emotional value that may be partially or totally inhibited by the OFC before exiting the model. It also sends a vector of the current emotional outputs to OFC.

The OFC binds together the stimuli (both bound and unbound) with the current context. It then compares the suggested output of the amygdala with the current reinforcer and inhibits as needed, using the context-dependent stimuli.

5.3 Simulations

In this section we present the results of a number of conditioning experiments simulated in the model. They are divided into four sections, according to the type of learning they represent.

There are two reasons to run simulations of experiments in this manner. First, well-known experiments are a good way to test any conditioning model, as the results are known. Second, and more importantly, it gives researchers a way of directly comparing their models with real animal data – and not only because we know what the animals did, but also because there are a lot of data and theories that explain why a given experiment works, and this can be compared to the inner workings of the model itself.

5.3.1 Basic Conditioning

These are a few experiments that are so basal to any conditioning model that serve mostly as tests to see whether the model works at all. Almost all of the more complex experiments make use of the mechanisms tested by these.

The first test of any conditioning model is basic acquisition and extinction. Our model manages this identically to the simulation in the preceding chapter. It also models reacquisition, where reconditioning to a previously conditioned, but extinguished, stimulus is faster than the original acquisition (see the previous chapter for details).

The model also handles blocking and inhibition in the same way as the previous model (see chapter 4). These protocols are not dependent on contextual effects or binding and thus run unchanged.

5.3.2 Discrimination

Here we test the ability to discriminate between similar or dissimilar stimuli in response to a reinforcer. As we saw in chapter 4, the model can do a simple discrimination learning protocol, where 'similar' stimuli are represented as collections of stimulus components with some components identical for both stimuli. The model was not able to handle positive and negative patterning.

Negative Patterning

Referred to as the XOR problem in connectionist literature, negative patterning is dependent on inhibition as an active process, as well as the ability to

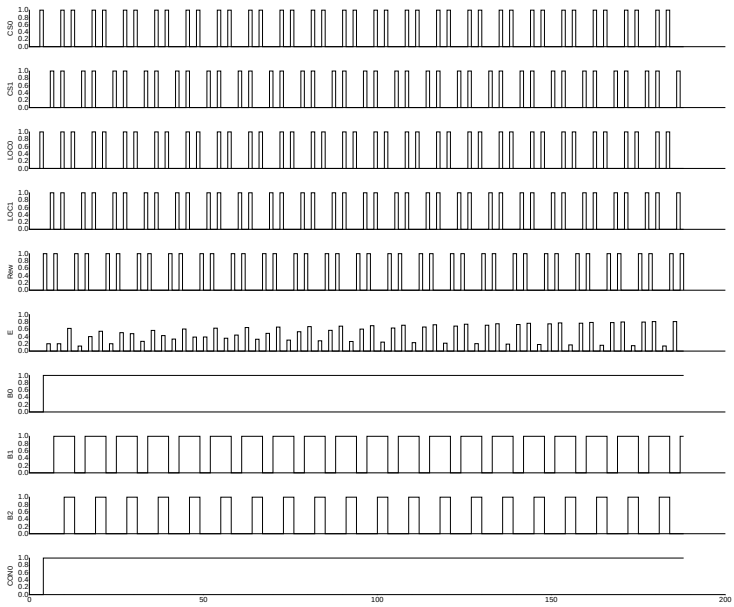

Figure 5.4: A negative patterning experiment. From top to bottom we have CS0 and CS1; Location 0 and 1; the reinforcer Rew; the output E; bind nodes B0 to B2; and Context 0. As seen in the output E, the model is able to learn to react to either stimulus, but not both together.

discriminate between stimuli. The idea is to learn that either of two stimuli is beneficial, whereas both taken together is not:

Condition:

CS0 + US
CS1 + US
CS0 & CS1

Test:

CS0 → CR
CS1 → CR
CS0 & CS1 → no-CR

This is quite a bit harder to learn than it might seem; CS0 and CS1 both predict a reinforcer, thus both together should, if anything, predict an even greater one. This implies that any mechanism by which a system is able to learn this experiment must necessarily be more complex than a simple collection of isolated nodes with excitatory or inhibitory connections.

The model handles this well. By being able to discriminate between a single stimulus and the compound of both stimuli it will gradually completely inhibit the combination as represented by the bind representation of the stimulus compound, while allowing a full reaction to either stimulus alone.

As we can see in figure 5.4, it is not the context that enables the model to handle this protocol. Instead, it is the binding nodes that combines stimuli with a place that allows this function. The bind node B2 is only activated when both stimuli are active, and it is the inhibition in the OFC connected to this bind node that inhibits the expression of a reaction to the combination, while allowing either stimulus to be conditioned by itself.

This is an example of configural learning from (Rudy and Sutherland, 1995; Bellingham et al., 1985), as discussed in (Rolls and Treves, 1998, p, 119), where it was hypothesized that the hippocampus is necessary for these kinds of discriminatory learning. For this model this is indeed the case.

Positive Patterning

Positive patterning is the flip side of negative patterning above. The subjects learn that while either of two stimuli do not predict a reinforcer, both together do:

Condition:
- CS0 + US
- CS1 + US
- CS0 & CS1

Test:
- CS0 → no-CR
- CS1 → no-CR
- CS0 & CS1 → CR

In practice, animals will respond weakly to the single stimuli even after extensive training.

This is the same protocol we ran in the previous chapter and showed how that version of the model could not handle these kinds of protocols. As seen in figure 5.5, the model manages to react much stronger to the combination of both stimuli than to either stimuli alone. Compare this with figure 4.9 on

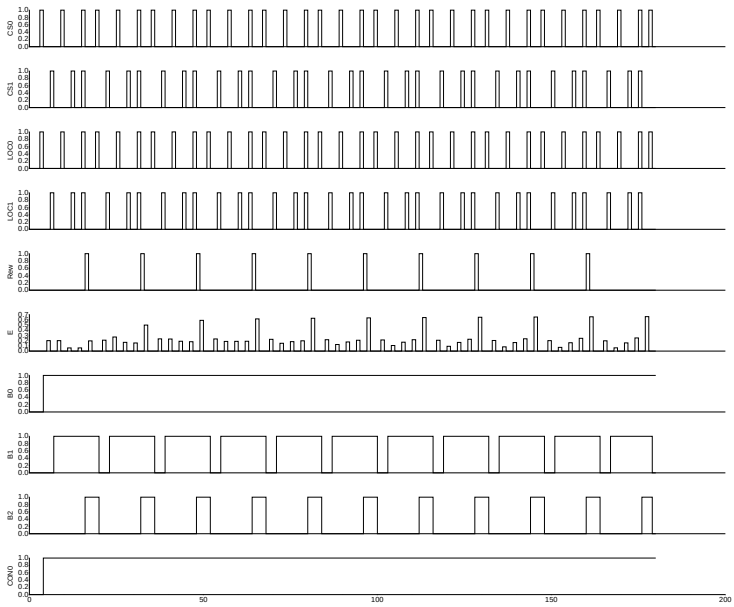

Figure 5.5: Positive patterning. From top to bottom we have CS0 and CS1; Location 0 and 1; the reinforcer Rew; the output E; bind nodes B0 to B2; and Context 0. The model is able to learn to react strongly to both stimuli, and weakly to either alone.

page 104. Note also that the model is not able to completely extinguish the reaction to a lone stimulus. This is in line with behavioral data, though we do not suggest that the mechanism responsible is the same.

Feature Discrimination

Simultaneous feature-positive discrimination (SFPD) and simultaneous feature-negative discrimination (SFND) are the simultaneous versions of positive and negative occasion setting, respectively (Schmajuk et al., 1998). As such, they are not as interesting as occasion setting proper (see chapter 3). SFPD has a simple protocol. Alternate between conditioning two stimuli, and non-conditioning of one of them:

Condition:
    CS0 &CS1 & + US
    CS1

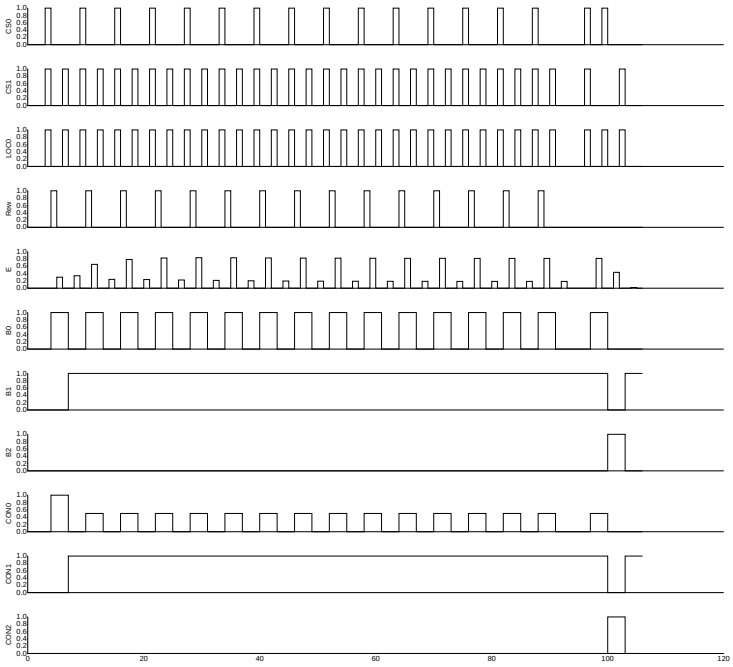

Figure 5.6: Simultaneous feature positive discrimination. From top to bottom we have CS0 and CS1; Location 0; the reinforcer Rew; the output E; bind nodes B0 to B2; and Context CON0 to CON2. We see that it manages to discriminate between the presence and absence of CS2.

Test:

CS1 → no-CR
CS0 → CR
CS0 & CS1 → CR

The animal will learn to react to to CS0 but not to CS1. The model also handles this, with greatest reaction to the compound stimuli, somewhat less to CS0 alone, and very little to CS1. The discrepancy between the compound stimuli and the single CS0 is due to the single stimuli not matching the conditioning context as well as the original compound.

The protocol for simultaneous feature-negative discrimination is identical to that of conditioned inhibition, unlike the serial version that is an example of occasion setting.

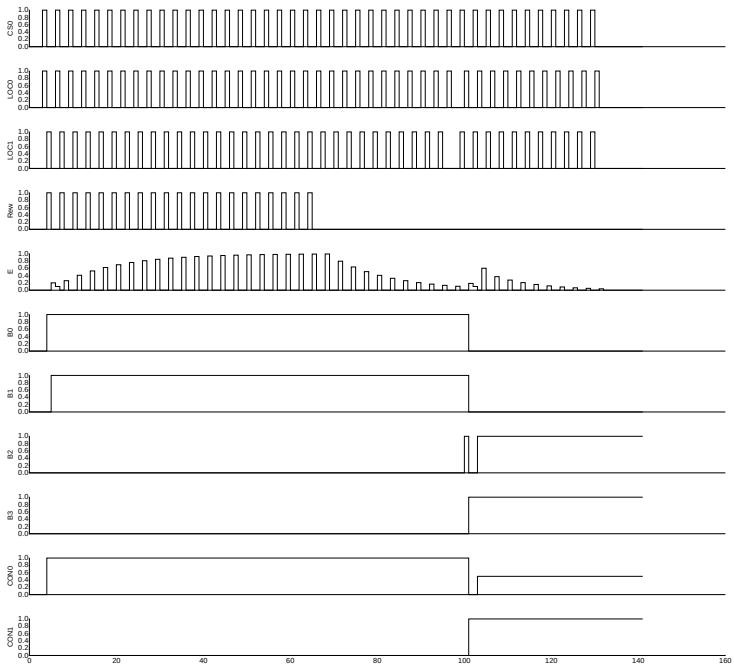

Figure 5.7: Contextual disinhibition. From the top we have CS0; LOC0 and 1; the reinforcer Rew and output E; Bind nodes B0 to B3; and context CON0 and CON1. CS0 is displayed at location LOC0, alternated with displays of no CS at LOC1. After conditioning and extinction, CS0 is moved to LOC1, which results in a disinhibition effect.

5.3.3 Context Effects

Context effects are those where the situation around the system becomes a factor in learning. Factors like the location of the stimulus or the location in which the experiment takes place become important.

Contextual Disinhibition

Contextual disinhibition occurs when a stimulus that is inhibited in one context loses its inhibition elsewhere (Schmajuk et al., 1998). It is a basic test of context-dependent inhibition.

the protocol:

Condition CS0 in context 1:
CX1: CS0 + US

Extinction:
CX1: CS0

Move to context 2:
CX2: CS0 → CR

Thus, the inhibition of CS0 is local to the context at which it is inhibited. The conditioning, on the other hand, is intact over all contexts. The protocol is shown simulated in figure 5.7. We reinforce CS0 at location LOC0, and alternate viewing that location with viewing the empty location at LOC1. Then we inhibit CS0 while still alternating the view of LOC0 and LOC1 (We could get the same result by just viewing LOC1 once). Then we move CS0 to LOC1, and view LOC0 and LOC1 as before. We get disinhibition. Unlike the patterning protocols above, this is a full context effect.

Disinhibition by Novel Stimulus

Disinhibition by novel stimulus is another experiment where inhibition is suppressed. This effect is achieved by presenting a novel stimulus for the animal to react to. It is similar to contextual disinhibition, but instead of changing the context, we change the stimuli present in the same context. We alternate the acquisition and extinction of CS0 at LOC0 with presentations of CS1 at LOC1. When CS0 is extinguished, we remove CS1 and replace it with CS2 at the same location, and continue to alternate between LOC0 and LOC1. The result is disinhibition of CS0.

As we see in figure 5.8, the model treats this protocol in the same way as contextual disinhibition. The appearance of a new stimulus CS2 in place of CS1 signals the creation of another bind node and a new context.

In animal experiment terms, contextual disinhibition would be moving the animal to a different experimental cage, while disinhibition by a novel stimulus would be presenting a novel stimulus in the original cage. Also, in contrast to contextual disinhibition, disinhibition by novel stimulus is a transient phenomenon; the animal will quickly generalize the inhibition to the changed surroundings.

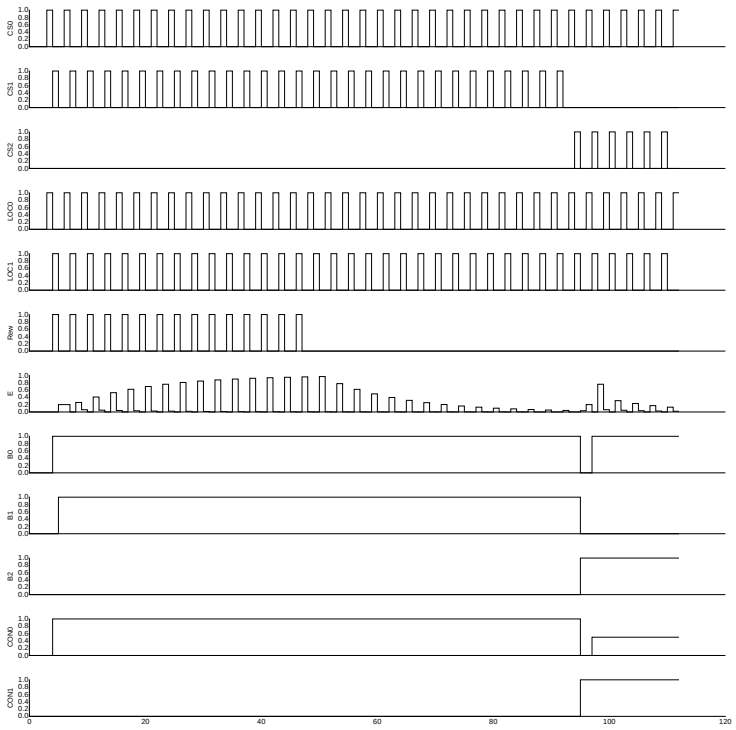

Figure 5.8: Disinhibition by novel stimulus. From the top we have CS0 to CS2; LOC0 and LOC1; the reinforcer Rew and output E; Bind nodes B0 to B2; and context CON0 and CON1. We alternate conditioning and extinction of CS0 at LOC0 with presentation of CS1 at LOC1. After extinction, we replace CS1 with CS2 at the same position and get disinhibition of CS0.

5.4 Discussion

As seen in this chapter, the integrated model has several very promising features: it handles most conditioning experiments, except for timing effects; it integrates context processing and conditioning; and it serves as a computational model for for interactions between the amygdala, the orbitofrontal cortex and the hippocampus.

The model takes inputs that are more realistic in nature than is usually the case. It needs stimuli and a simultaneous place representation, making it suitable for adding an attentional model. As it generates its contexts internally, it needs no externally imposed context representation. This makes the model suitable for implementation in a real-world setting such as a robot, as a stimulus collection and place signal are a good fit for the kind of data such a platform can provide.

The full model still has the limitation of only having one output, and of not being able to selectively inhibit responses on a stimulus level. the single most limiting factor, however is its inability to handle timing factors. All stimuli are a single timestep long, and must precede the US by a single timestep. This makes it impossible to run trace conditioning or any other timing-dependent protocols. The reason for the lack of this ability is that we have not yet focused on these aspects of learning. Further discussion on this will appear in the next chapter.

Chapter 6

Conclusion

In the last two chapters, we have looked at our model in detail. Now we will take a look at how this model would fit in from a systems level perspective. We will also summarize some of the results presented in earlier chapters and briefly discuss the model in light of these results.

As we saw in chapter 3 (page 78), the two-process model of instrumental conditioning proposed by Mowrer (1973) is based on the interaction of two distinct learning systems. There is an evaluation system, implemented as a classical conditioning system, and an action system implemented as an instrumental conditioning system. An alternative formulation of this idea within the realm of machine learning is the actor-critic framework proposed by Barto et al. (1983).

This is a general architectural structure that is applicable for any instrumental learning system. We view the amygdalo-orbitofrontal system as an evaluation system for several weakly interacting learning systems in the brain.

One such learning system is the amygdala and orbitofrontal cortices together with the basal ganglia as a system for motor reactions to emotionally significant stimuli. For this system, the amygdala model described here would be the classical conditioning module, and the basal ganglia would do instrumental conditioning based on the evaluation generated by the amygdala model.

The system is slightly different from the canonical two-process model discussed in chapter 3. As the OFC is inhibitory for both the amygdala and the basal ganglia, it has been broken out and added as a third module in figure 6.1.

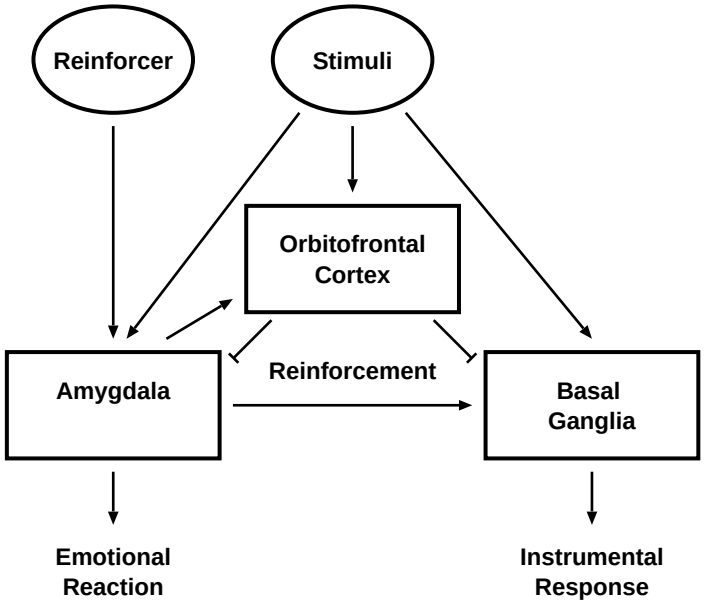

Figure 6.1: The amygdalo-hippocampal system and the basal ganglia as parts of a two-process model of conditioning. The figure is similar to the figure on page 78, but altered to illustrate the specific situation for the amygdalo-hippocampal system and the basal ganglia.

The connection between the OFC and the basal ganglia has been described in Fuster (1997), where he reports that among the many projections from the prefrontal cortex to the basal ganglia is a projection from the orbitofrontal cortex to the caudate. He also cites a number of other connections between the prefrontal cortex and various areas of the basal ganglia; it is less clear which prefrontal areas are involved in these connections, however.

What does this additional path mean? A speculative answer is that the role of the OFC for this connection is the same as for its connection to the amygdala: to inhibit emotional reactions in a context-dependent manner. This idea ties in with the work of Poremba and Gabriel (1999), that show that the amygdala is involved in the formation but not expression of responses to emotionally charged stimuli. Evidence discussed in section 2.1.1 supports this view. This is also consistent with the two-process architecture we have mapped the system on. If the OFC is to inhibit these responses, it needs both to inhibit the amygdala directly, to stop it from inducing conditioning to the stimulus, and inhibit the basal ganglia operation so the now undesired reaction is never expressed.

	Whole	-OFC	-Hipp	-OFC, Hipp
Delay Conditioning	+	+	+	+
Reacquisition	+	-	+	-
Extinction	+	-	+	-
Blocking	+	+	+	+
Conditioned Inhibition	+	-	+	-
Spontaneous Recovery	+-	-	-	-
Discrimination	+	-	+	-
Positive Patterning	+	+	-	-
Negative Patterning	+	-	-	-
Contextual Disinhibition	+	-	-	-
Feature Discrimination	+-	-	+-	-

Table 6.1: Running the model with 'lesions' of the main areas. Whole: the entire model is functioning; -OFC: the OFC is disabled; -Hipp: the hippocampus model is disabled (effectively the same result as when running the model in chapter 4); -OFC, Hipp: the model with only the amygdaloid part active. We do not show results for when disabling the amygdala, as the model would generate a zero response for all cases.

6.1 Review

In previous chapters, we have presented a model of emotional conditioning, and tested it with various conditioning protocols. We have also discussed other models of conditioning and context that share some of the goals or techniques of this one. Here we take a look at some of the results of these simulation experiments and summarize these results.

6.1.1 Lesions

With a model composed of semi-independent functional units comes the possibility of doing 'lesioning' experiments, vaguely analogous to studies being done in animal experiments. By selectively disabling parts of the model, we can gain some further understanding of the functional mechanisms by which it operates. This can also be a further way to assess how well the model accounts for the physiological mechanisms we are trying to simulate. At this stage, we do not presume that our model is in fact so detailed as to accurately exhibit the behavioral changes that the real lesions produce, but we feel it may nevertheless be instructive to see the results. The protocols were described on page 49 onwards.

Delay conditioning is the simplest, most paradigmatic of the conditioning protocols. This functionality rests entirely within the amygdala module of the model, and so is unaffected by the lesions in table 6.1. Equally simple in this respect is blocking. It rests on the ability of the model to assign credit proportion to the predictive value of any one stimulus, an ability that again rests within the amygdala module. Note, however, that partial impairment of classical delay conditioning has been found in hippocampus-lesioned rats; see cf. Bonardi (2001) for details.

Conditioned inhibition is typical for the simple protocols using inhibition. When we remove the OFC, there is no longer any inhibitory control of the system, so of course, no conditioned inhibition can be learned. The presence of absence of a hippocampal system does not affect the protocol.

For discrimination, there seems to be a partial effect when running one cycle without the OFC. This effect is entirely due to the non-reinforcement of the unique component of the second stimulus group, however, and is not indicative of real discrimination behavior. Schoenbaum et al. (2002) has found this pattern in a different discrimination task. They found that rats could acquire odor discrimination in a go, no-go task but could not learn reversal of the task. Our model does do this for the 'wrong' reasons however, so this similarity should probably not be taken too seriously.

In physiological studies, this effect is well known. As we saw in chapter 2, Shimamura (1995) and Kolb and Whishaw (1990) show that patients with damage in this area show difficulty in adapting to changing criteria in the Wisconsin card-sorting test. Gallagher et al. (1999) describe a set of experiments on rats that were subject to lesions in the OFC. The rats were trained to associate a visual stimulus (a light) to food, after which they were lesioned. The food was then associated with an aversive stimulus, and then they were tested with the visual stimulus. The rats had retained the CS-food association, and had learned the aversive association between the food and the aversive stimulus, but the lesioned rats still reacted positively to the visual stimulus. The OFC lesioned rats were thus unable to extinguish the positive emotional reaction to the visual stimulus, while the control group could do so.

Positive and negative patterning give different results in lesioning simulations. While positive patterning needs only a context to function, negative patterning needs both a context system and an inhibitory system. The task in both protocols is to distinguish the presence of either stimulus alone and the presence of both stimuli simultaneously. This accounts for the necessity of a context generation system. Negative patterning also needs active inhibition to produce the correct response for these contexts, however, which explains the additional need for an inhibitory system – and is also what makes this protocol

computationally strictly harder than positive patterning (see Hassoun (1995)). Gisquet-Verrier and Massioui (1997) show that rats with hippocampal lesions are indeed unable to learn positive patterning, or to retained the previously learned protocol.

Simultaneous feature positive discrimination needs both the OFC and the hippocampus to function. Without a hippocampal system, the alternately reinforced stimulus will elicit a response when testing the conditioning. Without an OFC, the alternately reinforced stimulus will not be suppressed during training. Compare with figure 5.6 on page 129 of the model without lesions, where CS1 is suppressed during training, and elicits almost no response when tested.

6.1.2 A Comparison of Models

In previous chapters we have presented a number of conditioning experiments. We have also discussed various models intended to model these experiments in a more or less biological manner. These models where not chosen necessarily to be 'state of the art', but to represent common approaches to modeling classical conditioning.

We have run several simulations on both our model and on other conditioning models. Here we summarize these simulations and briefly discuss the results. A more in-depth discussion of the results for the other models can be found in chapter 3 on page 61 and in (Balkenius and Morén, 1998a).

All models tend to handle basic conditioning protocols quite well. This should of course not come as a surprise; these protocols test the basis of conditioning, and a model that could not handle them is going to be crippled when attempting to model more complex phenomena. It's worth noting that all models can handle conditioned inhibition; this is commonly seen as a good test of a model's abilities.

Timing Effects

The other models can all handle various timing related effects to some degree. All of the protocols in that section of table 6.2 are dependent on the ability to handle trace conditioning, ie. the form of conditioning where there is a gap between the offset of the CS and the onset of the US. Our model is not able to handle this effect, and thus not any of the other timing dependent effects either.

	SB	TD	Klopf	Balk	SD	BM
basic conditioning						
Delay Conditioning	-	+	+-	+	+	+
S-shaped Acquisition	-	-	+	+	-	-
Reacquisition	-	-	+-	-	+	+
Extinction	+	+	+	+	+	+
Blocking	+	+	+	+	+	+
Conditioned Inhibition	+	+	+	+	+	+
Discrimination	+	+	+	+	+	+
timing dependent						
Trace Conditioning	+	+	+	+	+	-
ISI Curve	-	+	+-	+-	+	-
Secondary Conditioning	+-	+-	+	+	+-	-
Facilitation	+	+	+	+	+	-
Spontaneous Recovery	-	-	-	-	-	+-
context dependent						
Positive Patterning	?	?	-	-	+-	+
Negative Patterning	-	-	-	-	+-	+
Contextual Disinhibition	-	-	-	-	-	+
Feature Discrimination	-	-	-	-	+-	+-

Table 6.2: A summary of experiments and models talked about in this thesis. SB - the Sutton-Barto model; TD - the Temporal Difference model; Klopf - the Klopf model; Balk - the Balkenius model; SD - the Schmajuk-DiCarlo model; and BM - the Balkenius-Morén model (discussed in chapters 4 and 5).

An interesting exception is the Schmajuk-DiCarlo (SD) model (Schmajuk and DiCarlo, 1992). It can handle secondary conditioning (see page 53) only when CS1 and CS2 are paired simultaneously, but not when they are paired in sequence. This was discussed further on page 74.

What are we giving up at this time due to the lack of timing constraints? First, no learning protocols involving delay or trace conditioning will work. The same goes for secondary conditioning and higher-order reinforcement. In general, anything that depends on sequential, rather than simultaneous, presentation of stimuli will fail with this model.

There is of course no lack of models that can handle timing constraint to a greater or lesser extent (see Schmajuk (1997) for an overview; Sutton and Barto (1998) has a different perspective on the issue). This aspect has not been the focus of our work on this model. Instead, we have been aiming at having the larger-scale structure become a reasonable model of some of the interactions between the corresponding brain areas. When we feel that the structure

works, we can turn our attention to better, more complete models of learning within each module.

There is of course nothing that intrinsically prevents us from adding timing effects to the present model. Indeed, with care, one could choose any of the techniques used by the other models - delay nodes, trace or temporal difference measurement - and add to this model without losing the other features of the model.

Spontaneous Recovery

Spontaneous recovery is marked as a maybe ('+-') for our model. This is due to a question over what spontaneous recovery really is. Our belief is that it is a contextual effect, in that new context representations are created during the passing of time between the last trial and the test. It is possible that even the passing of time in itself may create new context representations (a given place in the morning is different than the same place in the afternoon). These new representations – reflecting the changed circumstances and experiences of the system since the last trial – will influence the representation the next time it is placed in the learning environment. An argument in favor of this interpretation is that contextual disinhibition and disinhibition by novel stimulus (Schmajuk et al., 1998) exhibits a similar effect by changing the context in various ways.

On the other hand, spontaneous recovery may be an effect simply of time passing, for one reason or another. This could be a side effect of consolidating memory from short-term memory systems into long-term systems, or it could be a time-dependent extinction of recently learned inhibitory connections; this view is partly supported in our model, as it sees acquisition as a far more permanent learning event than the context-dependent inhibition.

If we accept spontaneous recovery to be a contextual effect, our model supports the protocol. If it is a timing effect, it does not. Of course, the effect could be due to a combination of these and other factors, in which case we would get a partial effect.

Context Effects

For the context dependent protocols, the situation is reversed as compared to the timing dependent effects above. Our model – having explicit functional-

ity for handling context – performs quite well, while the other models tested largely lack this ability. The critical feature needed is the ability to allow one stimulus to modulate the reaction to another one. This is possible to add to these models just as timing effects can be added to ours.

No model can handle all forms of feature discrimination. Though all can handle simultaneous feature-negative discrimination, this is only due to its protocol being identical to conditioned inhibition. Our model and the SD model handles simultaneous feature-positive discrimination. Though the SD model does handle the serial versions in some respects, they can not pass the tests for determining whether the conditioning has resulted in occasion setting.

We have not run positive patterning protocols on the SB and TD models; thus the results for this protocol is unknown.

6.1.3 Future Research

This thesis has presented a model of emotional conditioning in the amygdala. This model is not complete in any way, but is rather a testbed for modeling neural structures at this level of detail. As a functional model, it has several good properties.

From a physiological standpoint, it shows that the connections found between the included areas do support the kind of functionality that has been inferred for them. Also, the 'lesions' performed on the model in some aspects match the results from animal studies and clinical observations. As a computational model of conditioning, it handles several conditioning protocols (excepting timing-related effects) and shows how the needed functionality can be distributed in the model, rather than attempting to add it all into one learning module.

Most of all, the model shows the feasibility of making a computational model at this level of detail, guided by both neurophysiology and by behavioral data.

There are a number of design changes and extensions that can improve this model:

Timing is crucial for modeling of many conditioning effects. This has been discussed earlier in this chapter.

A thalamic module was discussed in chapter 4 on page 108. This connection is important as an early pathway to the emotional system from the sensory

areas, and in earlier versions of this model we have attempted to include a simple thalamic module. While the presence of this module did give us some effects seen in animal studies, the module itself was too nonfunctional to make even a vague representation of this functionality and we decided in the end to remove it until we can do a better model of this area.

Larger scale networks in the model are necessary to be able to handle more complex environments. The model needs to be able to scale up for several hundred stimuli, and it needs to be able to emit more than a single scalar as a result of conditioning. This would as a first step be done by allowing each stimulus node to emit a signal instead of collecting them into the output. This will demand that the inhibitory system be able to inhibit outputs selectively, rather than everything at once. This was touched upon in chapter 4 on page 106.

That said, this model does have the critical features needed to represent the basic functionality of the chosen areas. With the addition of the features mentioned above, it should be feasible to implement this in a real-world robotic platform and thus be able to test it with real data as a part of a larger system.

Bibliography

Adolphs, R., Damasio, H., Tranel, D., and Damasio, A. R. (1996). Cortical systems for the recognition of emotion in facial expressions. *Journal of Neuroscience*, 16(23):7678–7687.

Amaral, D. G., Price, J. L., Pitkanen, A., and Carmichael, S. T. (1992). Anatomical organization of the primate amygdaloid complex. In Aggleton, J. P., editor, *The amygdala: neurobiological aspects of emotion, memory, and mental dysfunction*, pages 1–66. Wiley, New York.

Armony, J. L., Servan-Schreiber, D., Cohen, Jonathan, D., and LeDoux, J. E. (1997). Computational modeling of emotion: explorations through the anatomy and physiology of fear conditioning. *Trends in Cognitive Sciences*, 1(1):28–34.

Balkenius, C. (1993). Motivation and attention in an autonomous agent. In *the workshop on architectures underlying motivation and emotion - WAUME 93*. University of Birmingham.

Balkenius, C. (1995). *Natural intelligence in artificial creatures*. PhD thesis, Lund University Cognitive Studies 37.

Balkenius, C. (2000). Attention, conditioning and habituation: toward a computational model. *Cognitive Science Quarterly*, 1(2):171–214.

Balkenius, C. and Björne, P. (2001). Toward a robot model of attention-deficit hyperactivity disorder (adhd). In Balkenius, C., Zlatev, J., Kozima, H., Dautenhahn, K., and Breazeal, C., editors, *Epigenetic Robotics*, pages 61–67. Lund University Cognitive Studies, Lund, Sweden.

Balkenius, C. and Morén, J. (1998a). Computational models of classical conditioning: a comparative study. In Pfeifer, R., Blumberg, B., Meyer, J., and Wilson, S., editors, *From animals to animats 5: Proceedings of the Fifth International Conference on Simulation of Adaptive Behavior*. MIT Press, Cambridge, MA.

Balkenius, C. and Morén, J. (1998b). Dynamics of a classical conditioning model. In Bodén, M., Nicklasson, L., and Ziemke, T., editors, *ICANN 98*. Springer-Verlag, Berlin.

Balkenius, C. and Morén, J. (1999). Dynamics of a classical conditioning model. *Autonomous Robots*, 7(1).

Balkenius, C. and Morén, J. (2000a). A computational model of context processing. In Mayer, J.-A., Berthoz, A., Floreano, D., Roitblat, H. L., and Wilson, S. W., editors, *From animals to animats 6*, pages 256–265. MIT Press, Cambridge, MA.

Balkenius, C. and Morén, J. (2000b). Emotional learning: a computational model of the amygdala. *Cybernetics and Systems*, 32(6):611–636.

Balleine, B. W. (1992). The role of incentive learning in instrumental performance following shifts in primary motivation. *Journal of Experimental Psychology: Animal Behavior Processes*, 18:236–250.

Balleine, B. W., Garner, C., Dickinson, A., and Gonzalez, F. (1995). Motivational control of heterogeneous instrumental chains. *Journal of Experimental Psychology: Animal Behavior Processes*, 21:203–217.

Barto, A. G., Sutton, R. S., and Anderson, C. W. (1983). Neuronlike elements that can solve difficult learning control problems. *IEEE Transactions on Systems, Man, and Cybernetics*, 13(5):834–846.

Bellingham, W. P., Gillette-Bellingham, F., and Kehoe, E. J. (1985). Summation and configuration in patterning schedules with the rat and rabbit. *Animal Learning and Behavior*, 131:152–164.

Blazis, D. E. J., Desmond, J. E., Moore, J. W., and Berthier, N. E. (1986). Simulation of the classically conditioned nictitating response of a neuron-like adaptive element: a real-time variant of the sutton-barto model. In *Proceedings of the Eigth Annual Conference of the Cognitive Science Society*, Hillsdale, NJ. Erlbaum.

Bliss, T. V. P. and Lomo, T. (1973). Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. *Journal of Physiology*, 232:331–356.

Bonardi, C. (2001). Dorsal hippocampal lesions impair appetitive classical conditioning to localized cues. *European Journal of Neuroscience*, 13:1435–1443.

Bouton, M. E. and Nelson, J. B. (1998). Mechanisms of feature-positive and feature-negative discrimination learning in an appetitive conditioning paradigm. In Schmajuk, N. and Holland, P. C., editors, *Occasion setting: Associative learning and cognition in animals*, pages 69–112. American Psychological Association, Washington, DC.

Broks, P., Young, A. W., Maratos, E. J., Coffey, P. J., Calder, A. J., Isaac, C. L., Mayes, A. R., Hodges, J. R., Montaldi, D., Cezayirli, E., Roberts, N., and Hadley, D. (1998). Face processing impairments after encephalitis: Amygdala damage and recognition of fear. *Neuropsychologia*, 36(1):59–70.

Butler, R. and Harlow, H. F. (1954). Persistence of visual exploration in monkeys. *Journal of Comparative and Physiological Psychology*, 46:258.

Cahill, L. and McGaugh, J. L. (1998). Mechanisms of emotional arousal and lasting declarative memory. *Trends in Neurosciences*, 21(7):294–299.

Carpenter, G. A. and Grossberg, S. (1986). Neural dynamics of category learning and recognition: attention, memory consolidation and amnesia. In Grossberg, S., editor, *The adaptive brain*, volume 1 of *Advances in psychology*, pages 238–308. North-Holland, Amsterdam.

Clark, R. and Squire, L. (1998). Classical conditioning and brain systems: A key role for awareness. *Science*, 280:77–81.

Craig, A., Bushnell, M., Zhang, E., and Blomqvist, A. (1994). A thalamic nucleus specific for pain and temperature sensation. *Nature*, 372:770–773.

Crespi, L. P. (1942). Quantitative variation in incentive and performance in the white rat. *American Journal of Psychology*, 40:467–517.

Côté, L. and Crutcher, M. D. (1991). The basal ganglia. In Kandel, E. L., Schwartz, J. H., and Jessell, T. M., editors, *Principles of Neural Science*, pages 647–659. Prentice-Hall, London.

Davidson, R. J. and Irwin, W. (1999). The functional neuroanatomy of emotion and affective style. *Trends in Cognitive Science*, 2(1):11–21.

Davidson, T. L. (1998). Hunger cues as modulatory stimuli. In Schmajuk, N. and Holland, P. C., editors, *Occasion setting: Associative learning and cognition in animals*, pages 223–248. American Psychological Association, Washington, DC.

Davis, M. (1992). The role of the amygdala in conditioned fear. In Aggleton, J., editor, *The Amygdala: Neurobiological Aspects of Emotion, Memory and Mental Dysfunction*, pages 255–306, New York. Wiley-Liss.

Davis, M. (2000). The role of the amygdala in conditioned and unconditioned fear and anxiety. In Aggleton, J. P., editor, *The Amygdala: A functional analysis*, pages 213–287. Oxford University Press, Oxford, UK.

Desimone, R. (1991). Face-selective cells in the temporal cortex of monkeys. *Journal of Cognitive Neuroscience*, 3:1–8.

Desimone, R., Albright, T. D., Gross, C. G., and Bruce, C. J. (1984). Stimulus-selective properties of inferior temporal neurons in the macaque. *Journal of Neuroscience*, 8:2051–2068.

Desmond, J. E. (1990). Temporally adaptive responses in neural models: the stimulus trace. In Gabriel, M. and Moore, J., editors, *Learning and computational neuroscience: foundations of adaptive networks*, pages 421–456. MIT Press, Cambridge, MA.

Donahue, J. W. and Palmer, D. C. (1994). *Learning and complex behavior*. Allyn and Bacon, Boston, Massachusetts.

Edelman, G. M. (1989). *Neural Darwinism*. Oxford University Press, Oxford.

Eichenbaum, H. (1999). The hippocampus and mechanisms of declarative memory. *Behavioural and Brain Research*, 103:123–133.

Eichenbaum, H. and Buckingham, J. (1990). Studies on hippocampal processing: experiment, theory, and, model. In Gabriel, M. and Moore, J., editors, *Learning and Computational Neuroscience: Foundations of Adaptive Networks*, pages 171–232. MIT Press, Cambridge, MA.

Everitt, B. J., Cardinal, R. N., Hall, J., Parkinson, J. A., and Robbins, T. W. (2000). Differential involvement of amygdala subsystems in appetitive conditioning and drug addiction. In Aggleton, J. P., editor, *The Amygdala: A functional analysis*. Oxford University Press, Oxford, UK.

Fanselow, M. S. (1994). Neural organization of the defensive behavior system responsible for fear. *Psychonomic Bulletin and Review*, 1:429–438.

Fanselow, M. S. and LeDoux, J. E. (1999). Why we think plasticity underlying Pavlovian fear conditioning occurs in the basolateral amygdala. *Neuron*, 23:229–232.

Fendt, M. and Fanselow, M. S. (1999). The neuroanatomical and neurochemical basis of conditioned fear. *Neuroscience and Biobehavioral Reviews*, 23:743–760.

Ferster, C. B. and Skinner, B. F. (1957). *Schedules of reinforcement*. Appleton-Century-Crofts, New York.

Frank, M. J., Rudy, J. W., and O'Reilly, R. C. (2002). Transitivity, flexibility, conjunctive representations and the hippocampus: II: A computational analysis. *Hippocampus, in press*.

Fuster, J. M. (1997). *The prefrontal cortex: anatomy, physiology, and neuropsychology of the frontal lobe.* Lippincott-Raven, Philadelphia.

Gaffan, D. (1992). Amygdala and the memory of reward. In Aggleton, J. P., editor, *The amygdala: neurobiological aspects of emotion, memory, and mental dysfunction*, pages 471–484. Wiley, New York.

Gallagher, M. (2000). The amygdala and associative learning. In Aggleton, J. P., editor, *The Amygdala: A functional analysis*, pages 311–329. Oxford University Press, Oxford, UK.

Gallagher, M., McMahan, R. W., and Schoenbaum, G. (1999). Orbitofrontal cortex and representation of incentive value in associative learning. *The Journal of Neuroscience*, 19(15):6610–6614.

Gallistel, C. R. (1990). *The organization of learning*. MIT Press, Cambridge, MA.

Gallistel, C. R. and Gibbon, J. (2000). Time, rate and conditioning. *Psychological Review*, 107:289–344.

Gärdenfors, P. (2000). *Conceptual Spaces*. Bradford Books, Cambridge, MA.

Gisquet-Verrier, P. and Massioui, N. (1997). Selective hippocampal lesions in rats disrupt acquisition and retention of a positive patterning discrimination. *Physiology & Behaviour*, 61(4):577–589.

Gluck, M. A. and Myers, C. E. (1993). Hippocampal mediation of stimulus representation: A computational theory. *Hippocampus*, 3:491–516.

Gray, J. A. (1975). *Elements of a two-process theory of learning*. Academic Press, London.

Gray, J. A. (1982). *The neuropsychology of anxiety: an enquiry into the functions of the septo-hippocampal system*. Oxford University Press, Oxford.

Gray, J. A. (1995). A model of the limbic system and basal ganglia: Application to anxiety and schizophrenia. In Gazzaniga, M. S., editor, *The cognitive neurosciences*, pages 1165–1176. MIT Press, Cambridge, MA.

Gray, J. A., Feldon, J., Rawlins, J. N. P., and Hemsley, D. R. (1981). The neuropsychology of schizophrenia. *Behavioral Brain Science*, 14:1–20.

Hall, G. and Pearce, J. M. (1979). Latent inhibition of a cs during cs-us pairings. *Journal of Experimental Psychology: Animal Behavior Processess*, 18(5):32–42.

Hanson, H. M. (1959). Effects of discrimination training on stimulus generalization. *Journal of Experimental Psychology*, 58:321–334.

Hassoun, M. H. (1995). *Fundamentals of Artificial Neural Networks*. MIT Press, Cambridge, MA.

Hebb, D. O. (1955). Drive and the c. n. s. *Psychological Review*, 62:243–254.

Hitchcott, P. K. and Philips, G. D. (1998). Double dissociation of the behavioural effects of r(+) 7-oh-dpat infusions in the central and basolateral amygdala nuclei upon pavlovian and instrumental conditioned appetitive behaviours. *Psychopharmacology*, 140:458–469.

Holland, P. C. and Gallagher, M. (1999). Amygdala circuitry in attentional and representational processes. *Trends in Cognitive Sciences*, 3(2):65–73.

Houk, J., Davis, J., and Beiser, D. (1995). *Models of information processing in the basal ganglia*. MIT Press, Cambridge, MA.

Kaelbling, L. P., Littman, M. L., and Moore, A. W. (1996). Reinforcement learning: A survey. *Journal of Artificial Intelligence Research*, 4:237–285.

Káli, S. and Dayan, P. (2000). The involvement of recurrent connections in area ca3 in establishing the properties of place fields: A model. *Journal of Neuroscience*, 20:7463–7477.

Kamin, L. J. (1968). Attention-like processes in classical conditioning. In Jones, M. R., editor, *Miami symposium on the prediction of behavior: aversive stimulation*, pages 9–31. University of Miami Press, Miami.

Kamin, L. J. (1969). Predictability, surprise, attention and conditioning. In Campbell, B. A. and Church, R. M., editors, *Punishment and aversive behavior*, pages 276–296. Appleton-Century-Crofts, New York.

Kehoe, E. J. (1982). Conditioning with serial compound stimuli: Theoretical and empirical issues. *Experimental Animal Behavior*, 1:30–65.

Kehoe, E. J. (1986). Summation and configuration in conditioning of the rabbit's nictitating membrane response to compound stimuli. *Journal of Experimental Psychology: Animal Behavior Processes*, 12:186–195.

Kelly, J. P. (1991). The neural basis of perception and movement. In Kandel, E. L., Schwartz, J. H., and Jessell, T. M., editors, *Principles of Neural Science*, pages 283–295. Prentice-Hall, London.

Kling, A. and Steklis, H. D. (1976). A neural substrate for affiliative behavior in nonhuman primates. *Brain, Behavior and Evolution*, 13:216–238.

Klopf, A. H. (1982). *The hedonistic neuron: a theory of memory, learning and intelligence*. Hemisphere, Washington, D. C.

Klopf, A. H. (1988). A neuronal model of classical conditioning. *Psychobiology*, 16(2):85–125.

Klopf, A. H., Morgan, J. S., and Weaver, S. E. (1993). A hierarchical network of control systems that learn: modelling nervous system function during classical and instrumental conditioning. *Adaptive Behavior*, 1(3):263–319.

Klüver, H. and Bucy, P. C. (1939). Preliminary analysis of functions of the temporal lobes in monkeys. *Arch. Neurol. Psychiatry*, 42:979–1000.

Kolb, B. and Whishaw, I. Q. (1990). *Fundamentals of human neuropsychology*. W. H. Freeman, New York.

Kupfermann, I. (1991). Learning and memory. In Kandel, E. L., Schwartz, J. H., and Jessell, T. M., editors, *Principles of Neural Science*, pages 997–1007. Prentice-Hall, London.

LaBar, K., Gatenby, J., Gore, J., LeDoux, J., and Phelps, E. (1998). Human amygdala activation during conditioned fear acquisition and extinction: a mixed-trial fmri study. *Neuron*, 20(5):937–945.

LeDoux, J. (2000). The amygdala and emotion: a view through fear. In Aggleton, J. P., editor, *The Amygdala: A functional analysis*, pages 289–310. Oxford University Press, Oxford, UK.

LeDoux, J. E. (1987). Emotion. In Plum, F., editor, *Handbook of Physiology, Section 1, Neurophysiology: Vol. 5. Higher Functions of the Brain*, pages 416–459. American Psychological Society, Bethesda, Maryland.

LeDoux, J. E. (1992). Emotion and the amygdala. In Aggleton, J. P., editor, *The amygdala: neurobiological aspects of emotion, memory, and mental dysfunction*, pages 339–351. Wiley, New York.

LeDoux, J. E. (1995). In search of an emotional system in the brain: leaping from fear to emotion and consciousness. In Gazzaniga, M. S., editor, *The cognitive neurosciences*, pages 1049–1061. MIT Press, Cambridge, MA.

LeDoux, J. E. (1996). *The Emotional Brain*. Simon and Schuster, New York.

LeDoux, J. E. and Fellous, J.-M. (1995). Emotion and computational neuroscience. In Arbib, M. A., editor, *The handbook of brain theory and neural networks*. MIT Press, Cambridge, MA.

Leonard, C. M., Rolls, E. T., Wilson, A. W., and Baylis, G. C. (1985). Neurons in the amygdala of the monkey with responses selective for faces. *Behavioral Brain Research*, 15:159–176.

Lewin, K. (1936). *Principles of topological psychology*. McGraw-Hill, New York.

Lieberman, D. A. (1990). *Learning: behavior and cognition*. Wadsworth Publishing Company, Belmont, CA.

Lorenz, K. and Tinbergen, N. (1938). Taxis und instinkthandlung in der eirollbewegung der graugans. *Zeitschrift für Tierpsychologie*, 2:1–29.

Lysle, D. T. and Fowler, H. (1985). Inhibition as a "slave" process: Deactivation of conditioned inhibition through extinction of conditioned excitation. *Journal of Experimental Psychology: Animal Behavior Processes*, 11:71–94.

Mackintosh, N. J. (1983). *Conditioning and associative learning*. Oxford Psychology Series No. 3. Oxford University Press, Oxford.

Marr, D. (1971). Simple memory: a theory for archicortex. *Phil. Trans. R. Soc. B*, 262:23–81.

McClelland, J. L., McNaughton, B. L., and O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, 102:419–457.

McLean, J. and Shipley, M. (1992). Neuroanatomical substrates of olfaction. In Serby, M. J. and Chobor, K., editors, *Science of Olfaction*. Springer Verlag, New York.

McNish, K., Gewirtz, J., and Davis, M. (2000). Disruption of contextual freezing, but not contextual blocking of fear- potentiated startle, after lesions of the dorsal hippocampus. *Behav Neurosci*, 114(1):64–76.

Medina, J. F., Repa, C., Mauk, M. D., and LeDoux, J. E. (2002). Parallels between cerebellum- and amygdala-dependent conditioning. *Nature Reviews Neuroscience*, 3:122–131.

Mishkin, M., Ungerleider, L. G., and Macko, K. A. (1983). Object vision and spatial vision: two cortical pathways. *Trends in Neuroscience*, 6:414–417.

Morén, J. (1998). Dynamic action sequences in reinforcement learning. In Pfeifer, R., Blumberg, B., Meyer, J., and Wilson, S., editors, *From animals to animats 5: Proceedings of the Fifth International Conference on Simulation of Adaptive Behavior*, pages 366–271, Cambridge, MA. MIT Press.

Morén, J. and Balkenius, C. (2000a). A computational model of emotional learning in the amygdala. In Mayer, J.-A., Berthoz, A., Floreano, D., Roitblat, H. L., and Wilson, S. W., editors, *From animals to animats 6*, pages 383–391. MIT Press, Cambridge, MA.

Morén, J. and Balkenius, C. (2000b). Reflections on emotion. In Trappl, R., editor, *Cybernetics and Systems 2000*, pages 699–703. Austrian Society for Cybernetic Studies.

Morgan, M. A. and LeDoux, J. E. (1999). Contribution of ventrolateral prefrontal cortex to the acquisition and extinction of conditioned fear in rats. *Neurobiol Learn Mem*, 72(3):244–251.

Morris, J. S., Öhman, A., and Dolan, R. J. (1999). A subcortical pathway to the right amygdala mediating "unseen" fear. In *Proceedings of the National Academy of Sciences*, volume 96, pages 1680–1685, USA.

Mowrer, O. H. (1973). *Learning theory and behavior*. Wiley, New York.

NeuroNames (2002). Neuronames and template atlas of the primate brain. http://rprcsgi.rprc.washington.edu/neuronames/.

Öhman, A. and Mineka, S. (2001). Fears, phobias, and preparedness: Toward and evolved module of fear and fear learning. *Psychological Review*, 108(3):483–522.

O'Keefe, J. (1990). A computational theory of the cognitive map. *Progress in Brain Research*, 83:301–312.

O'Keefe, J. and Nadel, L. (1978). *The hippocampus as a cognitive map*. Clarenon Press, Oxford.

Olton, D. S. (1986). Hippocampal function and memory for temporal context. In Isaacson, R. L. and Pribram, K. H., editors, *The Hippocampus*, volume 4, pages 281–298. Plenum Press, New York.

Olton, D. S. and Samuelson, R. J. (1976). Remembrance of places past: spatial memory in rats. *Journal of Experimental Psychology: Animal Behavior Processes*, 2:97–116.

O'Reilly, R. C. and Rudy, J. W. (2001). Conjunctive representations in learning and memory: Principles of cortical and hippocampal function. *Psychological Review*, 108(2):311–345.

Packard, M. G. and Teather, L. A. (1998). Amygdala modulation of multiple memory systems: Hippocampus and caudate-putamen. *Neurophysiology of Learning and Memory*, 69:163–203.

Panksepp, J. (1981). Hypothalamic integration of behavior. In Morgane, P. and Panksepp, J., editors, *Handbook of the hypothalamus*, volume 3. Part B. Deker, New York.

Pavlov, I. P. (1927). *Conditioned reflexes*. Oxford University Press, Oxford.

Pearce, J. M. and Hall, G. (1980). A model of Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review*, 87(6):532–552.

Perrett, D. I., Heitanen, J. K., Oram, M. W., and Benson, P. J. (1992). Organization and functions of cells responsive to faces in the temporal cortex. *Philos. Trans. R. Soc. Lond. [Biol.]*, 335:31–38.

Petrov, T., Jhamandas, J. H., and Krukoff, T. L. (1996). Connectivity between brainstem autonomic structures and expression of *c-fos* following electrical stimulation of the central nucleus of the amygdala in rat. *Cell Tissue Res*, 283:367–374.

Pitkänen, A. (2000). Connectivity of the rat amygdaloid complex. In Aggleton, J. P., editor, *The Amygdala: A functional analysis*, pages 31–116. Oxford University Press, Oxford, UK.

Poremba, A. and Gabriel, M. (1999). Amygdala neurons mediate acquisition but not maintenance of instrumental avoidance behavior in rabbits. *The Journal of Neuroscience*, 19(21):9635–9641.

Premack, D. (1965). Reinforcement theory. In Levine, D., editor, *Nebraska symposium on motivation*. University of Nebraska Press, Lincoln.

Rawlins, J. N. P. (1985). Associations across time: The hippocampus as a temporary memory store. *The Behavioral and Brain Sciences*, 8:479–496.

Redish, A. D. and Touretzky, D. S. (1997). Navigating with landmarks: computing goal locations from place codes. In Ikeuchi, K. and Veloso, M., editors, *Symbolic Visual Learning*. Oxford University Press, Oxford.

Rescorla, R. A. (1968). Probability of shock in the presence and absence of cs in fear conditioning. *Journal of Comparative and Physiological Psychology*, 66(1):1–5.

Rescorla, R. A. and Wagner, A. R. (1972). A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In Black, A. H. and Prokasy, W. F., editors, *Classical conditioning II: current research and theory*, pages 64–99. Appleton-Century-Crofts, New York.

Rieke, F., Warland, D., de Ruyter van Steveninck, R., and Bialek, W. (1997). *Spikes: exploring the neural code*. MIT Press, Cambridge, MA.

Rolls, E. (1992). Neurophysiology and functions of the primate amygdala. In Aggleton, J., editor, *The Amygdala: Neurobiological Aspects of Emotion, Memory and Mental Dysfunction*, pages 143–165, New York. Wiley-Liss.

Rolls, E. T. (1986). A theory of emotion, and its application to understanding the neural basis of emotion. In Oomura, Y., editor, *Emotions: neural and chemical control*, pages 325–344. Japan Scientific Societies Press, Tokyo.

Rolls, E. T. (1989a). Functions of neuronal networks in the hippocampus and neocortex in memory. In Byrne, J. H. and Berry, W. O., editors, *Neural models of plasticity*, pages 240–265. Academic Press, San Diego.

Rolls, E. T. (1989b). Information processing in the taste system of primates. *Journal of Experimental Biology*, 146:141–164.

Rolls, E. T. (1990). Functions of the primate hippocampus in spatial processing and memory. In Kesner, R. P. and Olton, D. S., editors, *Neurobiology of comparative cognition*, pages 127–155. Lawrence Erlbaum, Hillsdale, NJ.

Rolls, E. T. (1995). A theory of emotion and consciousness, and its application to understanding the neural basis of emotion. In Gazzaniga, M. S., editor, *The cognitive neurosciences*, pages 1091–1106. MIT Press, Cambridge, MA.

Rolls, E. T. (1999). *The Brain and Emotion*. Oxford University Press, Oxford, UK.

Rolls, E. T. (2000). Neurophysiology and functions of the primate amygdala, and the neural basis of emotion. In Aggleton, J. P., editor, *The Amygdala: A functional analysis*, pages 447–478. Oxford University Press, Oxford, UK.

Rolls, E. T. and Treves, A. (1998). *Neural Networks and Brain Function*. Oxford University Press, Oxford, UK.

Roozendaal, B., Sapolsky, R. M., and McGaugh, J. L. (1998). Basolateral amygdala lesions block the disruptive effects of long-term adrenalectomy on spatial memory. *Neuroscience*, 84(2):453–465.

Rosenzweig, M. R. and Leiman, A. L. (1982). *Physiological psychology*. D. C. Heath and Company, Lexington, MA.

Rudy, J. W. and O'Reilly, R. C. (1999). Contextual fear conditioning, conjunctive representations, pattern completion, and the hippocampus. *Behavioral Neuroscience*, 113:867–880.

Rudy, J. W. and O'Reilly, R. C. (2001). Conjunctive representations, the hippocampus and contextual fear conditioning. *Cognitive, Affective, and Behavioral Neuroscience*, 1:66–82.

Rudy, J. W. and Sutherland, R. J. (1995). Configural association theory and the hippocampal formation: and appraisal and reconfiguration. *Hippocampus*, 5(5):375–389.

Sanghera, M. K., Rolls, E. T., and Roper-Hall, A. (1979). Visual responses of neurons in the dorsolateral amygdala of the alert monkey. *Experimental Neurology*, 63:610–626.

Schachter, S. (1964). The interaction of cognitive and physiological determinants of emotional state. In Berkowitz, L., editor, *Advances in experimental social psychology*, volume 1. Academic Press, New York.

Schachter, S. (1970). Some extraordinary facts about obese humans and rats. *American Psychologist*, 26:129–144.

Schmajuk, N., Holland, P. C., and Association, A. P. (1998). *Occasion setting : associative learning and cognition in animals*. APA science volumes. American Psychological Association, Washington, DC.

Schmajuk, N. A. (1997). *Animal Learning and Cognition: a neural network approach*. Cambridge University Press, Cambridge, United Kingdoms.

Schmajuk, N. A. and DiCarlo, J. J. (1992). Stimulus configuration, classical conditioning, and hippocampal function. *Psychological Review*, 99:268–305.

Schneiderman, N. (1966). Interstimulus interval function of the nictitating membrane response of the rabbit under delay versus trace conditioning. *Journal of comparative and physiological psychology*, 62:397–402.

Schneiderman, N. and Gormezano, I. (1964). Conditioning of the nictitating membrane of the rabbit as a function of the cs-us interval. *Journal of comparative and physiological psychology*, 57:188–195.

Schoenbaum, G., Chiba, A. A., and Gallagher, M. (1998). Orbitofrontal cortex and basolateral amygdala encode expected outcomes during learning. *Nature Neuroscience*, 1(2):155–159.

Schoenbaum, G., Nugent, S., Saddoris, M., and Setlow, B. (2002). Orbitofrontal lesions in rats impair reversal but not acquisition of go, no-go odor discriminations. *Neuroreport (In Press)*.

Schultz, W. (1998). Predictive reward signal of dopamine neurons. *Journal of Neurophysiology*, 80:1–27.

Schultz, W., Tremblay, L., and Hollerman, J. R. (1998). Reward prediction in primate basal ganglia and frontal cortex. *Neuropharmacology*, 37:421–429.

Sheffield, F., Wulff, J., and Backer, R. (1951). Reward value of copulation without sex-drive reduction. *Journal of Comparative and Physiological Psychology*, 44:3–8.

Shi, C. and Davis, M. (1999). Pain pathways involved in fear conditioning measured with fear-potentiated startle: Lesion studies. *The Journal of Neuroscience*, 19(1):420–430.

Shimamura, A. P. (1995). Memory and frontal lobe function. In Gazzaniga, M. S., editor, *The cognitive neurosciences*, pages 803–813. MIT Press, Cambridge, MA.

Siqueland, E. R. and DeLucia, C. A. (1969). Visual reinforcement of nonnutritive sucking in human infants. *Science*, 165:1144–1146.

Skinner, B. F. (1974). *About behaviorism*. Vintage Book, New York.

Smith, M. and Gormezano, I. (1965). Effects of alternating classical conditioning and extinction sessions on the conditioned nictitating membrane response of the rabbit. *Psychonomic Science*, 3:91–92.

Smith, M. C., Coleman, S. R., and Gormezano, I. (1969). Classical conditioning of the rabbit[1]s nictitating membrane response at backward, simultaneous, and forward cs-us interval. *Journal of Comparative and Physiological Psychology*, 69:226–31.

Solomon, P. R. (1979). Temporal versus spatial information processing theories of hippocampal function. *Psychological Bulletin*, 86(6):1272–1279.

Solomon, R. L. (1980). The opponent process theory of acquired motivation. *American Psychologist*, 35:691–712.

Spence, K. W. (1956). *Behavior theory and conditioning*. Yale University Press, New Haven.

Sutton, R. S. (1995). TD models: modeling the world at a mixture of time scales. In *Proceedings of the 12th International Conference on Machine Learning*. Morgan Kaufman.

Sutton, R. S. and Barto, A. G. (1981). Toward a modern theory of adaptive networks: expectation and prediction. *Psychological Review*, 88(2):135–170.

Sutton, R. S. and Barto, A. G. (1987). A temporal-difference model of classical conditioning. In *Proceedings of the Ninth Annual Conference of the Cognitive Science Society*, pages 355–378, Hillsdale, NJ. Erlbaum.

Sutton, R. S. and Barto, A. G. (1990). Time-derivative models of pavlovian reinforcement. In Gabriel, M. and Moore, J., editors, *Learning and computational neuroscience: foundations of adaptive networks*, pages 497–538. MIT Press, Cambridge, MA.

Sutton, R. S. and Barto, A. G. (1998). *Reinforcement learning: an introduction*. MIT Press, Cambridge, MS.

Tabert, M. H., Borod, J. C., Tang, C. Y., Lange, G., Wei, T. C., Johnson, R., Nusbaum, A. O., and Buchsbaum, M. S. (2001). Differential amygdala activation during emotional decision and recognition memory tasks using unpleasant words: an fMRI study. *Neuropsychologia*, 39:556–573.

Thompson, C. I. (1980). *Controls of eating*. Spectrum, New York.

Tolman, E. C. and Honzik, C. H. (1930). Introduction and removal of reward, and maze performance in rats. *University of California publications in psychology*, 4:227–272.

Tsvetkov, E., Carlezon, Jr., W. A., Benes, F. M., Kandel, E. R., and Bolshakov, V. Y. (2002). Fear conditioning occludes ltp-induced presynaptic enhancement of synaptic transmission in the cortical pathway to the lateral amygdala. *Neuron*, 34:289–300.

Waelti, P., Dickinson, A., and Schultz, W. (2001). Dopamine responses comply with basic assumptions of formal learning theory. *Nature*, 412:43–48.

Wagner, A. R., Siegel, S., Thomas, E., and Ellison, G. D. (1964). Reinforcement history and the extinction of a conditioned salivary response. *Journal of Comparative and Physiological Psychology*, 58:354–358.

Watkins, C. J. C. H. (1992). Q-learning. *Machine Learning*, 8:279–292.

Watson, J. (1930). *Behaviorism*. Norton, New York.

Weinberger, N. M. (1995). Retuning the brain by fear conditioning. In Gazzaniga, M., editor, *The Cognitive Neurosciences*, pages 1071–1089. MIT Press, Cambridge, MA.

Weinberger, N. M. (1998). Tuning the brain by learning and by stimulation of nucleus basalis. *Trends in Cognitive Science*, 2(9):271–273.

Weiskrantz, L. (1956). Behavioral changes associated with ablation of the amygdaloid complex in monkeys. *Journal of Comparative Physiological Psychology*, 49:381–391.

Whitelaw, R. B., Markou, A., Robbins, T. W., and Everitt, B. J. (1996). Excitotoxic lesions of the basolateral amygdala impair the acquisition of cocaine-seeking behaviour under a second-order schedule of reinforcement. *Psychofarmacology*, 127:213–224.

Widrow, B. and Hoff, M. E. (1960/1988). Adaptive switching circuits. In Anderson, J. A. and Rosenfeld, E., editors, *Neurocomputing: foundations of research*, pages 123–134. Mit Press, Cambridge, MA.

Wilson, F. A. W. and Rolls, E. T. (1993). The effects of stimulus novelty and familiarity on neuronal activity in the amygdala of monkeys performing recognition memory tasks. *Experimental Brain Research*, 93:367–382.

Zipser, D. (1985). A computational model of hippocampal place fields. *Behavioral Neuroscience*, 99:1006–1018.

Zipser, D. (1986). Biologically plausible models of place recognition and goal location. In Rumelhart, D. E. and McClelland, J. L., editors, *Parallel distributed processing*, volume 2, pages 432–470. MIT Press, Cambridge, MA.