A Model of Prefrontal-Hippocampal Interactions in Strategic Recall

by

Jean C. Lim

A thesis submitted in conformity with the requirements
for the degree of Master of Arts
Graduate Department of Psychology
McMaster University

MASTER OF SCIENCE (2000) McMaster University
(Psychology) Hamilton

TITLE: A Model of Prefrontal-Hippocampal Interactions in Strategic Recall

AUTHOR: Jean C. Lim, B.A. (Bryn Mawr College)

SUPERVISOR: Suzanna Becker, PhD.

NUMBER OF PAGES: 116

Acknowledgements

This thesis was written with the help of a number of people. Many thanks go to my advisor, Sue Becker for her enthusiasm, ideas and encouragement. My gratitude also extends to Bruce Milliken, Ron Racine, and Judy Shedden for their insight and understanding. Their patience was very appreciated.

I would also like to thank my colleagues Gaurav Patel, Arnold Liwanag and Steve Howell for their advice. Special thanks to Damian Jankowicz and Chris Lesner for proofreading. The lively, thought-provoking discussions were enjoyable and I appreciated your help in making the indecipheral understandable.

I thank my parents, Mali, and Andy for their love, support and for bringing me to this point in life. Last, but most of all, for providing all of the above and much more, I thank Kingsley Cudjoe.

Abstract

In this thesis, we look at evidence accumulated on the prefrontal cortical and hippocampal regions of the brain and review theories about the possible roles each structure has on human memory and behaviour. Aspects of these theories are tested via a self-reinforcing computational network model. We propose this model may simulate the underlying mechanisms or processes of the prefrontal-hippocampal interaction during performance of memory tasks that require intact prefrontal and hippocampal structures.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Our knowledge of memory depends upon the type of tests used to assess it. Memory can be assessed by explicit tests that query the individual's knowledge of past events directly. Explicit tests of episodic memory include variations of tests of recall, recognition, and paired associate learning that depend on conscious recollection. Memory can also be assessed by implicit tests. These are tests that examine the effects past experiences have on current performance of tasks that can be completed without conscious recollection. Solving a crossword puzzle a second time, as compared to the first, is an example of an implicit test of memory [58]. Memory is revealed by the facilitated performance as assessed by speed and accuracy for the second solution in comparison to the first. Although the subject may remember having attempted the puzzle, conscious recollection need not accompany all the solutions if enough time elapsed between the first and second attempt. In this thesis, we focus on explicit tests of memory and the processes and brain regions possibly involved in the performance of these tasks.

Due to the observed differential effects of brain damage on tests of memory, it has been suggested that multiple brain systems are involved in memory [78, 67, 52, 56]. Patients with bilateral damage to the medial temporal lobes - more specifically the hippocampus and its adjacent perirhinal and parahippocampal cortices - are severely impaired on some explicit tests of memory but have relatively normal implicit memory. These are the hallmark characteristics of amnesia: post-traumatic events are quickly forgotten and the patient cannot consciously recollect newly-acquired information [52, 58, 69, 97, 32]. For example, in amnesics the ability to learn a new motor skill (e.g., increased speed in tracing figures) and capability for priming remain intact. In a word fragment priming task, patients are presented a list of written words. Later, they are given a list of word fragments and asked to form real words from those fragments. There is a tendency in both normal subjects and amnesics to form the words that were previously presented to them [58]. This suggests

that amnesics possess a memory for the list of words presented to them or a memory for the motor skill that was just learned. However, when amnesics are asked to recall the words on the list or to describe the motor skill just learned, they are unable to do so. In short, evidence suggests that the acquisition of new motor skills and the capacity for priming are types of learning and memory that fall outside the domain of the hippocampal region which is damaged in amnesics, but the storage and retrieval of episodic, explicit memories - the memory for when an event occurred and who and what was involved in that event - are properties of the medial temporal lobes [78, 77, 89]. Various other functions have been attributed to the hippocampal region and some of these are described in Chapter 3.

Retrieval of episodic memories is believed by many to be modulated by the frontal lobes [33, 74, 64, 47, 15, 57, 69, 80, 6, 8]. It has been suggested that the prefrontal cortex is used to create retrieval strategies and to self-organize information - these abilities are most evident in performance on tasks of free recall [87, 73]. For example, experiments have shown that normal subjects will recall randomly presented lists of semantically related words in sequences of related words or clusters [7, 87, 31]. Repeated presentations of the word list increase clustering in recall and there is a direct relation between the degree of clustering and the number of correctly recalled words [34, 88, 31]. Even when words within the study list are not related, normal subjects will "impose a higher order organizational structure to the list that assists retrieval in the absence of external cues" [89]. That is, normal subjects organize their recall sequentially even in the absence of obvious organization in the stimulus lists, and this organization increases systematically with repeated exposures to, and recall of the material. This increasing sequential organization of words over trials is imposed upon the material by the learner, hence Tulving's label 'subjective organization' for the phenomenon. Studies comparing frontal lobe damaged patients to control subjects on semantically related word lists showed that control subjects exhibited more semantic clustering of words over repeated recall trials and more subjective organization - that is, an increasing sequential organization of words - during recall [79, 23, 73]. This suggested that the frontal cortex was involved in the use of search and retrieval processes. The frontal lobes have been attributed a variety of other functions and these are described in Chapter 3.

Although both the prefrontal cortex and the hippocampal region seem to play key roles in the encoding and retrieval of episodic memories, one can easily observe the different effects that damage to either region has on memory. Janowsky, Shimamura, et. al. (1989) found that patients with frontal lobe damage still do well on standardized memory tests on which amnesics do poorly. They studied patients with circumscribed frontal lobe lesions, patients with Korsakoff's syndrome (with damage to both medial temporal and frontal regions of the brain), non-Korsakoff amnesic patients and control subjects. Patients from the frontal lobe

damaged group performed similarly to patients from the medial temporal lobe damaged group on free recall tasks and tests of global intellectual function (the WAIS-R). Both groups performed well on the WAIS-R and badly on free recall. However, patients with frontal lobe damage could not be characterized as amnesic - they still performed well on other memory tests that classic amnesic patients do poorly on (the Wechsler Memory Scale, Wechsler Memory Scale-Revised, Story Recall, Complex Figure Recall, Paired Associate Learning and recognition of Word lists).

Frontal lobe damaged patients are also capable of performing at near normal levels on simple word-associate learning tasks (e.g., learning to associate the word "Window" to "Reason") [33] but they show an inordinate amount of proactive interference on tests of paired word-associate learning using the AB-AC paradigm introduced by Barnes & Underwood (1959) . Shimamura et al. (1995) presented a list of 12 related word pair associates (e.g., thief-crime; lion-hunter). After subjects learned those word pairs, proactive interference was increased by having subjects learn a second list in which each cue word used in the first list was paired with a new target word (e.g., thief-bandit; lion-circus). Although frontal damaged patients appeared to learn the first list as well as control subjects, they showed a disproportionate impairment when required to ignore the first associations and learn new ones [74].

In summary, the hippocampal region has long been recognized as a system involved with episodic memory. More recent studies have indicated that frontal cortical processes may also be engaged during retrieval of episodic information during some memory tests. However, the fact that frontal lobe damaged patients do well on recognition memory but show significant impairment on comparable tests of free recall for the same items have led some people to suggest that the prefrontal cortex is involved in the use of search and retrieval processes: free recall puts heavy demands on internally generated memory strategies and impaired free recall performance is associated with decreased use of memory organizational strategies [73, 87, 34, 79]. The prefrontal cortex also appears to play a role in creating and separating context - as evident in AB-AC list learning experiments [74].

In this thesis, we propose a computational model that focuses on the role of the prefrontal cortex and how this region may interact with other memory systems, namely the hippocampal regions, during performance of explicit memory tests - more specifically, during free recall and paired-associate AB-AC memory tasks.

There are many benefits in approaching this problem from a computational modeling perspective:

- Through implementation of simple models, computational modeling forces researchers to avoid the ambiguity of verbal theorizing. As Broadbent (1987) once complained:

  Terms are used such as "access to the lexicon," "automatic processing," "central execu-

> *tive," "resources"; formal definitions of such terms are rare, and even rarer are statements of the rules supposed to be governing their interaction. As a result one is left unclear about exactly what kinds of experimental data would invalidate such theories and whether or not they are intended to apply to some new experimental situation.*

Broadbent's point is that we need to tighten up our explanatory theories. Although it was theorized that the prefrontal cortex 'organizes' and 'strategically retrieves information', the question of how the frontal cortex does this still begs to be answered. Does it provide constraining cues to the other memory systems so that the necessary information is output? Or does it filter information from the other memory systems and sift desired information from the unwanted? Also important is how would such a 'constraining cue' or filtering system work? Or do other systems or alternative explanations need to be considered? Computational modeling forces us to examine a theory step by step. If crucial information is missing or garbled, the program will not run.

- Simulation models can be designed to work in a parallel, distributed manner and assume nonlinear or exponential activation functions. Experience with this type of model system can help one to develop new intuitions about the behaviour of systems having properties such as parallelism and nonlinearity. These are believed to be properties of the brain and are intuitively difficult to understand. In our case, we hope to understand how an aggregate of individual networked cells or units might exhibit organized encoding and strategic retrieval behaviours.

- Modeling can be used to question theories. For example, consider how memory for an event is influenced by the interpolation of conflicting information between the original learning and the test. On a forced-choice recognition test, subjects must choose between the original and the interpolated. Typically, subjects show poorer recognition memory than do controls who didn't see the interpolated material. The standard view in cognitive psychology was that the inconsistent information is either incorporated destructively into the original memory trace or interferes with its retrieval. However, McCloskey and Zaragoza (1985) showed that such a result can be consistent with a simple Markov model that assumes coexistence and noninterference between traces of the original and interpolated events.

- Modeling may lead to new theories and predictions. For example, consider the view put forward by McClelland, McNaughton & O'Reilly (1995) on the interaction between the hippocampus and neocortex. Previous computational modeling attempts had shown that even with multiple layers of units (analogous to layers of neurons), memory for previously learned material would be catastrophically

forgotten at the expense of learning new information. Although subjects typically show reduced memory for old information as new knowledge is acquired, the forgetting of previous information is not so severe as that observed in computational models. Modeling experiments showed that with 'interleaved' learning trials - that is, the continued presentation of old material interleaved among presentation of new materials - the models were capable of retaining both old and new information. This led to the theory put forward by McClelland et al. that the hippocampal region acts as a storage device that 'teaches' old and new information to the cortex in an interleaved fashion [48]. Another example is the theoretical advances in the analysis of memory for serial order as a result of Murdock's TODAM model [29]. Ebbinghaus (1913) proposed the simple view that word items on a list are connected by pairwise associations, like beads on a chain, and that retrieval proceeds in a given order from one end of the chain to the other. Like a chain, however, if the retrieval chain is interrupted because an item is forgotten then none of the following information should be retrievable. This was an unreasonable prediction falsified by hundreds of studies that show consistent serial position effects (i.e., elevated recall probabilities for terminal items at both beginning and ending of a list). As a result, the chaining theory was dropped and other theories emerged (e.g., Shiffrin and Cook (1978) proposed a relational model). Due to the difficulties facing theoretical developments, interest in serial order effects waned for many years. It was the distributed representational assumptions of TODAM, which arose from more basic considerations of how memory might work, that led to a more successful and intriguing reformulation of the classic chaining view. TODAM could predict a variety of findings in serial order memory, ranging from recall of short lists and build-up and release from proactive interference, to subtle effects of variables such as word frequency or presentation modality.

Before leaving the subject of computational models in general, I will differentiate between simplifying and realistic models. Simplifying models are generally used as tools to discover the principles of operation of systems, and so most of the details are abstracted away. Realistic models include most of the parameters known about the systems at the level of organization used in the models. They are generally used to explore parameter sensitivity (Sejnowsky et. al, 1988). The criticism has often been made that the simplifying neurobiological models investigated in neural network research are too abstract to give any valid insights about the brain. Implicit in this criticism is that realistic models would be more useful. However, as Sejnowsky et al. (1988) have noted, as we do not have such a complete understanding of the brain, a realistic model would contain so many free parameters that no specific predictions would be generated. There is also the danger that any completely specific model will become just as complex and as difficult to analyze as the brain itself. These considerations limit the usefulness of realistic models. However, with the

appropriate assumptions, simplifying models can be powerful explanatory tools [94].

Our intent was to create a simple model that would capture a few key features that we believe are characteristic of the prefrontal cortex and to engender a better understanding of the complex processes that occur between the prefrontal cortex and hippocampal regions during the retrieval of memory. With the tasks that our model was to simulate - the California Verbal Learning Test (CVLT) and the AB-AC list learning task - our simplistic prefrontal and hippocampal modules had to be capable of clustering temporally separated but related pieces of information, subjective organization, and maintaining contextual distinctions between lists of words [14, 3, 75]. Inherent in this design is the assumption that a common mechanism underlies the wide variety of dysfunctions often observed in frontal lobe damaged patients. This idea is not new [39] and the results of our simulations further support this view.

Chapters 2 and 3 provide the relevant background necessary for understanding our proposed model. Chapter 2 provides a preliminary introduction to the standard notations used in the neural network modeling literature and a detailed description of the architectures and learning algorithms as they are commonly used in this field and on which our model design is based. Chapter 3 reviews the anatomical, experimental and neuropsychological evidence on the hippocampal region, the prefrontal cortex and their interaction.

In Chapter 4 we discuss other computational models of the prefrontal cortex then describe in detail the architecture and the components of our model and the learning algorithms we used. The architecture and learning rules are based on algorithms (described in Chapter 2) often used by computational modelers. However, we modified the learning rules substantially and these modifications are described in this chapter.

Chapter 5 describes the California Verbal Learning test and the AB-AC list learning test paradigms and our computational model simulations. The first two experiments test individual components of the model. The third examines the performance of the whole model on a simple version of the CVLT and compares different variations of the learning algorithm: one learns during recall and study, the other learns only during study. The fourth tests the model on a more difficult version of the CVLT. The fifth experiment examines performance of the model on two versions of the AB-AC list learning test: one where the AC list is semantically similar to the AB (first) list, the other where the AC list is semantically different. This chapter is organized such that each simulation is followed by a report of the results and a brief discussion.

In chapter 6 we conclude with the overall implication of our results and suggestions for future work. Although our model did not perform as expected in some aspects, it did meet and even exceeded others. On the down side it showed gradual unlearning of the first list afer learning the second in the AB-AC list learning test. This is, as will be discussed in the chapter, a common problem that has faced many computational models that use distributed representations. On the up side, it was capable of independently creating context

and discriminating between the different lists. Context was developed by the network through the selection and strengthening of a few frontal 'neurons'. This contextual environment was self-generated by the model: we did not have to specify what particular units should become active under what specific conditions. On the CVLT, the network was able to cluster list words, temporally organize items it learned over trials and maintain recall consistency between trials despite the absence of external reinforcement. We conclude with possible directions for future research.

Chapter 2

The Foundations

In this chapter we provide the relevant background of artificial neural networks necessary for understanding our model, described in greater detail later in this thesis (see Chapter 4), and a brief description of the anatomical properties of the hippocampal region, the prefrontal cortex and the pathways by which they are thought to interact.

2.1 Neural Networks

In general, our model consists of three modules:

1. A *perceptual module* that performs word identification and acts as both the input and output layers of our model;

2. A *hippocampal module* that maintains an aspect of episodic memory (e.g., word-word associations);

3. A *prefrontal module* that permits strategic encoding and recall by learning then imposing constraining memory access cues to the hippocampal region.

These are described in more detail in Sections 4.3, 4.4 and 4.5. But in brief, our hippocampal module is represented by a recurrent attractor network - in particular, one that is known as a Hopfield network. This type of network can be trained and used as a pattern associator and will be further described in Section 2.1.11. The prefrontal cortex of our model is represented by a layer of units that are trained with a reinforcement learning rule. This learning rule allows the network to learn the best responses through trial and error despite delayed rewards. The learning algorithms and the pertinent standard architectures used

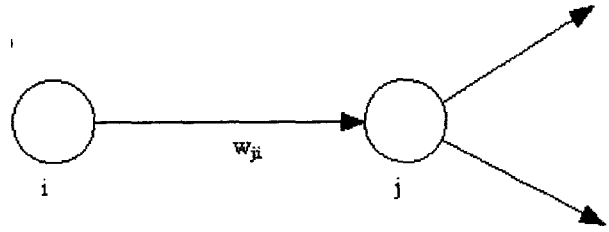

Figure 2.1: *The building blocks of a neural network. A single unit, j, has incoming connections. Its output is computed as some function f of its total weighted summed input and is transmitted along outgoing connections.*

in artificial neural networks are reviewed in detail in this chapter, but first a preliminary introduction of the basics and standard notations is in order.

2.1.1 Preliminary Notations

A neural network consists of parallel processing elements (units) connected to each other by directed links. The units receive input from other units, compute an output activation, and transmit a signal along their outgoing connections to other units. As illustrated in Figure 2.1, a connection from the ith to the jth unit has a weight w_{ji} which determines the effectiveness with which the ith unit can transmit its output signal to the jth unit.

The output of the jth unit, y_j, is computed as some function f of its total weighted summed input x_j:

$$y_j = f(x_j) \tag{2.1}$$

where

$$x_j = \sum_i w_{ji} y_i + b$$

where b is the unit's bias. To simplify the equation, the bias may be eliminated by creating an extra input unit with an activation value of 1 whose outgoing weight is trained in a manner similar to that used for the other weights. Henceforth, we will treat the bias in this manner, allowing us to omit the explicit bias, b, from the expression for x_j.

A common choice for f, and the function we use in our model, is the sigmoidal or S-shaped logistic function:

$$f(x_j) = \frac{1}{1 + e^{-x_j}} \tag{2.2}$$

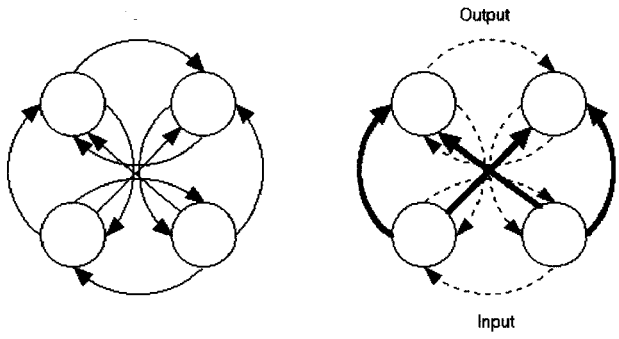

Figure 2.2: *The figure on the left (a) is a fully connected network. Each unit is connected to every other unit with positive or negative connection weights. The figure on the right (b) depicts a single-layer feedforward network. It may be considered to be a special case of the fully connected network where some connection weights are set to zero (denoted by dashed lines). The feedforward network has different sets of units that compose the input and output layers.*

This function limits the output of the unit to a minimum and maximum value of 0 and 1.

Computational models may assume stochastic binary valued units. These are units that usually output values of 0 or 1 (or -1/1) and can be randomly activated according to some probability function of the total input. Alternatively, each unit can take on any real, continuous value in the range 0 to 1 (or -1 to 1)[1]. The vector patterns in our model are continuous and centered around a zero mean which causes each unit to take on any real value in the range [-1,+1].

2.1.2 Architecture

The way units in a network are connected constitutes an important early design decision for the modeler. In a fully connected network, each unit is connected to every other unit. Connections may be excitatory (positive weights), inhibitory (negative weights), or irrelevant (zero weights). Every other architecture is actually a special case of a fully connected network and is obtained by setting some weights to zeros (Figure 2.2). Despite its generality and simplicity the fully connected network is seldom used due to the large number of degrees of freedom involved: in a network with n nodes, there are n^2 weights. It is difficult to devise learning schemes that work efficiently with such a large network. Furthermore, a biologically plausible design is desirable when modeling brain processes. A fully connected network is not biologically plausible because neurons rarely establish direct synapses with distant neurons.

The input units of a network are those whose pattern of activity, I, are determined by an environmental or external signal. When the values of units are held constant by some external influence, we say that their

[1]In neurobiology, the frequency with which a neuron fires is analogous to the unit's activation level.

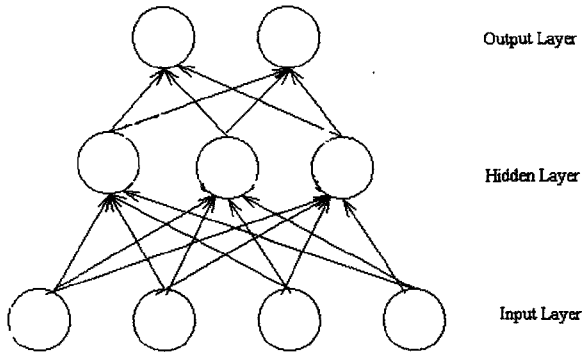

Figure 2.3: *A multilayered network consists of an input layer, an output layer and at least one hidden layer of units. This figure depicts a fully connected feed-forward network - each layer projects to the subsequent layer.*

activities are 'clamped' to those values. The output units are those whose pattern of activity, O, represents the output of the network. In our model, a presentation of an input pattern involves clamping the activities of the n input units to the n elements of the input pattern vector I for some period of time, and allowing this activation to propagate, weighted by the connection strengths, to the rest of the network.

A layered, feedforward network has its nodes partitioned into subsets called layers and each layer receives input from the preceding layer. Single layered networks consist of input units that project to a single output layer (Figure 2.2). Multilayered networks are formed by adding hidden layers of units between the input and output layers (Figure 2.3).

A recurrent network involves feedback. A unit in a recurrent network may feed its output signal to itself (self-feedback) and/or to the inputs of all the other units. The Hopfield network is a special case of a recurrent network. In the Hopfield net, each unit is connected to every other unit with bidirectional links but no unit has a recurrent connection to itself (Figure 2.4). Furthermore, the weights are *symmetric*. The connections to and from units i and j always carry the same weight (Conversely, in an *asymmetric* network, $w_{ij} \neq w_{ji}$. That is, the connection from unit i to unit j may carry a different weight than the connection from unit j to i).

2.1.3 Supervised versus Unsupervised Learning

Connectionist learning paradigms can be roughly divided into two approaches: supervised and unsupervised. A supervised network is one in which, during training, the network is 'taught' what output should be produced. Initially, the network's output is a function of randomly set initial weights. When the input pattern and desired response are provided, the network compares its calculated actual response with the

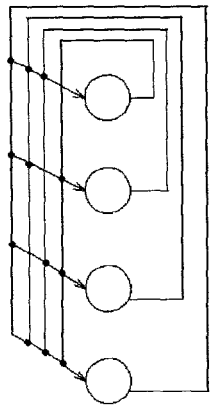

Figure 2.4: *In the recurrent network shown here, units are fully interconnected except that there is no self-feedback among the units. That is, each unit gives inputs to all other units except itself.*

desired (correct) response and adjusts its weights so the network's response in the future will be more desirable. Supervised learning is appropriate when the desired output is known for every input. However, a common criticism raised by modelers of brain functions is that exclusive use of supervised learning is not plausible: networks that use this type of learning are restricted to problems for which a teacher is available.

In unsupervised learning, input-output examples are not provided. Instead, the network uses an 'internal', task-independent measure of performance as the basis of learning. Rather than referring to an external performance measure, the connections are defined in terms of an equation for weight updates. An example of such an update rule is the Hebbian learning rule. This rule, expressed by neuropsychologist Donald Hebb, states:

> *When an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic changes take place in one or both cells such that A's efficiency as one of the cells firing B, is increased.*
>
> [*Hebb, 1949*]

Hebb had proposed this to be a mechanism of associative learning at the cellular level which would result in an enduring modification in the activity pattern of a spatially distributed 'assembly of cells'. Since its conception, Hebb's rule has been expanded (Stent, 1973 and Changeux & Danchin, 1976 as cited by Haykin, 1994):

- If two neurons are activated simultaneously then the strength of the connection between those two neurons is increased.

- If two neurons are activated asynchronously then the strength of the connection between those two neurons is selectively weakened or eliminated.

Although Hebb did not provide a formula for his rule, the following expression has been widely used to implement his postulate and its extensions:

$$w_{ij}(t+1) = w_{ij}(t) + \alpha x_i(t) y_j(t) \tag{2.3}$$

where $w_{ij}(t+1)$ and $w_{ij}(t)$ are the 'new' and 'old' values of the weight w_{ij} respectively, α is a positive constant determining the rate of learning, and $x_i(t)$ and $y_j(t)$ are the activities of the pre- and post-synaptic neurons (units) connected by w_{ij} at time t.

Another method of training network weights is analogous to Thorndike's Law of Effect. Thorndike posited the following as a theory of learning:

> *Of several responses made to the same situation, those which are accompanied or closely followed by satisfaction to the animal will, other things being equal, be more firmly connected with the situation, so that, when it recurs, they will be more likely to recur; those which are accompanied or closely followed by discomfort to the animal will, other things being equal, have their connections with that situation weakened, so that, when it recurs, they will be less likely to occur. The greater the satisfaction or discomfort, the greater the strengthening or weakening of the bond.*
>
> [*Thorndike, 1911*]

Thorndike's law corresponds with 'Reinforcement Learning' methods in neural networks:

> *If an action taken by a learning system is followed by a satisfactory state of affairs, then the tendency of the system to produce that particular action is strengthened or reinforced. Otherwise, the tendency of the system to produce that action is weakened.*
>
> [*Sutton, 1998*]

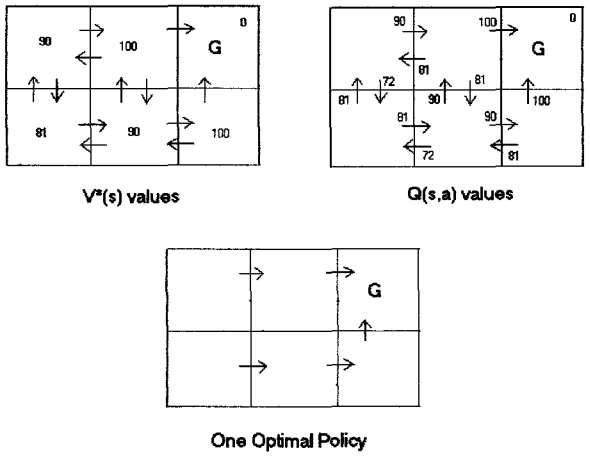

Figure 2.5: *A simple, deterministic grid-world to illustrate the basic concepts of reinforcement learning. Each grid square represents a distinct state, each arrow a distinct action. The immediate reward function, r(s,a) gives reward 100 for actions entering the goal state G, and zero otherwise. Values of* $V^*(s)$ *and* $Q(s,a)$ *follow from* $r(s,a)$*, and the discount factor* $\gamma = 0.9$*. The bottom gridworld depicts one optimal policy, corresponding to actions with maximal Q values. (Examples and figures adapted from Mitchell, 1997).*

2.1.4 Reinforcement Learning

Reinforcement learning is a combination of supervised and unsupervised learning in that it does not require detailed information about the errors made by individual units yet it still requires feedback about the 'goodness', or value, of performing certain actions.

In this section we describe reinforcement learning and its application to control optimization. Although the learning algorithm we used in our model is different from the equations we describe in this section and how it is traditionally applied, the core concepts are similar and may help the reader to understand the learning algorithm we used nonetheless.

In general, reinforcement learning deals with how an autonomous agent that senses and acts in its environment learns to choose optimal actions to achieve its goals. To illustrate the concepts that will follow, we present a simple example. Imagine you have an agent in a grid-world environment. Any one of the diagrams in Figure 2.5 depicts a sample grid-world. The six grid squares represent six possible states, or locations, for the agent. Each arrow in the top diagrams represents a possible action the agent can take to move from one state to another.

One way to define the goal of the agent is by this *reward* function that assigns a numerical value - an immediate payoff - to each distinct action the agent may take from each state. In reinforcement learning,

the task of the agent is to perform sequences of actions, observe their consequences and choose actions that maximize the reward accumulated over time.

In our example, the agent is allowed to perceive the set of states and the set of actions that it can perform. At each discrete time step t, the agent senses the current state s_t, chooses a current action a_t and performs it. In return, the agent receives a reward $r_t = r(s_t, a_t)$ and the succeeding state $s_{t+1} = \sigma(s_t, a_t)$.

The agent's task is to learn a *policy*, $\pi : S \rightarrow A$, for selecting its next action a_t based on the current observed state s_t; that is, $\pi(s_t) = a_t$. More precisely, the cumulative value $V^\pi(s_t)$ achieved by following an arbitrary policy π from an arbitrary initial state s_t is:

$$V^\pi(s_t) \equiv r_t + \gamma r_{t+1} + \gamma^2 r_{t+2} + ... \equiv \sum_{i=0}^{\infty} \gamma^i r_{t+i} \tag{2.4}$$

where the sequence of rewards r_{t+i} is generated by beginning at state s_t and by repeatedly using the policy π to select actions as described above. Here $0 \leq \gamma \leq 1$ is a constant that determines the relative value of delayed versus immediate rewards. Rewards received i time steps into the future are discounted exponentially by a factor of γ^i. Note that if we set $\gamma = 0$, only the immediate reward is considered. As we set γ closer to 1, future rewards are given greater emphasis relative to the immediate reward. It is reasonable to discount future rewards relative to immediate rewards because, in many cases, we prefer to obtain the reward sooner rather than later. The quantity $V^\pi(s)$ defined by Equation 2.4 is often called the *state value* or the ***discounted cumulative reward*** achieved by policy π from initial state s.

The agent must learn the policy that produces the greatest possible cumulative reward over time (such a policy is called an ***optimal*** policy, denoted here by π^*). This is the policy π that maximizes $V^\pi(s)$ for all states s:

$$\pi^* \equiv \overset{argmax}{\pi} V^\pi(s), (\forall s) \tag{2.5}$$

To simplify notation, the value function $V^{\pi^*}(s)$ will be referred to as $V^*(s)$. $V^*(s)$ gives the maximum discounted cumulative reward that the agent can obtain starting from state s when following the optimal policy.

Once the states, actions and immediate rewards are defined, and a value for the discount factor γ is chosen, we can determine the optimal policy π^* and its value function $V^*(s)$. The state labeled G (in Figure 2.5) is called the goal state. Once the agent enters state G, the only action available is to remain there. For this example, let us define the immediate reward to be zero for all state-action transitions except for those that lead to G whereupon it is given a reward of 100. In other words, the only way the agent can receive a

reward is by entering this state. Let us also define $\gamma = 0.9$.

The diagram on the upper left of Figure 2.5 shows the values of V^* for each state. For example, consider the bottom right state in this diagram. The value of V^* for this state is 100 because the optimal policy in this state selects the 'move up' action that receives an immediate reward of 100. Thereafter, the agent will remain in the goal state and receive no further rewards. Similarly, the value of V^* for the bottom center state is 90. This is because the optimal policy will move the agent from this state to the right (generating an immediate reward of zero), then upward (generating an immediate reward of 100). Thus, the discounted future reward from the bottom center state is

$$0 + \gamma 100 + \gamma^2 0 + \gamma^3 0 + ... = 90$$

This grid world environment is a deterministic, finite example where the agent has prior knowledge about the effects of its actions on the environment. The problem is finite because there is a final end or goal state. It is deterministic because the agent is allowed to perfectly predict the immediate result (the succeeding state and resulting reward) for every possible state-action transition. It is possible for the agent to gain perfect knowledge of the immediate reward function and the state transition function.

Under these conditions, a good evaluation function that the agent could learn is V^*. The agent would prefer state s_1 over s_2 whenever $V^*(s_1) > V^*(s_2)$, because the cumulative future reward will be greater from s_1. Thus, V^* can be used by the agent to choose among actions:

$$\pi^*(s) = \overset{argmax}{a} [r(s,a) + \gamma V^*(\sigma(s,a))] \tag{2.6}$$

where $\overset{argmax}{a}$ is the action that returns the greatest possible reward. Recall that $\sigma(s,a)$ denotes the state resulting from applying action a to state s. Thus, the agent can acquire the optimal policy by learning V^*. Unfortunately, learning V^* is useful ***only when the agent has perfect knowledge of the immediate reward function r and the state transition function σ***. This requires that it be able to perfectly predict the immediate result (the immediate reward and immediate successor) for every possible state-action transition.

The diagram at the bottom of Figure 2.5 shows one optimal policy for these settings. Because the agent is not rewarded until it reaches its goal state, the optimal policy directs the agent along the shortest path toward state G.

2.1.5 Q Learning

If the agent did not have knowledge of the reward function r and the state transition function σ, the agent would not be able to predict the next state that will result from an action. In this case, V^* is of no use for selecting optimal actions because the agent cannot evaluate Equation 2.6. The solution may be found through a method called Q learning [91].

The evaluation function $Q(s,a)$ is defined so that its value is the maximum discounted cumulative reward that can be achieved starting from state s and applying action a as the first action. In other words, the value of Q is the reward received immediately upon executing action a from state s, plus the value (discounted by γ) of following the optimal policy thereafter.

$$Q(s,a) = r(s,a) + \gamma \max_{a'} Q(\sigma(s,a),a') \tag{2.7}$$

$Q(s,a)$ is the quantity that is maximized in order to choose the optimal action, $\max_{a'}$, in state s.

Learning the Q function corresponds to learning the optimal policy but how can Q be learned? The answer is through iterative approximation. Let the symbol $\overline{Q}$ be the learner's estimate, or hypothesis, of the actual Q function. In this algorithm the learner represents its hypothesis $\overline{Q}$ by a large table with a separate entry for each state-action pair. The table entry for the pair $\langle s,a \rangle$ stores the value for $\overline{Q}(s,a)$ - the learner's current hypothesis about the actual but unknown value $Q(s,a)$. The agent repeatedly observes its current state s, chooses and executes some action a, then observes the resulting reward $r = r(s,a)$ and the new state $s' = \sigma(s,a)$. It then updates the table entry for $\overline{Q}(s,a)$ following each such transition, according to the rule:

$$\overline{Q}(s,a) \leftarrow r + \gamma \max_{a'} \overline{Q}(s',a') \tag{2.8}$$

This training rule uses the agent's current $\overline{Q}$ values for the new state s' to refine its estimate of $\overline{Q}(s,a)$ for the previous state s. Under this equation, the agent executes the action in its environment and then observes the resulting new state s' and reward r. With each repeated state-action transition, the agent refines its estimate $\overline{Q}$. Each time the agent moves forward from an old state to a new one, $\overline{Q}$ estimates are propagated *backward* from the new state to the old. At the same time, the immediate reward received by the agent for the transition is used to augment these propagated values of $\overline{Q}$. Given a sufficient number of training episodes, the information will propagate from the transitions back through the entire state-action space available to the agent, resulting eventually in a table containing the true Q values.

Consider applying this algorithm to the grid world and reward function shown in Figure 2.5, for which the reward is zero everywhere, except when entering the goal state. The Q value for each state-action transition

equals the r value for this transition plus the estimated $Q(s,a)$ value for the resulting state discounted by γ. The optimal policy corresponds to selecting actions with maximal Q values.

2.1.6 Temporal Difference Learning

The Q learning algorithm learns by iteratively reducing the discrepancy between Q value estimates for adjacent states. In this sense, Q learning is a special case of a general class of *temporal difference* algorithm that learns by reducing discrepancies between estimates made by the agent at different times. The training rule of Equation 2.8 reduces the difference between the estimated $\overline{Q}$ values of a state and its immediate successor, but we could just as well design an algorithm that reduces discrepancies between this state and more distant descendants or ancestors.

Thus far, we have looked only at Q learning training rules that calculate a value for $\overline{Q}(s_t, a_t)$ in terms of the values for $\overline{Q}(s_{t+1}, a_{t+1})$ where s_{t+1} is the result of applying action a_t to the state s_t. Let $Q^{(1)}(s_t, a_t)$ denote the training value calculated by this one-step lookahead

$$Q^{(1)}(s_t, a_t) \equiv r_t + \gamma \max_a \overline{Q}(s_{t+1}, a)$$

An alternative way to compute a training value for $Q(s_t, a_t)$ is to base it on the observed rewards for two steps

$$Q^{(2)}(s_t, a_t) \equiv r_t + \gamma r_{t+1} + \gamma^2 \max_a \overline{Q}(s_{t+2}, a)$$

or, in general, for n steps

$$Q^{(n)}(s_t, a_t) \equiv r_t + \gamma r_{t+1} + ... + \gamma^n \max_a \overline{Q}(s_{t+n}, a)$$

Sutton (1988) introduced a general method for blending these alternative training estimates, called TD(λ). The idea was to use a constant $0 \leq \lambda \leq 1$ to combine the estimates obtained from various lookahead distances in the following fashion

$$Q^{\lambda}(s_t, a_t) \equiv (1-\lambda)[Q^{(1)}(s_t, a_t) + \lambda Q^{(2)}(s_t, a_t) + \lambda^2 Q^{(3)}(s_t, a_t) + ...]$$

An equivalent definition is:

$$Q^{\lambda}(s_t, a_t) = r_t + \gamma[(1-\lambda) \max_a \overline{Q}(s_t, a_t) + \lambda Q^{\lambda}(s_{t+1}, a_{t+1})]$$

When $\lambda = 0$ we have our original training estimate $Q^{(1)}$, which considers only one-step discrepancies in the $\overline{Q}$ estimates. As λ is increased, the algorithm places increasing emphasis on discrepancies based on more distant lookaheads. At the extreme value $\lambda = 1$, only the observed r_{t+i} values are considered, with no contribution from the current $\overline{Q}$ estimate. Note when $\overline{Q} = Q$, the training values given by Q^{λ} will be identical for all values of λ such that $0 \leq \lambda \leq 1$.

The value of Q for the current state and action summarizes in a single number all the information needed to determine the discounted cumulative reward that will be gained in the future if action a is selected in state s. This means the agent can choose the optimal action without ever conducting a lookahead search to explicitly consider what state results from the action. The agent can choose globally optimal action sequences by reacting repeatedly to the local values of Q for the current state. The motivation for the TD(λ) method is that in some settings training will be more efficient if more distant lookaheads are considered.

2.1.7 Function Approximation: Generalizing Information

A large drawback of Q learning up to this point is that the target function is represented as an explicit lookup table with a distinct table entry for every state-action pair value. Thus, the algorithms we discussed make no attempt to estimate the Q value for unseen state-action pairs by generalizing from those that have been seen. Also, convergence only occurs if every possible state-action pair is visited infinitely often. This is clearly an unrealistic assumption in large or infinite spaces. As a result, more practical systems often use function approximation methods.

It is easy to incorporate function approximation algorithms into the Q learning algorithm, by substituting a neural network for the lookup table and using each $\overline{Q}(s, a)$ update as a training example. Backpropagation of error methods are usually used to train a neural network to learn the Q function [1]. In backpropagation of error, or 'backprop', the network's weights are adjusted to follow the gradient descent of its error surface. In traditional backpropagation, errors are calculated as the sum of the squares of the errors between the desired and target responses. However, in Q learning, error is calculated as [82, 1, 4]:

$$error = r_{t+1} + \gamma \max_{a_{t+1}} [Q(s_{t+1}) - Q(s_t, a_t)]$$

This temporal-difference error may be derived by using $\overset{max}{a_{t+1}} [Q(s_{t+1}) - Q(s_t, a_t)]$ as an approximation to $\sum_{k=0}^{\infty} \gamma^k r_{t+k+2}$ [1, 4].

Thus, a one time-step Q learning equation to update a neural network's weights is [65, 45]:

$$\Delta w_t = \alpha[r_t + \gamma \overset{max}{a \in A} Q_{t+1} - Q_t] \nabla_w Q_t \tag{2.9}$$

where α is the learning constant, $\overset{max}{a \in A}$ is the chosen action among the set of all possible actions that leads to the maximum reward, and $\nabla_w Q_t$ is the derivative of the error function with respect to the weight.

To adapt this equation to deal with multiple time steps Watkins (1989) suggested using the current update error to adjust not only the current estimate of Q_t, but also that of previous states by keeping a weighted sum of earlier error gradients,

$$\Delta w_t = \alpha[r_t + \gamma \overset{max}{a \in A} Q_{t+1} - Q_t] \sum_{k=0}^{t} (\lambda\gamma)^{t-k} \nabla_w Q_t \tag{2.10}$$

where λ weights the relevance of the current error on earlier Q value predictions. Lin (1993) reports that this form can also be re-expressed as:

$$\Delta w_t = \eta(Q'_t - Q_t) \nabla_w Q_t \tag{2.11}$$

where

$$Q'_t = r_t + \gamma[(1 - \lambda) \overset{max}{a \in A} Q_{t+1} + \lambda Q'_{t+1}] \tag{2.12}$$

We could encode the states and actions as network inputs and train the network to output the target values of $\overline{Q}$ given by the above training rules. An alternative is to train a separate network for each action, using the state as input and $\overline{Q}$ as output. Figure 2.6 illustrates these two architecture options.

2.1.8 Exploration Versus Exploitation

We have thus far not specified how the agent initially chooses which actions to take. One strategy is for the agent in state s to always select the action a that maximizes $\overline{Q}(s, a)$, thereby exploiting its current approximation $\overline{Q}$. However, with this strategy the agent runs the risk that it will over commit to actions that are found during early training to have high $\overline{Q}$ values and fail to explore other actions that have even higher values. To gain a more accurate estimate of every actions' true values, it is required that the agent make every state-action transition infinitely often. This will clearly not occur if the agent acts greedily and

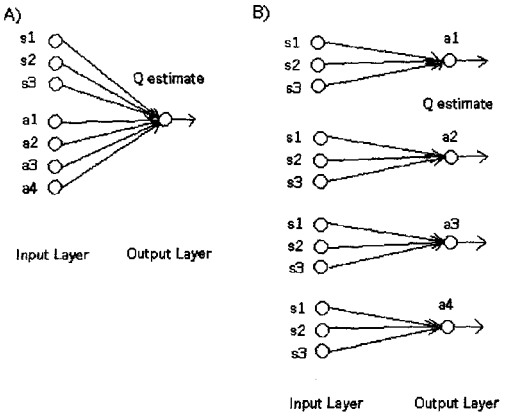

Figure 2.6: *Examples of how Q learning may be implemented in a neural network architecture. (A) Encode the states and actions as network inputs and train the network to output the target values of Q; (B) Train a separate network for each action, using the state as input and Q as output.*

always selects actions that maximize its current $\overline{Q}(s,a)$. For this reason, it is common to use a probabilistic approach to select actions.

Predetermining the probability that an agent will take a non-greedy action is a popular way of balancing exploration and exploitation. The drawback is that during exploration it chooses equally among all actions. This means that it is just as likely to choose the worst appearing action as it is to choose the next-to-best. This is unsatisfactory when the worst actions are very bad. The solution is to vary the action probabilities as a graded function of estimated value. More formally, the probability that an action a_i is selected given the agent is in state s is:

$$P(A = a_i|s) = \frac{e^{Q(s,a_i)}}{\sum_{j=1}^{n} e^{Q(s,a_j)}} \tag{2.13}$$

Actions with higher $\overline{Q}$ values are assigned higher probabilities and those with low $\overline{Q}$ values are assigned low probabilities of being selected. This results in the agent choosing potentially good actions and avoiding the bad ones when exploring.

2.1.9 Neural Applications

There are problems with mapping Q-learning or TD learning algorithms directly onto our particular task. With a network that represents the hippocampal and prefrontal regions of the human brain, what would

represent the agent? Would the 'actions' be the actions taken by an entire region or by an individual unit? What constitutes the 'environment' and, for that matter, how would this environment determine and execute rewards?

When modeling human performance on prefrontal-demanding tasks, incorporating the ideas set forth in reinforcement learning is attractive: theoretically, subjects may explore various possible strategies or actions and exploit the best one (or best several) to accomplish their goals. One may define this goal as maximizing accumulated returns over time. For example, in a word list free-recall task the goal may be to recall as many words as possible.

The problem that faces us now is how to incorporate these ideas into our model. The function approximation techniques described earlier are not suitable for several reasons:

- In our free recall task, there are no pre-established, state-action sequences. Suppose that the patterns of activity in our hippocampal module represent a 'state'. Let us further suppose that the activity of any unit in our prefrontal module represents an 'action' used to generate a new state in the hippocampal module. We have no predefined actions which the network can take as a result of being in that state because this is knowledge that we want the network to be able to learn over the course of trials. As a result, there is no pre-established value for taking any particular action.

- Because there is no well-defined sequence of states and actions between our hippocampal and prefrontal module, it is difficult to credit which actions taken by the prefrontal module led to the network's 'goal state'. For example, suppose that in its attempt to recall 16 list words, our network recalled five words correctly followed by eleven incorrect words. On another trial it recalled eleven words incorrectly followed by correctly recalling five words. On a third trial, it recalled eleven words incorrectly interspersed with five correct ones. In all three cases, the last action taken by the network is equally reinforced because the 'goal' state (i.e., the number of words correctly recalled on that trial) is the same. Backpropagation of the true values of any action would be impossible.

- We could possibly train the network to learn the sequence of words (states) as they are presented, but learning this direct state-action sequence during training would not necessarily be good. It would lead to the recall of words in the order that they were presented during study rather than a 'clustering' order.

2.1.10 Reinforcement Learning using Trace Values

We needed a rule that had the same properties as TD or Q learning but didn't depend on the value of the future states to calculate the value of the current state. We also wanted to use a more biologically plausible learning rule, similar to Hebb's rule, where connections between neurons are strengthened or weakened as a function of neural activity. Hebb's rule, however, is limited to expressing the simultaneous relationship between pre- and post- neurons' (units) activities. Sutton & Barto (1981) introduced the use of prolonged activity levels, called ***stimuli traces*** to track the intratrial temporal relationships between these units' activities. The learning rule we used in our model to train the prefrontal module combines the use of stimuli traces and reinforcement learning. This allowed the units in our network to be influenced by any reinforcement signals given to them and still maintain a 'blurred activity' of past activities.

As described in Section 2.1.1, Hebb's rule is often implemented in neural networks through the expression:

$$w_i(t+1) = w_i(t) + \alpha x_i(t) y(t)$$

where $w_i(t+1)$ and $w_i(t)$ are the 'new' and 'old' values of the weight w_i respectively, α is a positive constant determining the rate of learning, and $x_i(t)$ and $y(t)$ are the activities of the pre- and post-synaptic neurons (units) connected by w_i at time t. This learning rule can implement a simultaneous contiguity effect: w_i is incremented by α whenever an input pulse arrives and the unit fires and is unchanged (or decreased) otherwise. A modification of this rule, introduced by Sutton and Barto (1981), uses ***stimuli traces*** ($\overline{y}$ and $\overline{x_i}$):

$$w_i(t+1) = w_i(t) + \alpha[y(t) - \overline{y}]\overline{x}_i(t) \tag{2.14}$$

where

$$\overline{x}_i(t+1) = c\overline{x}_i(t) + x_i(t)$$

$$\overline{y}(t+1) = \beta\overline{y}(t) + y(t)$$

where c and β are positive constants and $0 \leq c, \beta < 1$. Under these equations, presynaptic activity at time t, indicated by $x_i(t) = 1$ initiates a prolonged trace given by nonzero values of a separate variable $\overline{x}_i$ for some period of time after t. This is accomplished by letting $\overline{x}_i$ be a weighted average of the values of x_i

for some time period preceding t. Similarly, there is a trace of the output y where $\overline{y}(t)$ denotes a weighted average of the values of the variable y over some time interval preceding t. Therefore, the network's behaviour is described by the values over time of the two variables y and $\overline{y}$ and the values of the three variables $x_i, \overline{x}_i$, and w_i for each input pathway $i = 1, ..., n$.

Whereas Hebb's rule detects correlations between input and output signals, this new rule detects correlations between traces of input stimuli and changes in output. This use of stimulus tracing permits the network to reproduce some of the intratrial temporal relationships between stimuli and responses made.

The learning rule that we used for our model is a combination of the reinforcement learning rules and stimuli trace rules proposed by Sutton & Barto (1986). Formally, the equation used to update the prefrontal module's weights was:

$$w_{ij}(t+1) = w_{ij}(t) + \alpha[r + \gamma y_i(t) - \overline{y_i(t)}]x_j(t) \tag{2.15}$$

where:

$$y_i(t) = f(\textstyle\sum_j w_{ij}x_j) \qquad\qquad f(\bullet) = \frac{1}{1+e^{-x_j}}$$

$$\overline{y_i}(t) = (1-\delta)\overline{y_i}(t-1) + \delta y_i(t) \qquad 0 \leq \delta \leq 1$$

In the above equations, α is a small positive constant called the learning rate parameter, r is a reinforcement signal that takes on the values $+1$ or -1, γ is a discount factor ($0 \leq \gamma \leq 1$) as in traditional reinforcement learning and $x_j(t)$ is the incoming pattern of activity. As indicated above, $y_i(t)$ is the total weighted summed input to unit y_i as a function of the sigmoidal function. The weights are also influenced by a trace of the output y, denoted by $\overline{y_i}$, which is a weighted average of the values of the variable y over some time interval preceding t. How much influence previous values of y has on the current value of $\overline{y_i}$ (and ultimately on the weight change itself) is determined by δ, a small positive constant between 0 and 1. In short, a unit's current activity depends in part on its past recent activity.

Upon further analysis of these equations, it is apparent that if α is large then any weight changes on future trials will be large providing neither $x_j(t)$ nor $[r + \gamma y_i(t) - \overline{y_i(t)}]$ is equal to 0. The direction and size of this change, however, depends on several factors. Weights will be weakened when $r = -1$ provided that it is less in value than the difference $\gamma y_i(t) - \overline{y_i(t)}$. Alternatively, the weight may still be weakened even when $r = +1$ if the input $x_j(t)$ is negative. The weight may also be weakened even when $r = +1$ if $\overline{y_i(t)}$ is larger in value. Thus, it cannot be easily said with confidence that weights will be strengthened or weakened based

on mere punishment or positive reinforcement signals (r) provided to the network - one must also take other variables into account.

As γ, the discount factor, approaches zero the effect that the present value of y has on updating the weight becomes negligible and the cumulative, weighted average of past y values play a larger role. As noted earlier, if $\overline{y_i(t)}$ is large, the likelihood that the weight is weakened becomes greater. This parameter is modulated, however, by the variable δ. As δ approaches 1, $\overline{y_i}(t)$ becomes nothing more than the current weighted summed input to the unit. The influence of the unit's previous activity on its current activity is negligible. The converse happens as δ approaches 0.

If the blurred, cumulative value of y (denoted as $\overline{y_i(t)}$) is greater than the current activity of y then the connections between the incoming unit x_j and that unit y will be weakened. An analogy may be drawn between this learning rule and Rescorla-Wagner's equations of reinforcement learning. When an animal has come to associate a reinforcing CS with a strong rewarding stimulus such that its expectations for that CS are great, then a decrease in the associative strength to that CS will lead to a decrease in its expectation of that stimulus. However, any continued positive reinforcement to the CS will lead to an increasing but asymptotic value of associative strength with the CS. This relationship between reinforcement learning in artificial neural networks and Rescorla-Wagner's theory of reinforcement learning is analyzed in greater depth by Sutton and Barto (see Sutton & Barto, 1998).

2.1.11 The Hopfield Network

In 1982, John Hopfield introduced the algorithm and architecture that is now known as the Hopfield Network. This architecture is a recurrent, symmetrically connected network with no self-feedback connections and the learning employs a one-trial Hebbian update. It is a design that can associatively store patterns and has often been incorporated in neural network models that simulate recognition and recall in memory tasks.

In a Hopfield network, memories are represented by spatial patterns distributed across units. After the weights are trained, each remembered pattern corresponds to an attractor state. An attractor state is a pattern of activity to which the network will converge given the same originally stored pattern or a noisy version of that pattern (a memory 'cue'). It is a state in which the activities of the units remain constant even after several update iterations. The update procedure for calculating unit activations is described in the next section. This convergence to an attractor state is necessary for the network to perform recognition or recall of stored memories.

Updating A Hopfield Network

In a Hopfield network, the weights are updated with every input pattern presentation by the rule:

$$w_{ij} = (1/N)\sum_{\mu=1}^{p} \xi_i^\mu \xi_j^\mu \tag{2.16}$$

where p is the total number of stored patterns labeled by μ, ξ_j^μ is the activity of the jth unit in pattern μ, and N is the number of memory units.

When a degraded version of a stored pattern is presented to the network, the units send signals back and forth in a closed feedback loop until their states become stable. Each individual unit calculates its summed input and produces an output through the sigmoidal function.

Tsodyks (1988) proposed a modification to the Hopfield algorithm that takes into account mean firing rates (i.e., continuous-valued units instead of binary valued ones) in sparse networks. This modification enhances the storage capacity of the neural network with low activity levels and correlated patterns. Using this modification and the update algorithm used by Becker et. al (1999) for her model of the hippocampus, we gradually updated units by the equation:

$$y_i(t) = \alpha[softmax(x_i) - 1/N] + (1-\alpha)y_i(t-1) \tag{2.17}$$

where $softmax$ is the softmax probability function described by 2.13.

One method for the update process is to update the units in sequence, then repeat the sequence until a stable state is attained. Another is to choose the next unit to be updated at random, which allows all units to have the same average update rate. In either case, the network will eventually reach a stable state. The recurrent connections allow the network to relax into an attractor in the absence of external input.

This stable state, or attractor, also corresponds to a minimum of the network's energy. Each state has an associated energy value defined by:

$$E = -(1/2)\sum_j \sum_{i}^{j\neq i} w_{ji}u_j u_i$$

for the binary-valued unit case, and for the continuous-valued case:

$$E = -1/2\sum_j \sum_{i}^{i\neq j} w_{ji}x_j x_i - \sum_i u_j I_j$$

where u_j is the activity of the jth unit of pattern u and I_j is an input bias term.

This energy function is the objective function that is minimized by the network. That is, the update procedure moves the state of the network until it corresponds to one of the network's 'memories'. At this point the network has reached a stable state: all units retain the same value upon updating and this state corresponds to the minimum of the network's energy.

Recall Limitations

The recalled memory pattern may not necessarily be most similar to the intended input pattern because of the network's initial state: the network uses the minimization procedure to find the *nearest* minimum energy state. Thus, it is possible that many different initial states may evoke the same minimum energy state. Also, the network may evoke spurious attractor patterns that were not on the list of originally stored patterns. This occurs when there are correlations between stored patterns. Storage capacity and correlated patterns are further discussed in Section 5.1.

2.2 Neural Networks and Biological Reality

Precisely how the human brain works is still a puzzle despite the combined efforts and centuries of studies by philosophers, doctors and scientists. Based on the information that has been compiled (and perhaps more so based on knowledge that we still do not yet have) it can be stated without quivocation that the brain is a complex organ. For the sake of computational feasibility, neural network modelers must abstract away many of its biological details. Biochemical aspects are often ignored as are cellular details such as the location of axonal connections along the dendrites or neuronal soma of one neuron from another. The lobes, its substructures, the multiple layers of cells in its cytoarchitecture and the number of connections between them are often schematized. Despite these simplifications, the model may still generally describe the processes of the modelled brain regions. Of course, the model is made more plausible when what it describes is supported by what is already known or accepted through empirical evidence obtained from other studies.

Having said that, it is worthwhile to review the biology of the hippocampus and prefrontal cortical regions before presenting our computational model that is designed to simulate their interaction.

2.2.1 Anatomy of the Hippocampal and Frontal Regions

The Hippocampus

The hippocampus comprises a portion of the limbic system and is found curled within the medial border of the temporal lobe. It interlocks with the dentate gyrus like two capital C's facing each other; hence the name hippocampal-dentate complex. As the oldest cortical portion of the limbic system, this complex contains only three recognizable layers instead of the more usual six layers of the neocortex.

The hippocampus and dentate gyrus are demarcated from the adjacent temporal cortex (parahippocampal gyrus) by the hippocampal fissure. The parahippocampal gyrus is in turn demarcated from the rest of the temporal lobe by the collateral sulcus. The parahippocampal gyrus is the zone of transition between the six-layered neocortex and the three-layered archicortex or hippocampal-dentate complex [16].

The parahippocampal gyrus is divided into the entorhinal cortex and the subicular cortex. Proceeding medially from the entorhinal cortex, the subicular cortex is organized into presubiculum, subiculum and prosubiculum, the last of which is continuous with the hippocampus. In continuation of this same progression, the hippocampus is divided into four sections: CA_1, CA_2, CA_3, and CA_4. Sections CA_1 through CA_3 exhibit an organized laminar structure but this pattern is lost in CA_4. Afferent fibers to the hippocampus approach from three sources: the parahippocampal gyrus, the septal-hippocampal pathway of the fornix, and the contralateral hippocampus via the hippocampal commissure. The entorhinal neocortex constitutes an important site of convergence for fibers from many sensory systems. Shown in Figure 2.7 are two slightly different paths, alvear fibers ('trench') and perforant fibers, running through the transitional subicular areas to enter the hippocampal-dentate complex. The fornix brings afferent fibers from the septal area, the basal forebrain area, and the diencephalon.

The largest proportion of efferent fibers leave the hippocampus and subiculum as the alveus and fimbria. On each side, the fibers of the fimbria become the crus of the fornix. Each crus arches superiorly and medially to become the body of the fornix and curves inferiorly to form the column of the fornix. About half of these fornical fibers project in front of the anterior commissure to become precommissural fibers. These pass to the septal nuclei, anterior hypothalamus, and the nuclei of the diagonal band of Broca. The postcommissural fibers project to the mammillary body and other areas of the hypothalamus, the anterior and intralaminar nuclei of the thalamus, and the midbrain tegmentum. In addition, some efferents cross with the anterior commissure to reach the contralateral hippocampus while others project to the subiculum and entorhinal cortex.

The unique multi-layered circuitry of the hippocampus is believed to provide the key to its rapid encoding

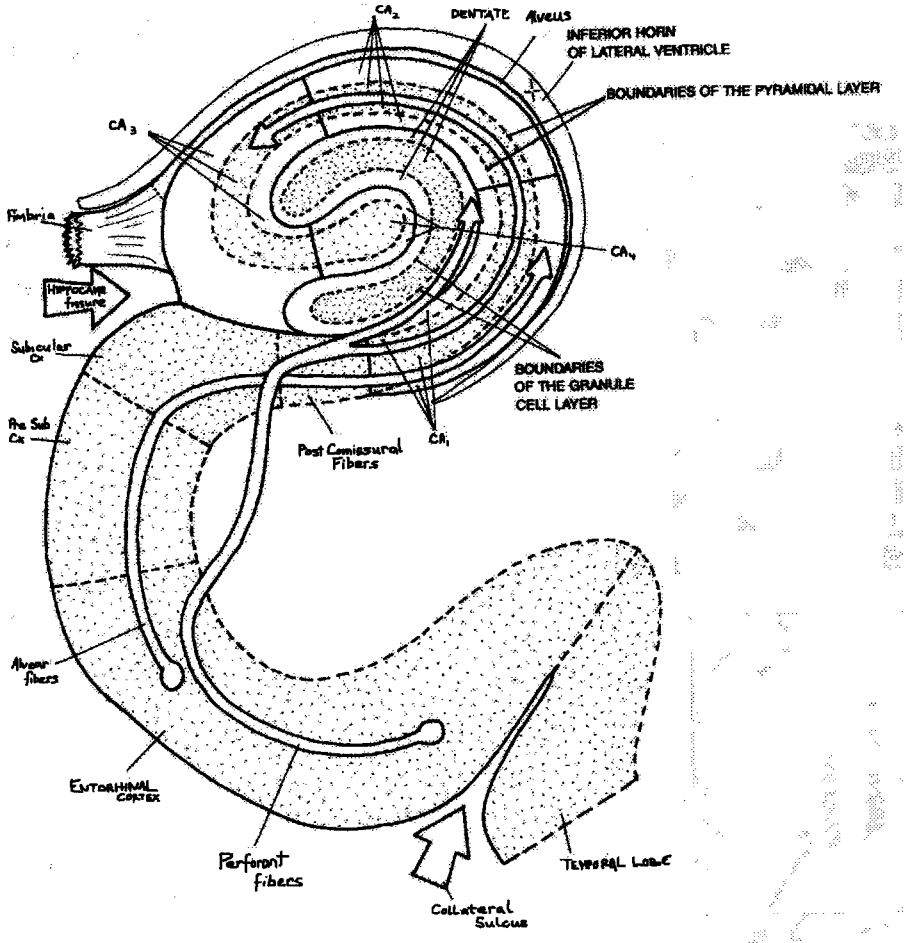

Figure 2.7: *Anatomy of the hippocampal region. The parahippocampal gyrus (entorhinal cortex and subicular cortex: presubiculum, subiculum, postsubiculum); the hippocampus (CA1, CA2, CA3, CA4) and Dentate; the afferent pathways (alvear fibers and perforant fibers) and the efferent pathways (alveus/fimbria); fornix; precommissural fibers and postcommissural fibers [16]*

and high capacity. The extensive recurrent connections between cells in the CA3 field have led many researchers to conclude that this area is similar to a recurrent autoassociative neural network, capable of encoding incoming patterns of activity rapidly after only a single exposure and associating arbitrarily related items. The CA1 layer has been attributed a hetero-associative role [63, 83] and many support the view that the dentate gyrus processes incoming cortical patterns before relaying them to the CA3 region [51, 83]. This preprocessing involves clustering and sparse encoding. The returning projections to the neocortex permit subsequent cortical reinstatement of a memory from partial or degraded input cues [5].

We implement a simple model of the hippocampus using a single Hopfield network. In reality, the hippocampus likely uses several layers with sparse coding in the intermediate layers which would lead to improved memory capacity.

The Prefrontal Cortex

The frontal lobes encompass roughly a third of the cerebral cortex in humans. They have strong reciprocal relationships with two large functional zones: (1) the visual, auditory and somatosensory zones via the parietal and temporal association cortices and (2) the telencephalic limbic system, including subcortical areas that monitor the internal milieu and provide information for affective and motivational responses. It has also been noted that the frontal lobes, with their major frontal-limbic associations, represent the major, if not the only, neocortical representation of the limbic system [60].

Each frontal lobe can be divided into three major subareas – the primary motor cortex, the premotor cortex, and the prefrontal cortex. The orbital frontal and dorsolateral regions comprise subregions of the prefrontal cortex (see Figure 2.8).

The frontal lobes are covered with neocortex which is divided into six layers of differing emphasis in different areas. The general significance of this stratification is that the sensory system (the thalamo-cortical projection) tends to end in the fourth and outer layers, while the motor pathways stem from the deeper ones. Subtle differences in the cytoarchitecture of the cortex have led to the development of anatomical systems which divide the neocortex into from twenty to more than two hundred areas. None of these schemes has escaped criticism or been universally accepted [62]. The differences between some of the suggested architectonic areas are small and variable, and it is often impossible to detect a distinct boundary between adjacent areas. However, in keeping with a majority of the literature, Brodmann's system, which divides the cortex into 47 numbered areas has been used here.

The frontal cortex projects to the limbic cortex (more specifically, the 'gyrus fornicatus' which is composed of gyrus cinguli, retrosplenial cortex and parahippocampal gyrus). There are fibres originating from the

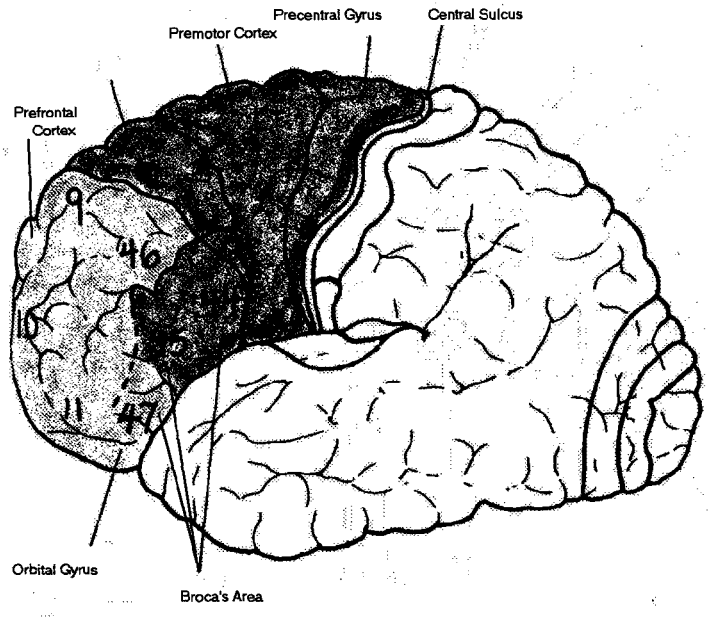

Figure 2.8: *Left cerebral hemisphere from the lateral surface. The frontal cortex (shaded areas) lies anterior to the central sulcus and above the lateral fissure. The prefrontal cortex comprises the anterior region of the frontal lobe. Brodmann Areas 9 and 46 are dorsolateral regions believed to be used for planning and control. Brodmann Areas 44, 45, and 47 are cytoarchitectonically similar to the prefrontal cortex but they comprise Broca's speech area so are considered to be premotor regions. Brodmann Area 47 may also be classified as prefrontal cortex, however. Dorsal to this area is the orbital prefrontal cortex - a region involved with personal affect and the ability to distinguish odors. The precentral gyrus is believed to be involved with discrete voluntary movments, the premotor cortex is involved with generalized movements, and the ocular motor area controls voluntary conjugate (bilateral) eye movements (figure from Diamond et. al., 1985; Stuss & Benson, 1986).*

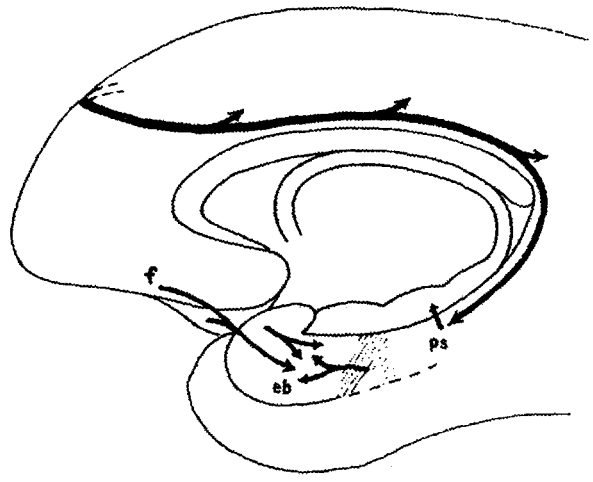

Figure 2.9: *Schematic indication of conduction routes leading from frontal cortex to the hippocampal region. The dorsal pathway originates in a dorsal convexity region and follows the fasciculus cinguli as far as the presubiculum (ps). The ventral path arises in the posterior orbito-frontal cortex (f) and terminates in the rostral half of the entorhinal area (eb) (figure adapted from [42]).*

dorsomedial surface of the frontal lobe that go caudalward in the cingulum bundle as far as the presubiculum and entorhinal area of the parahippocampal gyrus. The longest fibres even appear to extend into the subiculum immediately grading into Ammon's horn.

Another conduction route from the frontal cortex to the hippocampus originates in the caudal part of the cortex covering the orbital surface of the frontal lobe. It extends caudalward over a ventral route to the entorhinal area then projects massively to the hippocampus. Thus, there are at least two pathways by which the frontal cortex interacts with the neural mechanisms of the hippocampus: a fronto-entorhinal path that originates in the orbital frontal cortex and a fronto-limbic path that originates in the dorsomedial region and extends to the presubiculum and fornix bundle (Figure 2.9).

Chapter 3

Experimental and Neuropsychological Evidence

Clinical, neuroanatomical and experimental studies have led to many theories about the influence the hippocampal region and prefrontal cortex have on memory functions. In this chapter, we review some of the more popular theories and the relevant supporting evidence.

3.1 The Hippocampal Region

In recent decades, knowledge about the hippocampus has boomed in the areas of neuropsychology, anatomy, synaptic physiology and neurochemistry. Likewise, the number of theories posited about its function has also grown. Sutherland and Rudy (1989) noted that the hippocampal region has been attributed a unifying role for episodic memory, it has been considered as the intermediate-term buffer that bridges temporal gaps between events, it has been proposed to be responsible for working memory, anxiety, declarative memory, spatial mapping, temporal mapping, storing neocortical cell-assembly addresses, spatio-temporal context labeling, tuning out irrelevant events, response inhibition, memory-retrieval operations, and consolidation. Sutherland and Rudy, themselves, have suggested that the hippocampal formation is specialized for configural associations [81].

A popular theory, supported by McClelland and colleagues (1996), is that the hippocampus is a structure which learns faster than the cortex and is used to 'teach' memory patterns to the cortical structures through interleaved learning trials. This method reduces 'catastrophic interference' (the tendency to totally forget previously learned items as a result of acquiring newer information) - a problem that many previous

computational models faced. The cortex, in their model, is a slow learner which learns less detailed and more categorical representations than what is held in the hippocampus. A similar view from which many variations have grown (e.g., Jaap Murre's Trace Link model) is that the hippocampus serves as an 'index' in which compressed representations of a memory episode can be stored, which 'points' to the various cortical structures. When activated repeatedly over time, the intra-cortical links are strengthened until, eventually, the cortex alone will be able to recall an episode independent of the hippocampus [51]. Whether such a gradual consolidation process occurs is still controversial, but there is fairly broad consensus on the view that the hippocampal region can rapidly learn complex associations.

Interest in the hippocampal structures surged when it was discovered that ablation of this area produced the memory dysfunction called amnesia [52]. The characteristics of amnesia is an inability to acquire and retain new explicit memories despite normal intellectual functioning, normal immediate memory, preservation of most premorbid knowledge and skills, the ability to acquire new skills, and intact priming [68]. Such data suggest that there are multiple forms of memory in humans and distinct classes or clusters of memory processes. Indeed, a dissociation between intact skill learning and conscious recollection has been found in amnesics, and various studies have shown a dissociation between implicit priming tasks and explicit recall/recognition tasks in both normal controls and amnesic patients [69].

Interest in the hippocampal structures surged when it was discovered through patient cases such as H.M. and R.B., that ablation of these areas produce serious memory deficits. The hallmark characteristics of these and other amnesics is a deficit in the ability to acquire and retain new explicit memories yet an intact

One drawback of the human amnesic data is that amnesia is usually a consequence of neurological dysfunctions, including bilateral lesions of the medial temporal regions, Korsakoff syndrome, closed head injuries, ruptured aneurysms and anoxia. Not only is there a question of what side-effects these individual problems might cause, but also there is such a general gross impairment of brain regions, that it is difficult to pinpoint any one structure which may be responsible for the memory deficit.

Research on primates [78, 77, 97] has furthered our understanding of the specific temporal lobe structures involved in learning and memory. It is not only the hippocampus, but also the surrounding cortex (parahippocampal and perirhinal cortical structures) which, when lesioned, produce severe amnesic effects.

As mentioned earlier, several theoretical models depict the hippocampus as a structure which temporarily stores compressed representations of memory or maintains an index of it until, over time and through repeated exposures, the cortex can activate traces of memory patterns independently of the hippocampus. This theory of ***consolidation*** accounts for the neuropsychological data in which patients with closed-head injuries or other types of brain lesions show impaired retrieval of recent memories while old memories remain

preserved. Similarly, Salmon, Zola-Morgan and Squire (1987) showed that by varying learning-surgery intervals in primates it was possible to examine the hippocampus' role in long-term storage [66]. They found that the hippocampal area was needed not only during the time when the animal experiences and learns, but also for a 2-12 week interval following the experience. Damage to the hippocampus did not seriously impair memory for events which occurred long before the lesion [66]. The conclusion was that there is consolidation and transfer of information from the hippocampus to the cortex and that the hippocampus itself cannot be a permanent store for information retrievable for some period after exposure. Eichenbaum has further distinguished between hippocampus and neocortex by positing that the role of the parahippocampal region is to encode elements as fused, unitized or configural representations while the hippocampus promotes flexible associations by recognizing relations among items and differentiating overlapping patterns.

There are, however, two curious phenomena that contradict the theory that the hippocampal region is needed to 'teach' a slower-learning neocortex. One is that amnesics are capable of fast, very short-term memory. They can recall information for about a minute after its presentation, provided that distractions are minimized and uninterrupted rehearsal is permitted. Reasonably, therefore, we can infer that remembering events for about 60 seconds does not depend on the hippocampus - some cortical structure must be responsible for the initial short-term storage of a memory representation [1].

The second contradiction comes from a meta-study performed by Nadel and Moscovitch [59]. They examined retrograde amnesia (RA) following damage to the hippocampal complex in both humans and animals in previous studies and found that in many cases, RA can extend to as far back as 10-30 years, and in some cases an entire lifetime's episodic memory store could be absent. Such a flat temporal gradient seriously undermines the case for gradual consolidation. These findings suggest that, rather than transferring all memory patterns and traces to the cortex, the hippocampal region and surrounding temporal lobe structures are actively involved in autobiographical episodic and spatial memory for as long as they exist. Nadel and Moscovitch posit that the degree of retrograde memory loss may correlate with the size of damage, assuming older memories are distributed across more neurons.

A few case studies have supported the dissociation of episodic from semantic memory and underlined the hippocampal region's role in the former. One case study reported by Warrington and McCarthy (cited in Nadel & Moscovitch, 1997) suggested that a subject was able to acquire and retain general semantic

[1] Two intriguing theories have been proposed (Churchland and Sejnowski,1993, p.301) to explain working memory: one is that 'activity initiated perceptually would continue to circulate in a recurrent network for a brief time'. The second is that there exists a cellular mechanism with the same time course as the memory itself (i.e., a transient weight modification that lasts up to about 5 seconds). Post-Tetanic potentiation has been linked to the latter. The former theory has been criticized in that in order for a pattern to be held in working memory for five seconds, a signal would have to circulate in the loop many times (on the order of 102 to 103 times) depending on loop size. Considering the biological system, it is improbable that a signal can be so precisely preserved over hundreds of synapses

knowledge for words that came into use during the period for which he had a dense RA for autobiographical episodes and public events. He knew the meaning of words such as AIDS, telecom, shuttle and provisions, and could give detailed definitions of them even though his episodic memory during the times when these words came about was impaired. This would suggest that episodic memory (for which the hippocampal structures seem to be responsible) may not be necessary for the extraction of semantic memory. Similarly, a review of three case studies performed by Vargha-Khadem et al. showed that an acquisition of sustained bilateral hippocampal lesions very early in life (ages newborn, 4 and 9) in these patients produced amnesia and prevented the formation of episodic memory but it did not prevent the now-teenage subjects from attending mainstream schools, attaining levels of speech, literacy, language and factual knowledge that was within the low-average to average range [90]. This data further supports the view that episodic and semantic components of cognitive memory are at least partly dissociable, with only the episodic component being fully dependent on the hippocampus.

In our model, we have adopted the view that the hippocampal region is specialized for the rapid encoding of incoming information and is necessary for the retrieval of patterns from partial or degraded cues. A compressed, sparse representation of the incoming pattern is stored in the region while a less detailed and more categorical representation is maintained in the neocortex. In this simplified model of the hippocampus, we have sidestepped the issues and possible roles this region may have in spatial and temporal mapping, configural associations, and issues of retrograde amnesia and the length of time required for neocortical regions to retrieve episodic memories without aid from the hippocampus. For the purpose of simulating the California Verbal Learning Test and the AB-AC list learning test we found that our simple model of the hippocampal region was sufficient. As is evident from the paragraphs above, however, this region is more complex than we depict and to account for its other functions additional layers of units and extra connections and weights at a minimum would need to be added.

3.2 The Prefrontal Region

The prefrontal cortex is another brain region that has been the subject of interest and speculation among researchers. Patients with prefrontal cortical damage exhibit a wide array of cognitive deficits such as impaired initiation, planning and problem solving abilities; impaired memory for source and order of recalled material (temporal memory); difficulties in making behavioural shifts in attention, movement and attitude (perseveration); poor recall of thematic verbal material; increased levels of distractibility; increased trend in confabulation; abnormal failure to release from proactive interference (e.g., they do not benefit from a change

in the semantic category of words during a list-learning task); increased sensitivity of learning to interference; poor conditional learning; and diminished critical self-evaluation of behaviour and knowledge about what one knows, about how one learns and remembers best and about the accuracy of one's own memories (metamemory or monitoring, awareness and insight of ones knowledge) [33, 46, 53, 55, 68, 75, 80, 92]. Although some researchers view memory impairment as the primary deficit in frontal lobe dysfunction (e.g., working memory or temporal organization of memory) [53, 33, 25] , others view the impairment as secondary to other cognitive disorders, such as deficits in executive control, attention, inferential reasoning and cognitive mediation.

3.2.1 In Support of Memory

Experiments on lesioned primates indicate that the prefrontal cortex plays a role in the performance of delayed-response and delayed-alternation tasks [80, 66, 78, 77, 97]. These are standard memory tasks often administered to primates. In both tasks, the animal sits in a cage facing two identical, covered, shallow food wells situated at a specified distance from each other. The food wells can be shielded from the animal's view by a screen that can be raised or lowered by the examiner. In the delayed response task, food is placed in one of the containers in full view of the animal. Then, the wells are covered, and the opaque screen is lowered. Following a defined interval, the screen is raised and the animal must select the container with the food. In the delayed alternation task, the screen is lowered immediately after the animal selects one of the food wells, regardless of whether the response was correct or incorrect. If correct, food is placed in the alternate food well, both wells are covered and the screen is raised after the defined delay interval. If incorrect, food remains in the original well, and the trials are repeated until correct.

Monkeys without a prefrontal cortex perform badly on these tests. Evidence from neuronal intra-cellular electrical recordings and from imaging studies (positron emission tomography and magnetic resonance imaging data) also provide support to the theory that the prefrontal cortex is necessary for delayed memory tasks. Increased activity in prefrontal neurons is observed when subjects or animals are required to remember things from moment to moment. More specifically, whereas neurons in non-frontal regions of the brain 'forget' the target when a distracting stimulus appears (as measured by pre- and post- levels of activity), prefrontal cortical neurons maintain their rate of activity during a delayed-choice task even after the animal is presented with irrelevant, distracting stimuli [54]. Goldman-Rakic and colleagues (1991) have further suggested from similar studies that the prefrontal cortex may be organized into regions that temporarily store information about different sensory domains: one for the domain of spatial cues, one for cues relating to an object's appearance and others for various types of cues. Supporting studies have shown that neurons in the

prefrontal cortical region fire at different rates during the delayed-choice task, depending on the target the animal saw previously [25]. The domain specificity view of the prefrontal cortex is still under debate but, in general, the main conclusions that have been drawn from these studies is that the prefrontal cortex appears to be involved in keeping target information 'in mind' despite long delays and distractions [93, 25].

Other experimental studies have supported the view that the prefrontal cortex is involved in memory. Human patients with frontal lobe damage do badly on free recall tasks, have a decreased digit span, have problems completing tasks they initially set out to do, and have poor memory for sources of newly obtained information [75, 33, 80].

However, poor performance on the delayed-response tasks could alternatively be explained as a dysfunction of attention or a tendency to perseverate. Later experiments showed that performance improved if the distractions were minimized and attention fully captured. It was discovered that if the lights were turned off during the interval between target and test, performance improved [80]. Thus, the prefrontal cortex may be used to tune the visual and possibly other perceptual systems to the task at hand. It may be responsible for focusing an animal's attention and steering awareness. Performance on the delayed-alternation response task could be explained as perseveration on formerly rewarded behavior even after it had ceased to be rewarding [33]. Further studies led to the demise of the idea that the frontal lobes were directly involved in memory. Experiments with humans showed that there was no impairment on other standard memory tests (e.g., the Wisconsin General Testing Wechsler Adult Intelligence Scale and its subset of the memory scale, recognition, delayed-recall paired associate learning and nonverbal object span with 15 minute delays, Benton Visual Retention Test, recall of narrative prose) [80, 33]. Furthermore, the frontal-lobe damaged patient's capacity to acquire new information and form long-term memories was not impaired in the same fashion as those who exhibited classical amnesia.

3.2.2 So what is the role of the prefrontal cortex?

There have been many theories proposed on the role of the prefrontal cortex. Here, we describe a few of the more popular ones.

- Perseveration

  It was noted earlier that the delayed-alternation response task in primates could have been due to a perseverative-type dysfunction. Additional studies supported this idea. The Wisconsin Card Sorting Task (WCST) is a category 'shifting' test often considered to be sensitive to frontal lobe damage [33, 80]. On this test, the patient is presented with 4 target cards, each different in number, color and

form (1 red triangle, 2 green stars, 3 yellow crosses, 4 blue circles). The subject is given a deck of cards with similar stimuli to be sorted by placing each card in front of 1 of the 4 target cards based on 1 of 3 categories. The subject is only told 'correct' or 'incorrect' for each response. Using this feedback, the subject tries to get as many correct as possible. Unknown to the subject, the examiner selects one criterion as the correct one and bases the feedback on the relationship of the response to this criterion. After every 10 consecutive correct responses, the examiner alters the 'correct' sorting criterion without telling the subject. The subject must now 'shift' until a new principle is discovered. The final score reflects the number of sets of 10 correct responses attained by the subject, as well as the number of various types of errors. Results show that the frontal patient's perception of the relevant category is not impaired. In the beginning, patients are able to perform the categorization and verbalize their reasons for doing so. However, later in the task they are likely to continue to incorrectly sort by the originally approved criterion despite feedback that their responses are wrong. They tend to give impulsive answers with a tendency to perseverate. It was suggested that they had a deficit in 'shifting' or moving to a new task or criterion once they had an established position. Patients may even explicitly recognize that they are making mistakes, yet are unable to change their actions to correct them [46].

- Lack of inhibition

  The above Wisconsin Card Sorting task result has also been interpreted as an inability to inhibit prepotent or highly active responses. Another experiment which supports this is the Stroop task [8, 38, 37]. Frontal patients often make more word-naming intrusions than controls when asked to name the color. There is a failure to inhibit responses that are particularly available owing to their recent use. This is like perseveration in that frequently reinforced behaviours are repeated even when no longer rewarded or appropriate.

- Planning

  Frontal lobe damaged patients also do badly on tests of planning [8, 80]. They do poorly on tasks such as planning a trip into town [8] and on puzzles such as the Tinkertoy Test and the Towers of Hanoi [13, 80]. Both puzzles are tests that assess independent initiation, planning, and implementation of a potentially complex activity. The Towers of Hanoi is a 'look-ahead' puzzle that consists of rings and pegs. The examiner constructs a starting order of the rings and the subject must build a new order, as defined by the examiner. The level of difficulty is manipulated by the number of required moves. Although patients do badly on the overall problem, they are able to solve it if the task is broken down and step assistance is provided.

- Meta-Memory

Meta-memory involves the manipulation of information in order to strategically recall information from memory or the organization of information during encoding to solve future tasks. For example, it has been shown that frontal lobe damaged patients perform poorly on verbal fluency tests [33, 75, 55]. These are tests that require subjects to recall in an allotted time as many words that, say, begin with the letter 'A', or to list items that can be found in a supermarket. It has been suggested that they have problems chunking, categorizing, and separating lists or have an impairment in the use of external cues as a guide [68]. When the task is to memorize a long series of spoken/written words, those items that make a direct impression and require little effort to memorize are best remembered by patients with frontal damage [80]. The total number of items remembered does not significantly increase with repetitions of the list as it does with normal subjects, but different words may be recalled on subsequent trials. Luria (cited by Stuss, 1986) suggested that the passive imprinting of material is intact but a deficit exists in the control of retrieval. It has also been shown that frontal damaged patients have a greater tendency to repeat the more recent memory trace: when given one series of words to remember followed by a second series to remember and then requested to remember the first list, frontal damaged patients perform below normal. The second list may be presented, or, most likely, words from both lists are offered [75]. List-learning is a good measure of verbal memory because, unlike story recall, no semantic context is provided to organize information and aid recall. Frontal lobe damaged patients may do badly on free recall but they can easily recognize these same items they failed to recall. It is suggested that recall processes put heavier demands on internally generated memory strategies and require the use of strategic search and retrieval [95, 96].

On a test of free recall of word items, such as the California Verbal Learning Test (CVLT), frontal lobe damaged patients show low scores on a measure of organization (index of the degree to which same words are clustered together across multiple study-test trials) [34, 36, 19]. Unlike other list-learning measures, the CVLT is an auditorily presented test of learning and memory that uses words from conceptual categories which may aid in recall and can be used to discriminate between simple recall and information organization. The CVLT is a standardized clinical assessment that is administered as follows: four words each from four categories (16 words in total) are presented aurally in a mixed order. The task is to recall as many of the words as possible. The same list is presented five times with recall following each presentation. Measures include recall after the first trial, the fifth trial and the sum across the five trials. Immediately after the fifth trial, a new list is presented (list B) and the subject is asked to recall again. Short-delay memory is assessed with free recall (short-delay free recall) and

cued recall (subject is given the semantic categories as cues). Twenty minutes later, long-delay recall is assessed. In addition, a recognition task is given, where the patient is asked to identify the initial list from a set of distractor items. The ability to discriminate the original items from the distractors is assessed (discriminability). The tendency to repeat the same words (perseverations) and to say non-list items (intrusions) are also assessed [41]. The CVLT can also be modified to test proactive and retroactive interference effects in normal subjects. The concept of interference implies that if similar material is learned on two separate occasions, each learning episode reduces the accessibilty of what was learned during the other learning episode. Proactive interference refers to the decremental effect on later learning due to previously learned material on the retention of subsequently learned material (subject has a hard time recalling words at the end because of the items at the beginning); and retroactive interference refers to the detrimental effect on previously learned material due to later learned material (subject has a hard time recalling the beginning of a list, because of what was just recently learned). Kramer and Delis (1991) showed how the CVLT can be modified to illustrate these two types of interference effects in normal subjects. They presented two lists to their subjects. Both lists (List A and List B) were made up of 4 words each from 4 semantic categories. Two of the List B categories were the same as categories on List A and were referred to as 'shared categories. The other two categories were different and were referred to as 'nonshared' categories. List A was presented for 5 learning trials, and List B was presented once after List A. Immediately after each list presentation, the subject freely recalled as many words as possible. In addition, a short delay recall of List A was assessed after List B was presented, and recall and recognition tests of List A were performed after an additional 20 minute delay. As expected, they found that recall of shared items on List B was worse than recall of shared items on List A (a reflection of retroactive interference). Perhaps more surprising was that recall of non-shared items in list B was higher than recall of non-shared items in List A, which they attributed to a release from proactive interference.

Another important index measured by the CVLT is categorical clustering ability (the tendency for items from the same taxonomic category to be recalled consecutively at a level well above chance expectancy despite their separation during the study trial). Past studies have shown that normal, young subjects will organize several discrete items in a way that preserves some inherent relatedness among those items. These organizational processes are assumed to enhance memorability because recall performance is better when subjects use them. Furthermore, recall scores tended to be higher for a list of categorically related items than for a list of unrelated items [7, 35].

In a study performed by Hultsch (1975), a study list that contained 40 items and 4 instances from

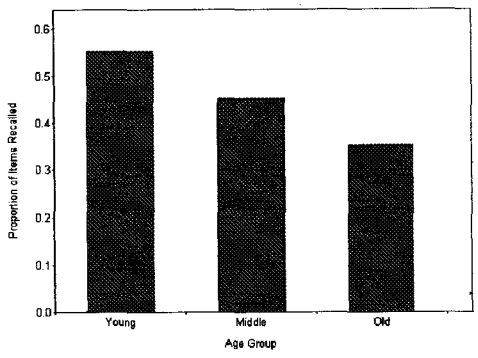

Figure 3.1: *The number of words correctly recalled under free-recall conditions of a list of related words (Hultsch, 1975).*

Figure 3.2: *The proportion of categories correctly recalled under free-recall conditions of a list of related words (Hultsch, 1975).*

each of 10 taxonomic categories was presented to normal, various aged subjects. The proportion of words recalled from this list by 3 age groups under free-recall conditions are plotted in Figure 3.1. Recall by young subjects significantly exceeded that of older adults with cued recall as well as with noncued recall. Similar results were found when comparing frontal lobe damaged patients against normal control subjects.

This impairment in free recall was attributed to the elderly or frontal lobe damaged patient's inability to successfully encode and use higher order categories to act as cues and aid in reconstructing the specific memory trace (Figure 3.2) [31, 73].

Frontal lobe damaged patients also perform worse than normal subjects on the AB-AC list learning task. In this test, subjects are given a list of word pairs to memorize. Later, they are presented with

a second list of word pairs in which the first words on the first list are now paired with new words. In the study conducted by Barnes and Underwood (1959) normal subjects learned a first paired-associate list (List 1) of 8 pairs until they achieved 1 perfect trial. Then List 2 was presented for 1, 5, 10, or 20 trials. Subjects were stopped at various points in learning List 2 and were asked to give both List 1 and List 2 responses to each stimulus. In the AB-ACD paradigm, the target words in the first list were dissimilar to the target words in the second list (e.g., desk-lion and pen-window in the first list, desk-thief and pen-cup in the second list). That is, there was no apparent interlist response similarity and intralist response similarity was low. In the AB-ACS paradigm, the target responses in one list were highly similar to the target words in the other list. The responses had low intralist response similarity but high interlist response similarity (e.g., desk-insane and pen-afraid in the first list, desk-crazy and pen-scared in the second list). For both groups combined, it took an average of 10 trials to learn List 1 to perfection. Figures 3.3 and 3.4 show the results obtained in this study. In general, more List 1 intrusions occur in the learning of List 2 for the AB-ACS than for the AB-ACD paradigm. These results suggest that learning a list of word items can impede the storage and recall of a new set of word items (also known as proactive interference). This is especially true when the target responses in one list are similar to the target words in the second list. They found that:

1. More List 1 intrusions occurred in the learning of List 2 for the AB-ACS paradigm than for the AB-ACD paradigm.
2. Rate of acquisition of the second list was retarded due to the first list learned.

As the number of A-C learning trials increased, recall of the C responses improved steadily, from 3.46 out of 8 (43%) after one A-C trial, to 7.33 (92%) after 20 trials. At the same time recall of the B responses steadily declined, from 6.67 (83%) after one trial on the A-C list, to 4.12 (52%) after 20 trials. Thus, learning of the C responses interfered with recall of the B responses, and this retroactive interference effect was larger the greater the amount of training on the A-C list.

Recall of the second list was higher in the AB-ACS test than in the AB-ACD test. This was attributed to a 'mediation' effect in which after learning A-B, the subject will recall the CS response because of its high associative connection with B.

Shimamura and his colleagues (1995) performed an experiment similar to Barnes and Underwood's AB-ACS experiment, but with frontal lobe damaged patients. Figures 3.5 and 3.6 show both groups' performance on both lists. Their results supported two of their hypotheses:

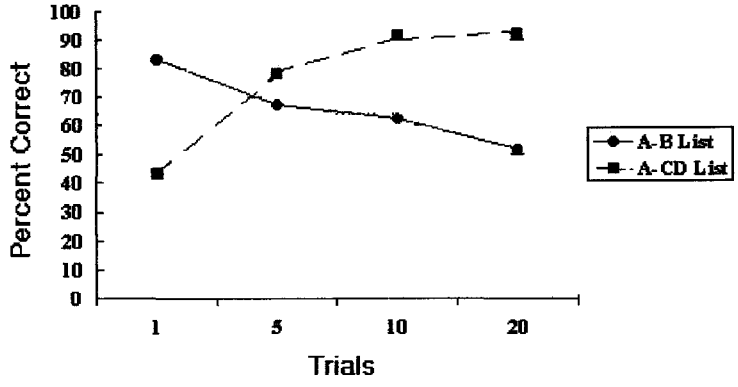

Figure 3.3: *Percent correct on the A-B and A-CD list in the Barnes and Underwood (1959) study as a function of number of trials on the A-CD list.*

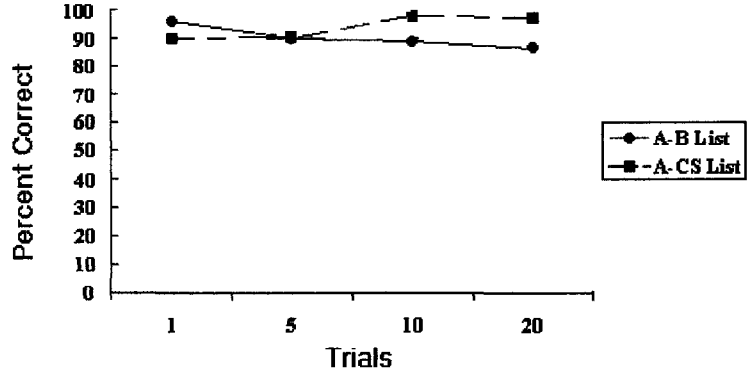

Figure 3.4: *Percent correct on the A-B and A-CS list in the Barnes and Underwood (1959) study as a function of number of trials on the A-CS list.*

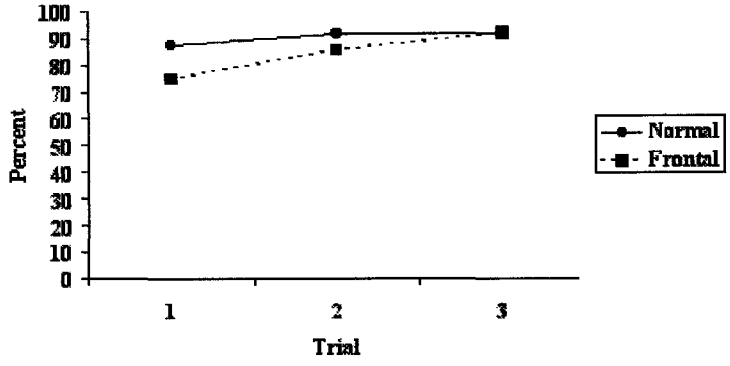

Figure 3.5: *Recall performance of both normal and frontal subjects on the A-B list (Shimamura et. al, 1985).*

1. Patients with frontal lobe lesions exhibited disproportionate impairment (proactive interference) on second-list learning.
2. The most significant decrement in performance occurred on the first learning trial of the second list, as this trial was most susceptible to interference from the first list.

- Contextual and Spatio-Temporal Information

Considering the AB-AC list learning task results, one could also conclude that there is a problem in distinguishing one list from another or maintaining context. Some have suggested that frontal lobe damage leads to a failure to make temporal distinctions and/or keep separate trials distinct [68, 53, 61]. That is, patients have problems discriminating recent from prior stimuli, or have a deficit in source memory. Other studies have shown that patients can often recall actual information given to them with few problems but cannot remember when or where those facts were learned. In addition, they have problems in discriminating items in memory temporally. In one study by Shimamura and colleagues [75], frontal lobe damaged patients were tested on the recall and recognition of words and facts. The task was to reproduce a previously presented list of words in order and arrange, in chronological order, a random display of 15 factual events that occurred between 1941 and 1985. The patients were impaired on placing the items in the correct temporal order despite normal item memory (normal recall and

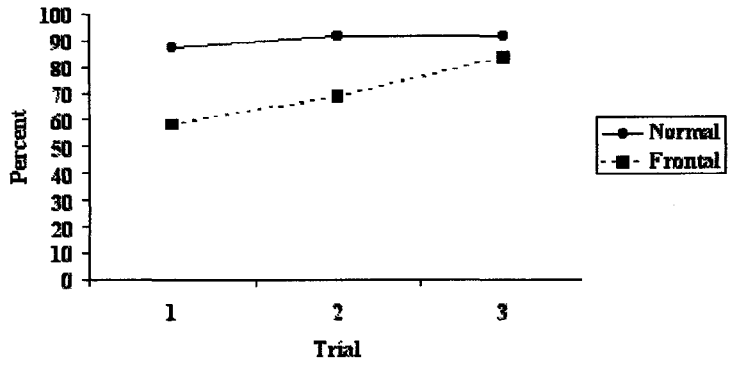

Figure 3.6: *Recall performance of both normal and frontal subjects on the A-CS list (Shimamura, 1985).*

recognition memory for the words and facts). In another experiment [53], 184 cards were prepared on which two items were shown. In the verbal modality test, two words were presented (eg, 'cowboy', 'railroad'). In the nonverbal modality test, reproductions of abstract art/representational drawings were shown. In the recency task condition the cards were presented one at a time. Several cards had a question mark between the two words or works of art. On those particular cards, both items had been presented on previous cards. When cards with question marks were presented to subjects, their task was to report which of the two items had been seen most recently. In the recognition task condition, a question mark occurred between the two items presented on one card, but only one item had been presented previously and the subject was asked to identify the item seen before. Results showed that although they performed at normal levels on the recognition tasks, patients with unilateral left-hemispheric frontal damage performed badly on the verbal recency test, and right-hemispheric frontal damaged patients performed badly on the nonverbal recency judgments as well as the verbal recency task.

- The Control of Action Selection

It has been suggested that the prefrontal cortex is used for response sculpting. That is, it is used when the subject must select one response or action from many rather than few possible alternatives [21]. In

a series of functional magnetic resonance imaging studies performed by Nathaniel-James & Frith (1997) subjects were given the task of generating words to complete sentences. In one condition they had to generate a word that best fit the sentence and in the other they had to generate a word that did not fit. In the 'high constraint' condition, few possible appropriate words were available (e.g., 'He posted the letter without a ——'). In the 'low constraint' condition, many possible appropriate words could have completed the sentence (e.g., 'The police had never seen a man so ——'). They found that high levels of activity were observed in the left prefrontal cortex during inappropriate completion conditions. Also, there was no effect of constraint in the inappropriate condition, but under the appropriate completion, there was more activity under conditions of low constraint. In another experiment, the subjects' task was to generate words on the basis of word stems. Word stems with many or few possible completions were used. When activity was compared between these two conditions, it was found that the many-completion condition (the less constrained condition) produced greater activity in the left dorsolateral prefrontal cortex. The results from this study also suggest that it is the lack of constraints and not the difficulty of the task that leads to activity in the prefrontal cortex. The task is more difficult when the word stem has few rather than many possible completions, but the less constrained condition produced greater activity.

Similar results were obtained when subjects performed a verb generation task [21]. The task in this test is to generate a verb that goes with a noun (e.g., cake-eat, knife-cut). Some of the nouns used (e.g., 'cat') had many possible associate verbs. Positron Emission Topography data showed high activity in the left prefrontal cortex during recall and activity in the right prefrontal cortex during encoding. Furthermore, this activity was a function of the constraint levels used. High activity levels in the prefrontal cortex were most evident in the low-constraint conditions.

All of the data obtained from these experiments suggest that the prefrontal region is used in the *selection* of appropriate responses from a set of many possible responses when the alternative options are plausible or equally appropriate to the desired response (low-constraint conditions). On relatively simpler tasks, where there is only one appropriate or prepotent response, there is little prefrontal cortical activity.

In conjunction with the experimental evidence described above, several more well-defined theories have been proposed to describe the role of the frontal lobes. Some executive system theories are described below.

1. Baddeley (1996), who first coined the phrase 'dysexecutive syndrome', introduced the idea of the frontal cortex as a 'central executive system'. This system directs and monitors the activities of lower-level

systems to achieve its goals in an efficient manner. The central executive is one component of a complex system called working memory (the other components include verbal and visuo-spatial storage buffers) and is responsible for control and decision processes such as directing attention and allocating cognitive resources.

2. Petride (1982) has suggested that there are two levels within the frontal cortex. The first is responsible for retrieving data from long term memory stores elsewhere in the brain. The second is responsible for more abstract thinking. Together they monitor brain processes and enable it to keep track of several events.

3. Shallice (1982) proposed that the prefrontal cortex houses the 'Supervisory Attentional System' (SAS) - a mechanism that exerts high-level control in situations where automatic processing is inadequate. The SAS is required only in cases of novelty, or in tasks requiring non-routine decision making. This theory accounts well for the Stroop task in which routine responses (namely, reading the word) need to be overridden in favour of a more novel response (naming the colour of its ink).

4. Moscovitch (1994) posited that the frontal cortex is necessary for controlled storage and retrieval. In terms of a memory system, the prefrontal cortex provides and controls information going to and coming from the slave-like, automatic hippocampal system that mandatorily encodes whatever information it is given.

The anti-executive proponents support theories of 'simpler and more fundamental' systems or processes ([38]):

1. Based on single-cell neuronal recordings in the primate that is performing delayed-response tasks, Goldman-Rakic and colleagues (1991) have proposed that the prefrontal cortex instantiates a form of working memory that Goldman-Rakic terms 'representational memory'. She further suggests that the prefrontal cortex is segregated into different functional regions. Neuronal cells in one prefrontal region are used to keep spatial representations in mind over short time intervals.

2. Cohen and Servan-Schreiber (1992) have posited that frontal dysfunction is a consequence of the difference in a single low-level parameter (the gain in dopamine levels) that directly affects the maintenance of context.

3. Fuster (1988) suggests that rather than superimposing a steering or directing function on the remainder of the nervous system, the prefrontal cortex expands the temporal perspectives of the system thereby

allowing it to integrate longer, newer and more complex structures of behaviour. He proposed three distinct prefrontal functions: active, or working memory; preparatory task set, or the ability to prepare for future action; and inhibitory control.

4. Contrary to Fuster's view, Kimberg & Farah (1998) proposed that the function of the prefrontal cortex lies not on inhibition, but rather on the connections and the levels of activation in what they call 'working memory' (which they define to be 'functions and properties of the human cognitive system that allow representations to have levels of activation appropriate to the current task') [38].

3.3 Conclusion

The division between 'executive' and 'non-executive' type theories may be a blurred one. In general, the theories that would fall under the non-executive class tend to be biologically or computationally driven. Interactions within the brain consist of small subsystems, units, or cells cooperating with each other as a network on an equal level. Executive-type theories, on the other hand, appear to be more behaviourally inspired. The frontal cortex, in conjunction with other brain areas, is necessary for directing attention, storing and retrieving information, and performing non-routine tasks. Yet a good model of the frontal region should require a combination of the two. One cannot propose a model of the brain without acknowledging that it is a network of interacting units and systems. Yet, as some 'low-level and simple' models (e.g., Cohen & Servan-Schreiber (1993) and Kimberg & Farah (1998) - described in Section 4.1)- have shown, these non-executive and non-suppressing subsystems may interact and give rise to controlling, executive type behaviours.

Our approach to the problem has been a combination of the two views. In our model, there are subsystems of neural networks (a prefrontal network and a hippocampal network) that iteratively communicate with and cue each other. As one system learns and acquires new information it feeds the modulated information to the other system. This iterative passing of information occurs during both encoding and retrieval. Patterns of activity among the neurons are stored, changed and passed on to the other system which uses these cues to produce a new pattern of neuronal activity. Although no system in our model actively 'directs' the operation of the other through suppression, both systems do 'suggest' operations by 'cueing' each other. This does not exclude the possibility of any inhibitory or suppression mechanism, however, as we will discuss in later chapters.

In our model, information from different sensory and semantic cortical areas is consolidated and stored in the hippocampal region. The recurrent attractor network we use for our hippocampal module is capable

of the rapid learning, pattern association and pattern completion necessary for recognition and recall of this episodic information. This representation of the hippocampal region is extremely simplistic as it captures only a small fraction of its structure and what we believe the hippocampus to be capable of. As discussed earlier, there is evidence that the hippocampal region is used in spatial and temporal mapping, consolidation and other functions. However, for the purpose of our present task, the simple elegant architecture and learning algorithm provided by the Hopfield network to simulate rapid encoding and pattern recall was sufficient.

The prefrontal cortex module of our model learns a less-detailed representation of the information given to the hippocampal module. It learns to extract and represent only the features that are shared among the input. These shared features may be, for example, semantic category memberships or contextual list memberships. In addition, the reinforcement learning algorithm we use to train the network allows it to find and use the best sequence of 'states' and 'actions' the network should use for optimal recall. Without the prefrontal module, our hippocampal module can perform pattern recognition and recall but it is not capable of organizing retrieval strategies.

To simulate human performance on free- and cued- recall tests via these simple, single-layered hippocampal and prefrontal modules was a true challenge. Does there exist a single underlying learning rule by which they could learn? Would the activities between these sets of 'cells' manifest into the complex human behaviours described earlier? As will be seen later in this thesis, the simple underlying learning rule by which the prefrontal module was governed led to a variety of behaviours. The simulation of one function led to the natural simulation of many others. Yet, as is described in the discussion chapter of this thesis, not all prefrontal attributes could be imitated which suggests that additional layers or sets of 'brain regions' or different learning rules may be needed to model a fully functioning prefrontal cortex.

Chapter 4

The Model

4.1 Other Computational Models of the PFC

Other computational modeling attempts have been made to account for deficits seen in patients with frontal lobe damage. We describe some of these models before describing our own. Guignon et al. (1995) and Moody et al. (1998) [27, 54] designed models to account for learning in the prefrontal cortex in primates during a delayed matching-to-sample task. Moody et al's model also simulated the basic use of temporal context. They did so with a real-time recurrent learning neural network architecture that was capable of holding internal state information over some delay period. Guignon et al.'s connectionist model accounts for stimulus discrimination, motor initiation and sequence of actions (temporal organization). Their model of the prefrontal cortex was also used for temporal organization. This involvement was expressed by processing units that switched between two stable states of activity (that is, they exhibited bistable behavior) in response to synaptic inputs. Similar to our model, there was sustained activity during the delay between a sensory cue and the production of a response, and there was a global increase for reinforced sequences and a decrease for non-reinforced sequences. Also similar to our model, unit activations were correlated with behavioral processes like expectation of forthcoming signals and preparation of a behavioral action. Before learning, weights had low values, thus were unlikely to become activated. Repeated presentation of a reinforced sequence led to increased weights. After learning, a unit became activated when the first event of the sequence was presented, predicting the occurrence of a reinforcement. This rule related the quantity of sustained activity to the predictability of reinforcement.

Cohen and Servan-Schreiber (1993) [10, 11] modeled the underlying cognitive and linguistic impairments in schizophrenia which is thought to be due to frontal lobe damage. A test of predictions was performed

in simulating the stroop task, continuous performance task and interpretation of ambiguous words using sentential context. Their neural network simulation differs from ours in that they offered a framework for exploring how dopamine (DA) effects at the physiological level affects behavior: DA modulates neuronal activity by enhancing the ability of neurons to transmit signals. This is achieved by increasing the "gain" of individual neuron-like units. In their model, frontal impairment results from an abnormality in the activation of these units that represent context. Specifically, there is a decrease in the "gain" parameter which corresponds to the effect of lowered DA levels in real neurons. Our model is similar to theirs in that it attempts to account for failure on several superficially distinct tasks by hypothesizing an underlying change in one system. It differs in that Cohen and Servan-Schreiber propose that this system is a reflection of the response properties of the individual neurons rather than the strength of the network associations between them. A model by Braver, Cohen and Servan-Schreiber (1995), also simulated performance on the continuous performance task and predicted behavior based on prepotent responses. Their model predicts an increased tendency of responses made more "prepotent" due to an increased frequency of presentation on trials. Our model is similar in that this high frequency of rewarded responses will lead to a higher evaluation of a specific action.

Dehaene and Changeux's (1991), Bapi and Levine's (1994), and Levine and Pruiett's (1989) models [13, 2, 43] simulated WCST performance. In these models, selection of a sorting criterion can be affected by mutually inhibitory nodes that correspond to the three possible sorting criteria. The activation of these nodes is determined in part by a reinforcement signal (the feedback of the experimenter) and by a representation of what Levine calls "habit nodes" - a system that reflects the frequency with which particular criteria have been selected in the past. Within these models, frontal lobe damage is modeled as a weakening in the gain of the reinforcement signal: a parameter that determines the weight given to feedback. Unlike these models, our model's units are not pre-assigned any contextual meanings (e.g., 'sort cards by colour'). Rather, our network adaptively discovers the best way to represent context that captures the information it was given.

Kimberg et al [38, 37] modeled the prefrontal cortex via a simple system based on levels of activity within what they termed "working memory". Their model is based on Anderson's ACT-R model of cognition. That is, behavior on the Stroop, the WCST, motor sequencing and contextual memory tasks was modeled by a series of "productions", or rules for actions. These were formulated in the form of "If X...THEN...do Y". A specific example would be: "If the goal is to name colours and the stimulus is red THEN say 'red'". This system modeled the association between goals, stimuli and stored knowledge as well as the ways in which these associations guided behavior. Our model is closest in spirit to Farah et al.'s model (with a few adaptations from Guignon's model), but implemented within a neural network and reinforcement learning

framework.

Most of the aforementioned models focus on the perseverative and inhibitory properties of the frontal cortex. Our model is unique from these models in the following aspects:

1. The other computational models are hand-tailored or hard-wired to solve a particular task. For example, a cell or cluster of neurons is pre-assigned to represent 'recall all things red' or 'recall all spices', or whatever particular concept is needed for that specific experiment[1]. We assume that the prefrontal cortex can generate unique cues for the memory system and can dynamically form new concepts that alter stimulus response pathways.

2. Other models depend on an immediate, external signal of reinforcement. Those models simulate performance on tasks that allow immediate feedback from the experimenter, such as the Wisconsin Card Sorting Test, the delayed matching-to-sample task, the continuous-performance task, and the Stroop test. The paradigms of these tasks allow the experimenter to give immediate, external reinforcement to the network during performance. This provides a direct teaching signal to the weights that influences the network's next output. Take a simulation of the Wisconsin Card sorting test. The network may be following the rule 'sort cards by colour' because it has been consistently rewarded for following that rule. Suddenly, the experimenter punishes the network because the sorting rule has been changed. This allows the network to easily suppress the unit that represented the 'sort by colour' rule and allow other eligible rules to compete for future sorting actions [13]. However, there are many other tasks that are sensitive to prefrontal damage and do not allow immediate feedback from the experimenter (e.g., the California Verbal Learning test (CVLT), verbal fluency test, temporal source memory tests, and the AB-AC list learning test). Free recall of a short list of words comes so easily to most humans that the simulation of such a task may seem simple. The mechanisms that underlie such a process are actually somewhat complex. During free and cued recall, the network is neither punished nor rewarded by the experimenter during recall. It is merely presented with the list of words again, if the experiment is not finished. In this thesis, we attempt to tackle some of the latter, more difficult to simulate tasks and we use internally generated reinforcements to accomplish this.

[1] An exception to this is the computational model proposed by Cohen and Servan-Schreiber (1992). Their model simply treats the prefrontal cortex as a recurrent hidden layer with no special control function apart from saving context. In this respect, our model is similar to theirs.

4.2 Our Model

Our goal was to create a model of the hippocampal-prefrontal regions of the human brain, to capture their interaction and the prefrontal cortex's ability to strategically recall and use context to aid in information retrieval. For reasons discussed in the previous chapter, free recall of a related word list is a good task that demands the prefrontal cortex's ability to categorically organize study list items and generate adequate retrieval cues. Also previously described was how the AB-AC list learning task taps into the prefrontal region's purported ability to maintain distinct contextual and temporal information on the lists. Both tasks require encoding, manipulating and organizing information - characteristics attributed to the prefrontal cortex. Besides learning to recall words correctly, other related behaviours our network had to exhibit in order to perform successfully was:

- Non-perseveration. Not only did the network have to self-organize or 'cluster' information into appropriate cues, it must also detect *which* cue is appropriate to use during retrieval and *when* to use that cue. Every prefrontal cortex model created so far used external negative reinforcement signals to tell the network when to 'turn off' a stimulus cue and when to 'turn on' a new one.

- Retention of other types of memory. In neuropsychology, function of prefrontal cortex is inferred by studying patients with damage concentrated on the frontal lobes. Likewise, the effect our prefrontal cortex layer has (if any) on the network can be observed by 'damaging' its prefrontal cortex. In any case, we should bear in mind that damage to the prefrontal cortex does not lead to outright amnesia or extreme memory deficits. Patients with prefrontal damage are still capable of recalling some list words in a task like the CVLT and remembering associated word lists (like the AB list in the AB-AC list test). They are merely very susceptible to proactive interference and have poor subjective organizational abilities.

Our job was made more challenging by our stipulation that:

- The network had to self-organize the context and set of cues it would later use during retrieval. As mentioned earlier, previous models had hard-wired units to represent these cues.

- The 'vocabulary' of words that the network possessed was to be larger than what was needed to perform these tasks.

To model this hippocampal-prefrontal interaction, we created the network architecture shown in Figure 4.1. We wanted to test our model on free- and cued-recall word memory tasks, so we needed a way to

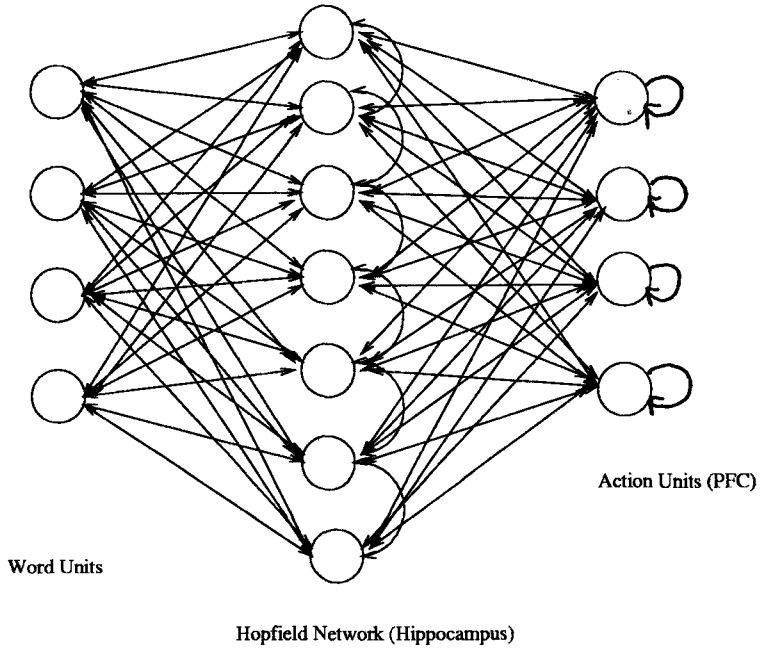

Figure 4.1: *The architecture of our model.*

represent the network's vocabulary, its word inputs and its outputs. On the left side of this figure there is a set of units which is referred to as 'Word Units' that function as the input and output layer of the network. These are localist representations of the network's 'vocabulary' that convey both word form and semantics. Since these units are trained prior to simulating any experiments, they also model the pre-experimental knowledge possessed by the network. This layer projects divergently to the hippocampal region simulated by an autoassociator. The connections between the Word Units and the hippocampal module are trained before each study phase of an experiment so they are considered to model pre-existing representations of vocabulary items.

Our model has two hidden layers that learn during the actual experiment simulations. These components are:

1. an autoassociator (the hippocampal module) that maintains an aspect of episodic memory (e.g., word-word associations);

2. a set of units (the prefrontal module) that provide cues to the hippocampal module. The activities of the units in this prefrontal cortical layer correspond to a particular action and represent the cues that probe the other memory systems.

Both of these components will be discussed in detail later.

4.3 Input/Output Layer

The input layer consists of units coding for a single word. Each unit in the Word Unit layer is a localist representation of each word that forms the network's vocabulary and serves as both the input for the hippocampal layer and the output of the network. These units signal for the occurrence of a single word by an all-or-none code (a unit's value is 1 if the word it represents is being currently studied or recalled, and is zero otherwise). Because this layer projects divergently to the hippocampal module, the representation of a word at the hippocampal layer is a distributed vector of semantic and perceptual features. The distributed representation of a word was pseudo-randomly created so that semantically related words were correlated with each other (Figure 4.2). Because the connections between the Word Unit layer and the hippocampal layer are bidirectional and symmetric, a distributed representation of a word can be reversibly mapped to its localist representation at the Word Unit layer. The final output of the network is a single 'word' represented by the winning Word Unit. Note that we are not attempting to model how we think semantic memory actually operates; rather, we are simply using a localist representation at the word layer as a convenient way of generating input for the hippocampal layer and mapping its output into discrete responses.

For every pattern that comprised the network's entire vocabulary of words, the weights between the Word Unit layer and the Hippocampal layer were trained using the Hebbian learning rule (Equation 2.3, Section 2.1.3) prior to the first experimental study phase. Thus, this part of the network could recognize even experimental non-studied list words.

Another mechanism was included that affected what words were output. We made the assumption that the network was capable of remembering the most recently recalled words. This could be modeled by an external 'short term memory' system that could store five to seven word items. We used such a system to store the six most recently recalled word items. When a word was output by the network, it was stored in

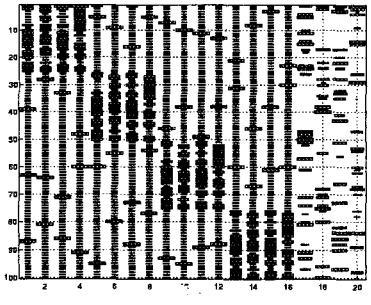

Figure 4.2: *Sample input patterns (size 100; sparseness .4; and correlation .5) that may compose the network's vocabulary. A column represents 1 word. Shown here are 20 words: 4 categories of 4 words each and a set of 4 vectors of the same sparseness randomly generated.*

this short-term memory system and the oldest stored word item (the word that was recalled 7 recall attempts ago) was displaced. When a word was output, it was compared to the words kept in this short-term memory store. If it matched any of those words then a negative reward signal was given to the network. In short, the network was punished when it attempted to recall the same word it had already recalled recently. One could view this as a 'filtering' mechanism: a subject's natural inclination may be to recall a recently retrieved word, but conscious and deliberate comparison of this word to the words kept in this short term memory buffer would cause the subject to try to recall a different word.

4.4 The Hippocampal Autoassociator

At the center of Figure 4.1 is an autoassociator. This is our simplified representation of the hippocampal region. Our choice of an autoassociator to represent the hippocampus is based on:

- Previous work that suggests rapid encoding and pattern recognition/recall capabilities are desirable
- On the choice of other modeling attempts for modeling episodic or event memories in the hippocampus, especially within hippocampal CA3 field which shares many of the basic connectivity requirements for an autoassociator [24, 71, 84, 28].

Our autoassociator was a Hopfield network that was modified to store the sparse, continuous valued patterns that represented each word item on the study list. Hence, the network's weights and units were updated according to Equations 2.16 and 2.17 (Section 2.1.11).

During the recall phase, the hippocampal network settled into an attractor state as the result of an incoming cue from the prefrontal region of the model. Thus, if the autoassociator settled into a stable pattern that was similar to, say, the word 'chisel' then the particular Word Unit that represented the word 'chisel' became active and was likely the network's next output. The weights from the hippocampal region to the Word Units were bidirectional and symmetric, thus the winning Word Unit output also served as input to the hippocampal region for the next recall attempt. The cycle continued until the network had enough opportunities to recall all the words from the study list.

4.5 The Frontal Cue Units

The third component of our network is a layer of units called the Frontal 'Cue Units' that is bidirectionally and symmetrically connected to the hippocampal module. This layer forms the basis of prefrontal cortical

activity that is used for tasks that require strategic recall. The frontal Cue Units' weights were trained using Equation 2.15. As described in Section 2.1.10, this reinforcement learning rule allows the units to maintain a temporally blurred 'history' of past activities or actions taken. It will be shown that this learning algorithm aids the network in clustering words that are separated in space and time and in generating a form of context or source memory during learning.

Chapter 5

Experimental Methods and Results

We tested the individual components of the model on simple tests before evaluating the entire neural network on the California Verbal Learning and AB-AC List learning tests. In this chapter, we present and discuss the results of each experiment.

Experiment 1 We focused on the Hippocampal component of the model. Its ability to perform rapid learning and recall was assessed under conditions of varying pattern sparseness, sizes, and correlations.

Experiment 2 We focused on the interaction between the Hippocampal and Word Unit layers. A simple test was used to evaluate its ability to select the correct Word Unit when cued with words from the original list under various 'background vocabulary' size conditions.

Experiment 3 We added the Frontal Cue Units layer and simulated the California Verbal Learning Test. The network in this experiment had a background vocabulary of 24 words. One set of results obtained in this experiment reflects how the network performs when allowed to learn only under the Study Phase of the experiment and not during the Recall Phase. That is, weights were updated with reinforcement learning as the word list was presented to the network and 'clamped', or frozen, during the recall stage. This was compared to how the network performed when it learned during both study and recall.

Experiment 4 Using the same architecture, we tested the network on a more difficult version of the CVLT. The network was trained on and possessed a larger background vocabulary of 52 words, many of which were similar to and highly correlated to the original study list items. Two versions of a 'frontal-lobe damaged' model were tested against our 'normal' network. In one frontal lesion design, the frontal weights were prevented from learning. In the other frontal lesion design, the frontal units were prevented from maintaining a memory of its past activities.

Experiment 5 We used the same network from the previous experiment and evaluated model performance on the AB-AC list learning task. This model was tested on two sets of data. The first, which is referred to as the AB-ACS data set, is composed of a list of word patterns in which the second list word pairs are semantically similar to the first list of word pairs (correlation between patterns from both sets were high). The second, referred to as the AB-ACD data set, is composed of a list of word patterns in which the second list word pairs are semantically different from the first list of word pairs.

5.1 Experiment 1 - The Hippocampal Autoassociator

It has previously been proven that the memory capacity of the Hopfield network depends on the sparseness, dimension size and correlation of stored patterns ([30, 85, 86]). The storage capacity of the original Hopfield network with non-sparse, orthogonal [1] patterns is approximately 0.14 times the dimension size of the patterns to be stored. It has been shown that as patterns become more correlated, the capacity of the network decreases. In a similar vein, if the patterns are more sparse (less bits are turned on), the likelihood of the patterns becoming decorrelated is greater and the capacity of the network improves [9, 26, 86].

5.1.1 Correlation of Patterns

For the California Verbal Learning Test (CVLT) and the AB-AC list learning task, our network must be able to work with highly correlated patterns. For the CVLT the network must store a minimum of 16 patterns - four different 'categories' of four correlated patterns. The representation of a word consists of a distributed vector of semantic features, with semantically related words correlated with each other. Thus, words within a category will be more highly correlated than words between categories. Correlation between vectors was calculated by taking the cosine angle between the two vectors.

5.1.2 Methods

To examine the effects sparseness, dimension size and correlation of patterns have on the performance of our hippocampal module, the autoassociator was tested on sets of sixteen pattern vectors of varying sparseness, dimension sizes, and similarities. We tested the hippocampal component on 16 patterns with dimension sizes of 100, 200, 300, 400, 500 and 600. For each dimension size, we applied sparseness and correlation levels of 0.25, 0.50, 0.75 and 0.90. The Hippocampus' weights and units were updated using Equations 2.16 and

[1] To be orthogonal, the inner product between pattern vectors must be zero.

2.17 (Section 2.1.11). Recall performance was tested by presenting the network with the actual stored input patterns degraded by 25%.

5.1.3 Results and Discussion

Two measurements were used to assess autoassociator performance across all experiments. Hamming distance is the number of bits that differ between two equal-sized vectors. The lower the hamming distance value the more similar the two vectors are; thus a hamming distance of zero indicates identical vectors. Another measurement of performance is the dot product between the input pattern and the actual output of the network. A higher value indicates better autoassociative performance than a lower value. Overall performance is determined by whether the network's output has the maximum dot product with the intended input pattern - that is, if it made the best associative match.

Figures 5.1 - 5.4 show error levels as a function of sparseness and overlap. We compared varying sparseness levels (0.50, 0.75, and 0.90) at overlap levels of 0.25, 0.50, 0.75 and 0.90.

In general, as sparseness increased, errors decreased. When pattern dimension size was 100, a sparseness level of 0.25 led to a 9% error rate whereas a sparseness level of 0.75 yielded 1% errors (Figure 5.1). Even as overlap levels were increased to levels of 0.50, 0.75 and 0.90, the 0.25 sparseness level yielded 17.5%, 11%, and 6% errors respectively whereas the 0.75 sparseness level yielded 8.5%, 6%, and 2% errors respectively.

As pattern dimension size increased, errors decreased even more (Figures 5.1 and 5.2). However, as overlap between patterns increased, these trends became less defined. This was expected since the number of overlapping elements between patterns was being increased. However, in general, higher sparseness levels still yielded less errors. There were less errors produced at a 0.75 sparseness level than those produced under a 0.25 sparseness level and errors decreased as pattern dimension size increased. At extremely high overlap (0.75 to 0.90) levels, however, the benefits of using very sparse patterns was negligible. Very similar patterns (due to the high number of overlapping elements) were difficult to distinguish from each other despite high sparse levels. In fact, it would seem that a 0.90 sparseness level exacerbated declining performance. This is not surprising since only 10% of the bits are 'turned on' and of those bits, a high percentage of them are the same between patterns.

Extremely low sparseness levels (0.05) were also tested as a function of varying overlap values. At overlap values of 0.50 and higher, performance was extremely poor (90-100% errors) regardless of pattern size. At overlap levels of 0.25, performance was only slightly better at 85% errors.

In summary, it has been shown that the capacity of a Hopfield network is generally 10-15% of the stored pattern dimensions but this capacity depends on the sparseness of the patterns being stored and their orthogonality. For the purpose of our experiments, sufficient overlap must exist in the patterns to denote 'category membership'. As can be seen from these results, too much overlap will make unique, individual patterns indistinct from each other. Storage capacity improves if sparse vector patterns are used but if vector patterns are too sparse the shared semantic features will not be detected by the network. A balance must be struck. We decided to work with a minimum pattern overlap within categories of 0.50 and a maximum sparseness among patterns of 0.75. Therefore, for the specific task of storing only 16 patterns a minimum vector size of 200 is necessary to obtain no more than 5% error.

5.2 Experiment 2 - The Word Units

The purpose of this experiment was to ensure adequate performance of the Hippocampal-Word Unit layer components of the model. The size of the network's 'background vocabulary' was determined by how many Word Units were used. Therefore, if the Word Unit layer had 16 units, the network had only 16 words to choose from during recall. Given that the study list in the CVLT has only 16 words, if the network adopted a strategy of only recalling exactly one different word in the recall phase it would show perfect performance. To prevent trivially successful recall, we made the vocabulary size larger than 16 word items.

5.2.1 Methods

Again, we pretrained the Hippocampal-Word Unit weights using Hebb's learning rule on either a vocabulary size of 24 words (Experiment A) or 32 words (Experiment B). The 'study list' was composed of 16 of these patterns (4 different 'categories' of 4 unique patterns - within category correlation of .20 and between category correlation of .01). The remaining word patterns were randomly generated vectors of equal size and sparseness (.25).

A single trial consisted of a study phase and a recall phase. During study, the hippocampal weights were trained on the 16 word patterns using Equation 2.16 and the Hippocampal-Word Unit weights were trained with the Hebb rule. During recall, the Hippocampus was cued with each of the word patterns from the original study list. This pattern of activity served as input to the Word Unit layer. Each unit's activity in

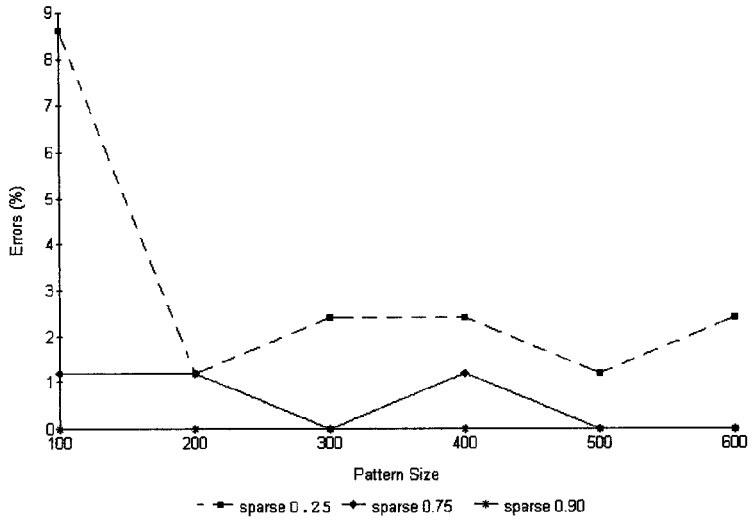

Figure 5.1: *Performance of the autoassociator as a function of sparseness (Overlap = 0.25).*

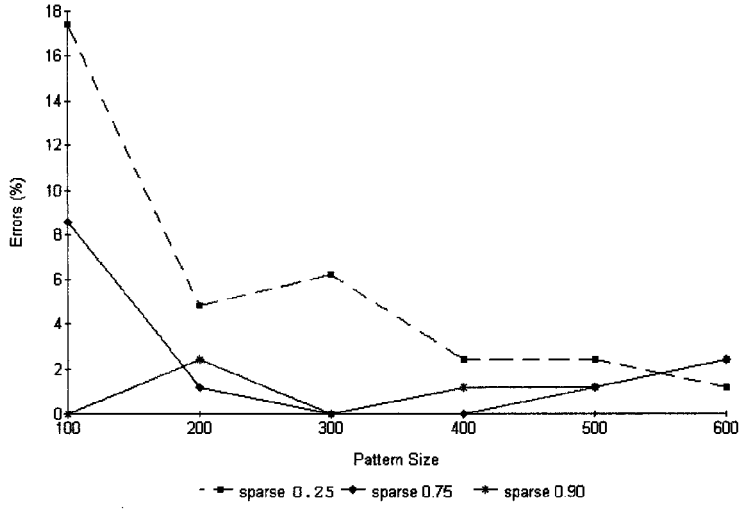

Figure 5.2: *Performance of the autoassociator as a function of sparseness (Overlap = 0.50).*

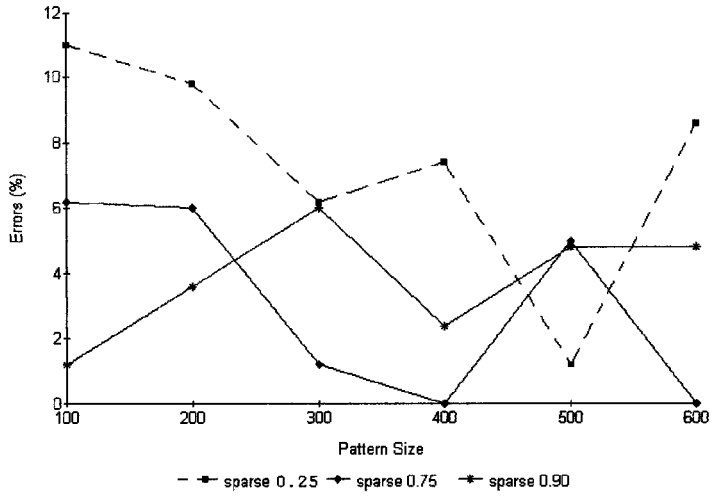

Figure 5.3: *Performance of the autoassociator as a function of sparseness (Overlap = 0.75).*

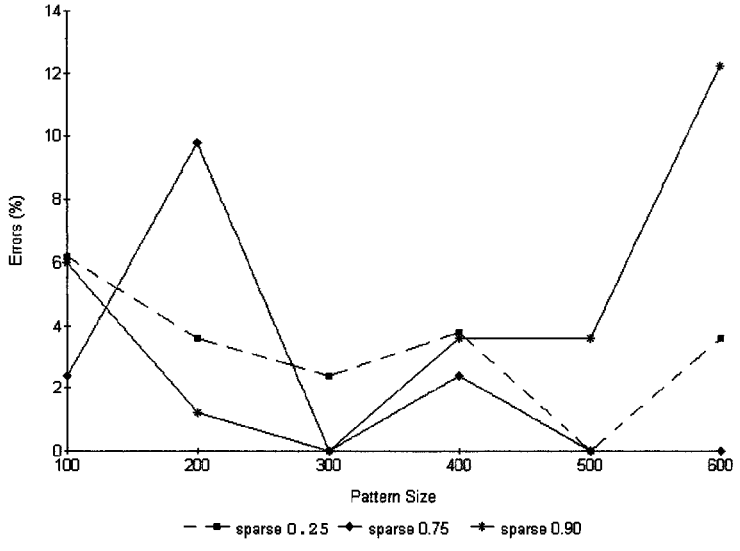

Figure 5.4: *Performance of the autoassociator as a function of sparseness (Overlap = 0.90).*

that layer was transformed with the softmax probability function proposed by Bridle (1990):

$$P(A = a_i) = \frac{e^{x_i}}{\sum_{j=1}^{n} e^{x_j}} \tag{5.1}$$

and the Word Unit with the highest probability value was selected to be the network's output.

5.2.2 Results and Discussion

	Trial 1	Trial 2	Trial 3	Trial 4	Trial 5
Correct Words Recalled (%)	56.25	68.75	73.75	82.75	95.25
Correct Class Recalled (%)	80	100	100	100	100
Nonlist Words Recalled (%)	8.33	0	0	0	0

Table 5.1: *Performance of the Hippocampal-Word Unit system when background vocabulary size is 24 (sparseness .25; learning rate .04). This table shows the percentage of correct words recalled, correct classification of the word (i.e., recalled word is of the same semantic category as the desired target word), and non-list words recalled.*

	Trial 1	Trial 2	Trial 3	Trial 4	Trial 5
Correct Words Recalled (%)	25.00	43.75	68.75	81.25	93.75
Correct Class Recalled (%)	81.25	100	100	100	100
Nonlist Words Recalled (%)	8.33	0	0	0	0

Table 5.2: *Performance of the Hippocampal-Word Unit system when background vocabulary size is 32 (sparseness .25; learning rate .04). This table shows the percentage of correct words recalled, correct classification of the word (i.e., recalled word is of the same semantic category as the desired target word), and non-list words recalled.*

Tables 5.1 and 5.2 show the results of testing the network with a background vocabulary of 24 and 32 words respectively. With each additional trial, the network's ability to select the correct word increased. However, even if the network did not select the correct word it often selected a word from the same category membership as the cue word pattern (at least 80% of the time on the first trial). The network showed only a small tendency to erroneously recall non-list words at the beginning and didn't recall any by the second trial. The data also indicate that an increase in vocabulary size does not increase the likelihood of recalling non-list words.

These results show that even if recall performance is low, the networks' errors lie in recalling word patterns prototypically similar to the cue word pattern instead of recalling non-list words, regardless of background vocabulary size.

5.3 California Verbal Learning Test Paradigm

Building upon the architecture of the first two experiments, Frontal Cue Units and a bidirectional set of weights from this layer to the Hippocampal module were added.

In the California Verbal Learning Test (CVLT), subjects undergo five successive trials of learning and recalling a word list. A trial, in our simulation, is defined as one study and one recall phase. So to simulate this experiment, our network was allowed five trials of study and recall - this simulated a 'subject' undergoing a single experimental condition. The list of words was presented in the same order for each study session and the connection weights were not reset to zero between list presentations.

During study, each word pattern was presented to the Hippocampal module. After the network settled into the closest stored memory attractor, the corresponding Word Unit was trained. The selection of a Frontal Cue Unit for each pattern depended on several factors. It was influenced by the pattern of activity exhibited by the Hippocampal module at that time: each Cue Unit receives input from the Hippocampal component in the form of that layer's weighted summed activations. It was also influenced by its past activity levels. As explained in Section 2.1.10, the learning rule by which we update the Frontal layer's weights is influenced by $\overline{y_i}$ - a weighted average of a unit's previous outputs. For ease of reference, this learning rule is reprinted here:

$$w_{ij}(t+1) = w_{ij}(t) + \alpha[r + \gamma y_i(t) - \overline{y_i(t)}]x_j(t)$$

where

$$y_i(t) = f(\textstyle\sum_j w_{ij}x_j) \quad f(\bullet) = \frac{1}{1+e^{-x_j}}$$

is the summed current weighted input coming from the Hippocampal layer transformed by the sigmoidal function, and

$$\overline{y_i}(t) = (1-\delta)\overline{y_i}(t-1) + \delta y_i(t) \quad 0 \leq \delta \leq 1$$

This last term shows that the parameter δ determines how much influence previous values of y have on the current value of $\overline{y_i}$. As δ approaches 1, $\overline{y_i}$ is determined solely on the current output of the Hippocampus.

The selection of a Frontal Cue Unit for each pattern also depended on r - the reinforcement signal that, in our case, takes on the values $+1$ or -1. When $r = 1$, as it was during the study phase, weights between

active Hippocampal Units and the selected Frontal Cue Unit were strengthened.

Only a select few Frontal Cue Units were allowed to correspond with each pattern vector. To achieve this, the sigmoidal outputs of each Frontal Cue Unit were transformed with a 'soft k-winner-take-all activation function'. For each new activation vector, a bias term was dynamically chosen such that only k 'winners' would have a positive net input; each unit's net input was shifted by this amount, and then transformed through a sigmoid nonlinearity (Equation 2.2) to compute a real-valued activation. We then applied the softmax nonlinear activation function to each unit. This allowed us to obtain a differentiable form of soft-k-winner-take-all for each layer.

The selection of a Frontal Cue Unit for each pattern also depended on how greedy our network was. The output of every Cue Unit was transformed by the softmax probability function. This forces the set of n units in the Frontal layer to represent a probability distribution over the n states and adopt states whose probabilities sum to one. If our model was fully greedy, it would select that unit with the highest probability value for each word and strengthen only that unit's weights. If it was set to be 50% greedy, then it would 'explore' other actions half the time and strengthen weights to units that did not have the highest probability value.

During recall, a Frontal Unit was chosen to serve as the cue to the rest of the system. Selection of a Cue Unit depended on the weighted average of its previous outputs: thus, a Cue Unit that was selected often during the study phase would likely serve as a cue during the recall phase. In other words, the Cue Units maintained a blurred history of previous activities and often active units were encouraged to remain active.

During recall, the selected Cue Unit served as the input to the Hippocampal module, which in turn provided the network with the information to choose the best Word Unit. The 'winning' Word Unit became the network's output. It also served as a cue to the Hippocampus, which in turn led to the selection of the next best Cue Unit to serve as a cue for the next recalled word. This 'volley' of cues between the Frontal layer, Hippocampal layer, and Word Units was allowed until the network had produced at least 16 words.

In Experiment 3A, we tested a model that is reinforced only during the study phase. In Experiment 3B, we tested a model that is positively reinforced during the study phase and is negatively reinforced ($r = -1$) during recall under conditions that will be made clearer later.

In the actual CVLT task, subjects are allowed to recall as many words as they can per trial. They may end up recalling only 6 words or they may recall 20 words. Obviously, by recalling 20 words they will make a minimum of 4 errors. To better understand the way our network was recalling words and to examine the types of errors it tended to make, it was allowed 20 attempts to recall a list for each trial - that is, it recalled 20 words per trial.

Word	Class	Word	Class	Word	Class	Word	Class	Word	Class
Parsley	spice	Plums	fruit	Drill	tool	Vest	clothes	Rug	*Other*
Paprika	spice	Grapes	fruit	Wrench	tool	Sweater	clothes	Aspirin	*Other*
Chives	spice	Tangerines	fruit	Chisel	tool	Jacket	clothes	Film	*Other*
Nutmeg	spice	Apricots	fruit	Pliers	tool	Slacks	clothes	Books	*Other*

Table 5.3: *The 20 words that represent the network's vocabulary. 'Other' words are the patterns that compose the network's non-list vocabulary. There are 16 'List' words, and 4 'Other' words. The 'other' words are unrelated to any of the other patterns and are randomly generated.*

5.3.1 Pattern Description

For the CVLT simulations, the dimension size of our pattern vectors was 400 and had a sparseness level of 0.25. As in Experiments 1 and 2, the representation of a word consisted of a distributed vector of semantic features, with semantically related words correlated with each other. Non-list words were randomly generated vectors of the same length as words on the list. Words within a category had a correlation of 0.20; between categories the correlation was 0.01.

For the first experiment, the network's vocabulary consisted of 24 words - 4 of which were random and 4 were 'members' of the same categories represented in the study list. Table 5.3 shows the words used for this experiment. In Experiment 4, the network's vocabulary consisted of 52 word patterns - 16 of which were list words - which possess phonemic and prototypic similarity to each other. How we conveyed phonemic and prototypic similarity between words is described below.

In the 52-word CVLT test, the pattern vectors were more complex than the 24-word CVLT test. Sixteen words were 'list' words. Words within a category had a correlation of .20, while between categories the correlation was .01. This left 36 extra-list or 'other' words. Twenty of these words were semantically similar to the 'list' words (thus, had a correlation of .20 with the actual list words) but were not studied during the experiment. Eight words formed 2 new semantic categories (each category contained four words each). These words had a .01 correlation between the studied list categories as well. Eight words were random, sparse vectors. These patterns had no or very little correlation with any of the other patterns. The sparseness levels of all pattern vectors was 0.25. The size of each vector was 500. Table 5.4 shows the words that were used. Although the 'other' words are based on phonetic similarity (e.g., 'grapes' and 'tapes') as well as semantic similarity (e.g., 'grapes' and 'plums'), no distinction was made between phonetics and semantics in our pattern representation. Thus, the patterns for 'grapes','plums', and 'tapes' all had the same within-category correlation with each other.

Word	Class	Word	Class
Parsley	List	Tires	*Other-PhS*
Paprika	List	Grill	*Other-PhS*
Chives	List	Hammer	*Other-PrS*
Nutmeg	List	Vest	List
Ginger	*Other-PrS*	Sweater	List
Oregano	*Other-PrS*	Jacket	List
Sage	*Other-PrS*	Slacks	List
Cinnamon	*Other-PrS*	Wax	*Other-PhS*
Pepper	*Other-PrS*	Shoes	*Other-PrS*
Pastry	*Other-PhS*	Racket	*Other-PhS*
Chimes	*Other-PhS*	Toaster	*Other-Utensils*
Plums	List	Spatula	*Other-Utensils*
Grapes	List	Skillet	*Other-Utensils*
Tangerines	List	Bowl	*Other-Utensils*
Apricots	List	Halibut	*Other-Fish*
Cherries	*Other-PrS*	Flounder	*Other-Fish*
Pineapple	*Other-PrS*	Cod	*Other-Fish*
Lemons	*Other-PrS*	Salmon	*Other-Fish*
Peaches	*Other-PrS*	Rug	*Other*
Apples	*Other-PrS*	Aspirin	*Other*
Drums	*Other-PhS*	Film	*Other*
Tapes	*Other-PhS*	Briefcase	*Other*
Drill	List	Clock	*Other*
Wrench	List	Books	*Other*
Chisel	List	Vitamins	*Other*
Pliers	List	Soap	*Other*

Table 5.4: *The 52 words that represent the network's vocabulary. These words were obtained form the actual CVLT administration tests (Delis et. al, 1987). 'List' words are those patterns that were used in the study list. 'Other' words are the patterns that compose the rest of the network's vocabulary. There are 16 'List' words, and 36 'Other' words. Within 'List', the correlation of patterns between categories was 0.01, and 0.20 within a category. This was modeled after the actual CVLT administration scale in which 12 of the 'Other' words were prototypically similar (PrS), and another 8 of the 'Other' words were phonemically similar (PhS) to the actual List items. Eight of the 'Other' words formed their own categories ('Fish' and 'Utensils'), in which words within the category were correlated in the same way as those words in the List categories. The final eight 'Other' words were unrelated to any of the other patterns and were randomly generated.*

Trial	Word Recalled (in order)	Category	FU	Trial	Word Recalled (in order)	Category	FU
1	Drill	Tool	3	5	Wrench	Tool	1
	Pliers	Tool	3		Pliers	Tool	1
	Wrench	Tool	3		Chisel	Tool	1
	Chisel	Tool	3		Drill	Tool	3
	Plums	Fruit	3		Aspirin	Other	1
	Aspirin	Other	3		Nutmeg	Spice	2
	Nutmeg	Spice	2		Parsley	Spice	2
	Parsley	Spice	2		Chives	Spice	2
	Chives	Spice	2		Paprika	Spice	2
	Paprika	Spice	2		Rug	Other	2
	Rug	Other	2		Film	Other	2
	Grapes	Fruit	1		Grapes	Fruit	5
	Apricots	Fruit	1		Tangerines	Fruit	5
	Rug	Other	1		Plums	Fruit	5
	Sweater	Clothes	6		Apricots	Fruit	5
	Slacks	Clothes	4		Aspirin	Other	1

Table 5.5: *Example output of the network on trials 1 and 5. 'FU' indicates the most active Frontal Cue Unit that served as the cue to the hippocampal memory systems to retrieve the recalled word.*

5.4 Experiment 3 - CVLT Methods (Part I)

An aspect of free recall tasks that makes biologically plausible computational simulations difficult is the absence of immediate, external reinforcement. For many human memory tasks, the subject is not provided with the correct answer after recalling each and every word item. The purpose of this experiment was to determine if training the weights during the study phase of the experiment (when external teaching signals are available) is sufficient for adequate performance. If not, then additional methods must be used. Preferably, any additional methods used would allow the network to make necessary adjustments based on self-monitoring and without external guidance.

After running a few preliminary simulations, the following parameter values were settled upon: $\delta = 0.40$, $\gamma = 0.40$, $\alpha = .05$. The learning rate of the weight update rule used for the Word Unit weights was 0.4. The network was set to be greedy only 70% of the time.

5.4.1 Experiment 3A

Using the 24 word patterns and the methods described in Section 5.3, we ran 20 simulations. Table 5.5 gives an example of the network's output for the first and fifth trials.

Figures 5.6, 5.7, 5.8, and 5.9 show the number of correctly recalled words and errors, the types of errors made, clustering performance and serial recall tendencies of the network across all 20 simulations. A full analysis of these errors, the correctly recalled words, categories recalled and cluster performance of the network is given in Section 5.4.3. For now, it should be noted that the network tended to intersperse non-list

Trial	Highest	Second Highest	Difference	Transition
1	vest	sweater	0.000381	*within*
2	sweater	slacks	0.000527	*within*
3	slacks	jacket	0.000353	*within*
4	jacket	apples	0.007188	*other (fruit)*
5	apples	wax	0.000196	*other*
6	vest	nutmeg	0.007202	*between*
7	parsley	nutmeg	0.000099	*within*
8	nutmeg	grapes	0.000122	*between*
9	grapes	tangerines	0.000579	*within*
10	apricots	tangerines	0.000145	*within*
11	tangerines	plums	0.000068	*within*
12	plums	tapes	0.000496	*other (within fruit)*
13	rug	grill	0.000019	*other (tool))*
14	drill	pliers	0.000082	*within*
15	pliers	chisel	0.000017	*within*
16	chisel	racket	0.008716	*other (clothes)*

Table 5.6: *Each time the network recalls a word, it makes its selection based on the softmaxed probability distribution over the Word Units. Shown here is the word with the highest and second highest probability values, the difference between the two values, and the relationship between the two words. 'Within' indicates two words within the same category. 'Between' indicates two words between different correct categories. 'Other' are non-list words (phonetic or semantic similarity these words have to correct list words are in parentheses). See Figure 5.5 for the corresponding graph.*

words among the correct words it recalled. Further analysis showed that when a non-list word occurred it led to the recall of another non-list word (usually of the same category as the first non-list word) 53% of the time; 44% of the time, the non-list words led to the recall of a correct word, but from a category different from the one that the network was in prior to recalling the non-list word; and only 3% of the time did it lead to the recall of a list word within the same category as the previous list word. This suggests that the network had a tendency to recall a non-list word before transitioning from one category to the next. Perhaps the analogy would be subjects 'searching' for another cue stimulus when they have exhausted all possibilities with the current cue or 'line of thought'. Thus, we decided to examine more closely the network's decision - making and choice of words during recall.

During recall, the current pattern of activity in the Hippocampus causes the layer of Word Units to become active. A Word Unit is selected based on the softmax probability distribution of its activities and is the network's output. We examined the difference in the probabilities between the 'winning' word and the 'runner-up' (that is, the Word Unit with the second highest probability value) throughout recall of the list. We found that when a word is followed by another word from within the same category, the difference between the probability values between the winner and the runner-up was negligible. However, when a word followed another that was not of the same category, the difference was relatively large and usually signalled the upcoming recall of a non-list word. Table 5.6 shows the output of the network at each recall

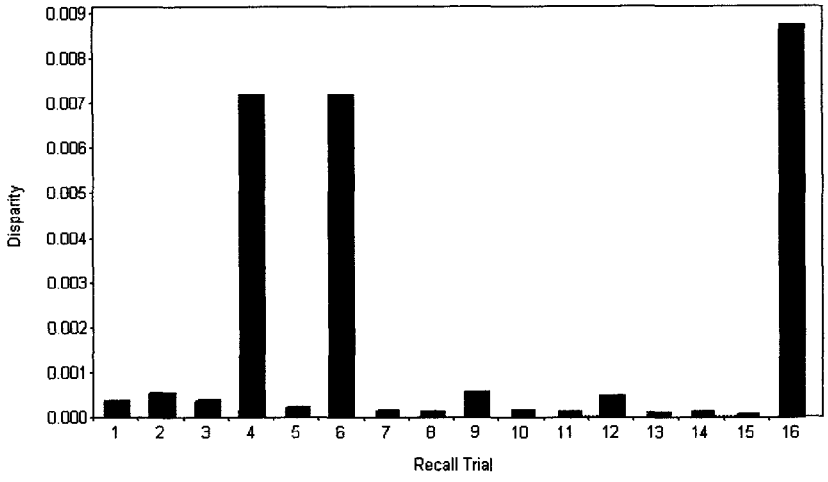

Figure 5.5: *Disparity between the softmaxed probabilities of the winning Word Unit (network's actual output) and the 'runner-up' Word Unit. See Table 5.6 for the corresponding data.*

step (the Word Unit with the highest probability value), the next highest word, the difference between the two probability values, and the relationship between the two words. Figure 5.5 is a graph of the same data.

When the runner-up was of the same category as the winner, the disparity was relatively small. When the winner and runner-up were members of different categories or one of the words was a non-list word of a different category, this disparity was generally large.

It may be useful to interpret this disparity as the network's 'lack of confidence' in choosing a particular Word Unit and use it as an internal punishment signal. After the network has recalled as many words as it can using the current Frontal Cue, it begins to detect the lack of available appropriate responses. The remaining words, whether they be studied or non-studied ones, are undifferentiable under the current Frontal Cue the network is using. Instead of selecting a Word Unit about which the network is 'uncertain', it may reconsider its options and select another Frontal Unit to serve as a cue to choose another Word Unit about which it *is* confident. If, after reconsidering its options, it decides that the same Cue Unit is still the best, or only, cue the network may keep that Cue Unit and end up producing that word about which it was originally uncertain as its output anyway. However, we may be able to use this disparity to prevent non-list words from being chosen and strengthen its confidence in selecting studied list words on future trials.

5.4.2 Experiment 3B

The patterns and methods used in this experiment are the same as in Experiment A except for one modification. During recall, the network provides itself with a negative reinforcement signal when it detects a large disparity between the activities of the 'winning' Word Unit and the 'runner-up'. This notifies the network when a change of cues is in order. The network only gives itself this negative reinforcement on the next recall attempt, *after* the large disparity has been detected. So, for example, let us say the network has already recalled a series of tools. When it recalls the word 'chisel' it notes that the softmaxed probability of the runner-up candidate word ('book') was extremely low. The disparity between the probabilities of the winning word and the runner-up from being chosen is beyond some threshold. This disparity primes the network to punish the sets of Frontal Cue weights when the next word is recalled. The punishment occurs regardless of what the next word being recalled is. As noted earlier, the probability that the next word would be a non-list word was greater than the probability that it would be a correct list word. Again, as in Experiment 3A, 20 simulations were run.

5.4.3 Results

Recall Performance

As explained earlier, the network was allowed 20 attempts to recall the word list items. As a result, the network made a minimum of 4 errors. The data in our graphs and tables include this minimum four errors baseline. It remains constant across all the simulations and measurements so it does not make a difference when comparing data.

A recalled word was classified as either 'correct' or as an 'error'. Errors consisted of perseverations (repetitions) or intrusions (not on the original list). Figure 5.6 shows the number of correct words and errors averaged over 20 experiments on a trial when learning was allowed to occur during recall and when it was not. It is evident that the first network makes more errors than the second. Figure 5.7 shows the breakdown of the types of errors made by each network.

Cluster Analysis

For each trial, three semantic clustering indices were derived that quantify the degree to which the network used a semantic strategy in recalling a list:

- An observed *correct* semantic clustering score (one point is obtained whenever the network reports a correct word after another correct word from the same category);
- An expected correct semantic clustering score (which indicates the chance value of semantic clustering

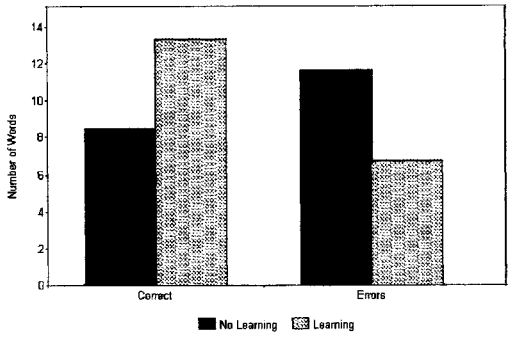

Figure 5.6: *The number of correctly recalled words and errors by two networks - one that learns during recall (Experiment 3B) and one that does not (Experiment 3A).*

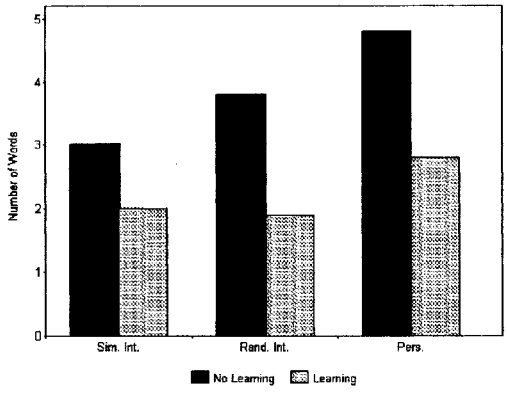

Figure 5.7: *The types of errors made by two networks - one that learns during recall (Experiment 3B) and one that does not (Experiment 3A). 'Sim. Int' are similar intrusions: non-list words that are semantically similar to words on the list. 'Rand. Int' are random intrusions: non-list words that are dissimilar to list words. 'Pers' are perseverative errors: list words that have already been recalled are recalled again.*

given the total number of words and the number of different categories represented in the examinee's recall on a particular trial), and

- The ratio of observed to expected correct clustering scores (where a ratio of 1 indicates a chance clustering performance and a ratio greater or less than 1 indicates above or below chance clustering performance, respectively).

The 'correct semantic cluster' score is computed by counting the number of correct responses that immediately follow another correct response from the same semantic category. Thus, the maximum correct semantic cluster score for each trial is 12. To be properly interpreted, the raw score has to be compared with the expected score based on chance [14], which is calculated by:

$$\sum_{n=1}^{4} \frac{Tn_i(Tn_i - 1)}{MX_i} \tag{5.2}$$

where n is the category type (four categories per list), Tn_i is the number of correct words recalled from category type n on trial i, and MX_i is the total number of words recalled on trial i, including intrusions and perseverations. This value indicates the chance value of semantic clustering given the total number of words and the number of different categories represented in the examinee's recall on a particular trial [14].

The observed correct semantic clustering scores were similar between the two networks (0.74, no learning during recall; .7212, learning during recall). When the ratio of observed to expected (1.62, no learning; 0.47 learning) clustering scores were calculated, however, it became evident that the network that did not learn during recall did not cluster above chance (cluster score = 0.63) whereas the network that learned (cluster score = 1.68) during recall did (Figure 5.8).

Serial Recall

A common alternative strategy to semantic clustering for humans is serial-order clustering – recalling items in the same order as they were presented. Indeed, a problem with training networks with a repeated, non-varying order of patterns is that the network tends to memorize the patterns *in order*. When cued with a pattern from the study list, the network uses it as a starting point to recall the patterns in the order it was studied. However, it has been noted [14] that on the CVLT, serial-order recall is generally a less effective learning strategy for people than semantic clustering and it correlates with poor performance on the CVLT.

An observed serial clustering score, a chance-expected serial clustering score and an observed/expected serial clustering ratio were also calculated based on methods used by Delis et. al. (1987) on the CVLT scales. The observed serial clustering score is the number of times that the network recalled two words in succession that also appeared in the same succession on the stimulus list. The expected probabilities of serial-order

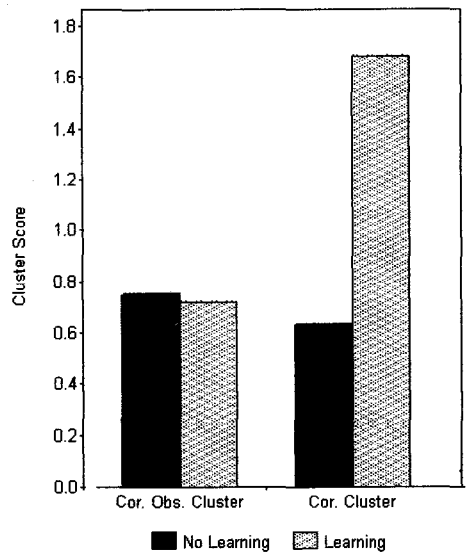

Figure 5.8: *The clustering scores of two networks - one that learns during recall (Experiment 3B) and one that does not (Experiment 3A). 'Cor. Obs. Cluster' is the observed correct semantic clustering score (one point was obtained whenever the network reported a correct word after another correct word from the same category). 'Cor. Cluster' is the ratio of observed to expected correct clustering score.*

clustering were calculated using a power function derived from a monte carlo study of 10,000 random lists of varying length [14].

The expected serial-order clustering score for trial i is:

$$SC_i = (.135 \times nC_i^{62}) - .135 \tag{5.3}$$

when $nC_i \geq 1$. Otherwise, when $nC_i = 0$, the expected serial order clustering score for trial i is zero. nC_i is the number of correct words recalled on trial i. According to Delis et. al. (1987), this equation was obtained from a power function derived from a monte carlo study of 10,000 random lists of varying length.

The indices described above allow a comparison to be made between the network's use of semantic and serial recall strategies. These indices also indicate if the network is using neither recall strategy, but rather is recalling words from the list in an idiosyncratic or haphazard manner. These results are shown on the left side of Figure 5.9. Both network models serial-order cluster above chance.

Another form of serial recall is possible; recalling words in the same order as the previous recall. The right side of Figure 5.9 shows the networks' tendencies for this type of serial recall. Where the network that did not undergo learning during recall was more likely to recall words in the same order as presented during

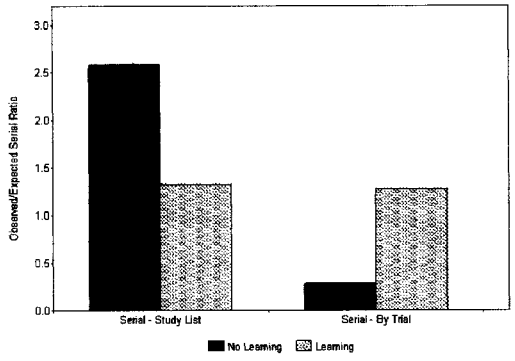

Figure 5.9: *The serial recall scores of two networks - one that learns during recall (Experiment 3B) and one that does not (Experiment 3A). 'Serial-Study List' indicate the number of times that the network recalled two words in succession that also appeared in the same succession on the stimulus list. 'Serial-by Trial' indicate the number of times that the network recalled two words in succession that appeared in the same succession as the network's previous recall trial.*

study, the network that did undergo learning tended to recall words in the order similar to its last recall trial. The network that did not undergo learning during recall failed to exhibit trial-serial learning above chance levels.

5.4.4 Discussion of Results (Experiments 3A & 3B)

The results from this experiment suggest that, in general, performance improved when the network was allowed to learn during the recall stage of the experiment. The network that learned during the recall phase produced less errors. Of the types of errors that are produced, perseverations are more common for both networks. Both networks produce intrusion type errors, but the network that does not learn during recall produced a disproportionate amount of random intrusions compared to similar intrusions. This is not evident in the network model that learns during recall. This suggests that the network that learns during recall has at least learned the relationship or similarities between the 'similar non-list' words and the list words.

Semantic clustering strategies typically result in effective encoding into long-term memory in people [12, 14]. Our data suggest that when a network was allowed to update its weights during recall, its clustering ability improved.

We presented the word list in the same order for each trial. The network that did not learn during recall showed a high tendency to recall words in an order similar to the order of words that were presented in the study list compared to the network that did undergo learning during recall. The fact that the network that learned during recall showed a lower tendency to recall words in the order that they were presented likely

reflects the network's tendency to cluster the items and recall them by category instead of serial order. The network that learned during recall also showed the tendency to recall words in the order similar to the order it had recalled words on the previous trial, whereas the network that did not learn during recall did not show this tendency.

5.5 Experiment 4 - CVLT Methods (Part II)

The methods used in this experiment were identical to those used in Experiment 3B except that the network possessed a 52-word vocabulary. The characteristics of the pattern vectors and their correlations were described in Section 5.3.1.

We also simulated two models of a frontal-damaged network. In one model, henceforth referred to as Lesioned Model A, we clamped the weights between the Hippocampal autoassociator and the Frontal Cue layer to zero. This effectively prevented any learning of the Frontal Cue Units to take place and is equivalent to removing the Frontal Cue Unit layer altogether. In the other model, henceforth referred to as Lesioned Model B, we restricted the Frontal Cue Units' ability to maintain a long-term 'blurred history' of past actions by changing the values of the parameters δ and γ. This would be a test of our model from a computational perspective. The weights will still learn in a Hebbian manner (connections between simultaneously active units are strengthened), but the advantage of using the temporal reinforcement learning algorithm will not be available.

The values of the parameters were the same in the current normal network as it was in the previous simulations. However, for Lesioned Model B, we set $\delta = 1$ and $\gamma = 0$. When $\delta = 1$, the term $\overline{y_i(t)}$ becomes $y_i(t)$. Since $\gamma = 0$, one of the weight update terms drops out leaving the weight update rule to be:

$$w_{ij}(t+1) = w_{ij}(t) + \alpha[r + y_i(t)]x_j(t)$$

With this equation, weights between the Hippocampal autoassociator and Frontal Cue Units are strengthened between any two simultaneously active units in the presence of positive reinforcement (or weakened in the case of negative reinforcement), but no Frontal Cue Unit is able to maintain a memory of its previous activity.

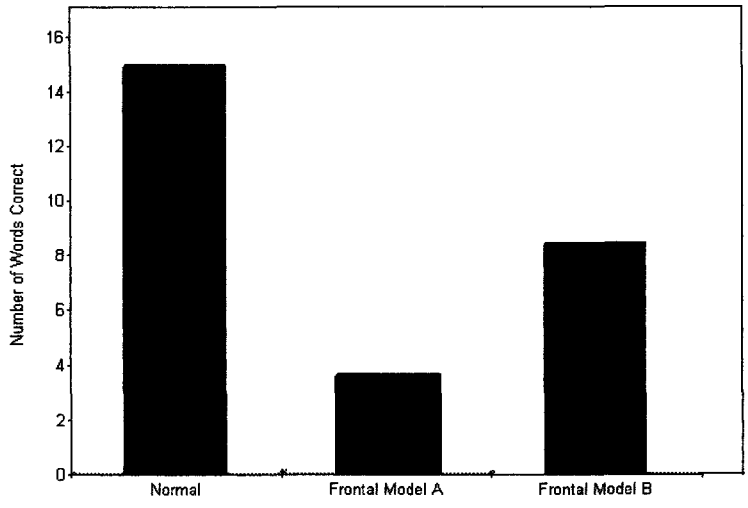

Figure 5.10: *The average number of correct words recalled by three networks: a 'normal' network, a 'lesioned' model with weights clamped to zero (Frontal A), and a 'lesioned' network with no memory for past actions and events (Frontal B). Data is averaged over 20 experiments.*

5.5.1 Results

In this section, we show the results of a 'normal' network and two versions of our 'frontal lesioned' networks. For comparison purposes, we present the number of words correctly recalled, the number of errors, types of errors, clustering and serial recall scores within the same tables and graphs.

Recall Performance

Errors and correct recalls were analyzed as described in the previous experiment. Figure 5.10 and Figure 5.11 show the number of correct words and the types of errors recalled on the third trial averaged over 20 experiments for a normal network, Lesioned Model A, and Lesioned Model B.

Analysis of variance revealed a main effect of differences between the three networks ($F = 424.94$; $d.f. = 2, 57$; $p \leq .01$). Sheffe tests showed that the normal network recalled significantly more correct words than Lesioned Model B, which in turn recalled more correct words than Lesioned Model A. Although the normal network produced significantly less intrusion errors than Lesioned Model B ($F = 153.817$; $d.f. = 2, 57$; $p \leq .01$), there were no significant differences in the number of perseverative errors produced ($F = 39.340$; $d.f. = 2, 57$; $p \leq .01$). However, both Lesioned Model B and the normal network produced significantly less perseverative errors than Lesioned Model A. Across all three models, similar intrusions tended to occur more frequently than random intrusions. That is, the network recalled more frequently non-list words that

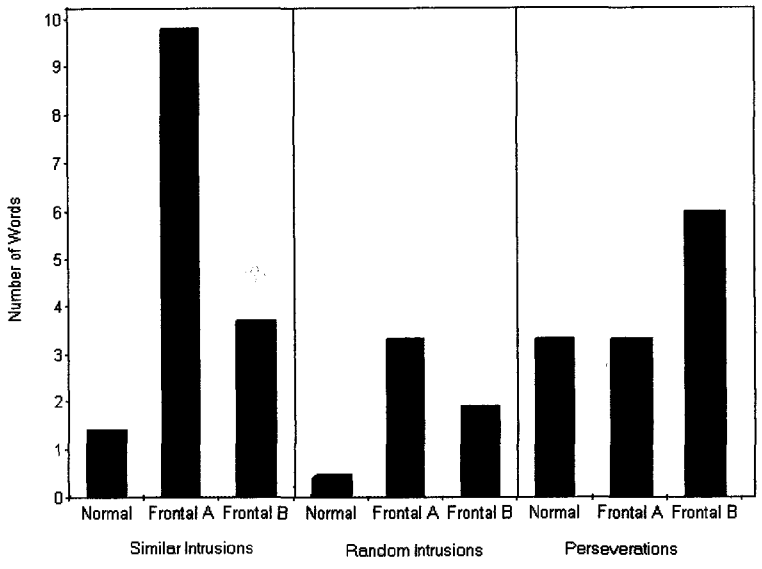

Figure 5.11: *The types of errors made by three networks during recall: a 'normal' network, a 'lesioned' model with weights clamped to zero (Frontal A), and a 'lesioned' network with no memory for past actions and events (Frontal B). Similar intrusions are non-list words that are semantically similar to words on the list. Random intrusions are non-list words that are dissimilar to list words. Perseverations are list words that have already been recalled but are recalled again. Data is averaged over 20 experiments.*

belonged to a category represented in the study list than other non-list words. There were 4 categories represented in the study list: spices, fruits, tools and clothing. There were 11 spices represented, therefore there was a 7 out of 11 (64%) chance that a non-list spice would be recalled. The same probability applied to any words recalled from the fruits category. Only 7 tools and 7 articles of clothing were used in the network's vocabulary so there was a 43% chance that a non-list tool or clothing was recalled.

We did a further analysis of the types of errors the three network models made. There were, in total, 6 categories of words available to the network during recall: spices, fruits, tools, clothes, kitchen utensils, and fish. In addition to naming the wrong spices, fruits, tools and clothes, the network could make an error by recalling any kitchen utensils, fish or random/stand-alone words.

Tables 5.5.1 and 5.5.1 indicate whether the correct categories were even mentioned during list recall. A correct category is one which is represented by the word list (spices, fruits, clothes and tools). Incorrect categories are those which were not on the list (e.g., fish and kitchen utensils).

Table 5.5.1 shows the average number of words recalled from each possible category in the network's vocabulary. All three models (normal, Lesioned Model A and Lesioned Model B) recalled more words from among the studied categories over words from non-list categories. If any words at all were recalled that

Normal							
Trial	Spices	Fruits	Tools	Clothes	Kitchen	Fish	Random
1	4.70	5.15	5.05	3.85	0.05	0.45	0.75
2	5.00	0.20	4.30	5.55	3.95	0.35	0.20
3	5.00	0.50	4.85	4.15	5.10	0.15	0.05
4	5.2	0	5.60	4.60	3.80	0.25	0.20
5	5.35	5.25	3.50	4.70	0.25	0.45	0.50
Frontal A							
Trial	Spices	Fruits	Tools	Clothes	Kitchen	Fish	Random
1	4.50	5.75	3.70	2.55	0.80	0.90	1.80
2	0.45	4.65	5.80	3.40	2.80	0.85	1.00
3	0.20	4.65	5.85	3.60	2.60	0.75	0.70
4	0.35	4.60	5.95	3.20	2.60	0.80	1.10
5	4.40	5.85	3.65	2.55	0.90	1.00	1.65
Frontal B							
Trial	Spices	Fruits	Tools	Clothes	Kitchen	Fish	Random
1	4.65	3.60	5.90	4.30	0.25	0.30	1.00
2	1.50	4.50	3.80	5.85	4.00	0.40	0.45
3	1.85	4.15	3.80	6.10	4.05	0.40	0.50
4	1.75	4.10	3.95	5.65	4.45	0.35	0.60
5	4.15	3.80	5.70	4.50	4.35	0.55	0.95

Table 5.7: *The average number of times a word from each possible category in the network's vocabulary were recalled. This table shows the results of a 'normal' network and two models of the frontally-lesioned networks over 5 trials. Data is averaged over 20 experiments.*

fell outside of the studied categories, they tended to be the randomly generated word patterns. Table 5.5.1 shows the percentage of times the network recalled at least a single word from a correct category. Lesioned Model A recalled a word from the correct categories less often than either the normal network and Lesioned Model B. The normal network and Lesioned Model B showed a near equivalent recall of words from correct categories.

Cluster Analysis

Four clustering indices were derived as described in the previous experiment.

Figure 5.12 shows the semantic clustering results of the normal and lesioned networks. Sheffé analysis showed no significant differences in the observed global clustering score between the normal network and Lesioned Model B, although both showed significant differences compared to Lesioned Model A. When the ratio of observed global clustering to expected chance clustering was used, the normal network showed significantly more global clustering tendencies than either Lesioned Model A or B, and Lesioned Model A showed significantly higher global clustering tendencies than Lesioned Model B ($F = 32.70$; $d.f. = 2, 57$; $p \leq .05$). When only correctly recalled observed cluster performance was examined, analysis revealed a difference between the three networks: the normal network showed higher raw clustering scores than Lesioned Model A ($p \leq .01$) and Lesioned Model B ($p \leq .05$). Lesioned Model B showed significantly less raw correct

Normal				
Trial	Spices	Fruits	Tools	Clothes
1	85	80	95	85
2	95	95	100	80
3	100	100	95	100
4	100	100	100	95
5	100	95	85	95
Frontal A				
Trial	Spices	Fruits	Tools	Clothes
1	50	65	70	40
2	55	70	80	70
3	50	60	75	60
4	50	70	75	55
5	45	75	75	55
Frontal B				
Trial	Spices	Fruits	Tools	Clothes
1	85	85	80	90
2	75	85	70	85
3	70	80	75	80
4	70	80	70	85
5	80	90	75	90

Table 5.8: *The percentage of times each list category was hit. This table shows the results of a 'normal' network and two models of the frontally-lesioned networks over 5 trials. Data is averaged over 20 experiments.*

clustering than Lesioned Model A ($p \leq .01$). These differences still held when the ratios of observed correct clustering to expected chance clustering were analyzed ($F = 58.90$; $d.f. = 2, 57$; $p \leq .01$).

5.5.2 Serial Recall

None of the networks showed subjective ordering serial recall above chance levels, but the network that simulated a normal subject showed higher raw scores of serial clustering and a linear trend in this direction. Given more trials, our normal network would show subjective ordering serial recall above chance levels.

5.5.3 Discussion of Results (Experiment 4)

Recall Performance

Examination of the types of errors exhibited by all three models showed that incorrect words recalled tend to come from the same categories that were represented in the original study list. Despite the addition of two more categories to the network's vocabulary, all three networks tended to recall words from the original four categories studied. Randomly generated patterns were recalled more often than extra-list categories,

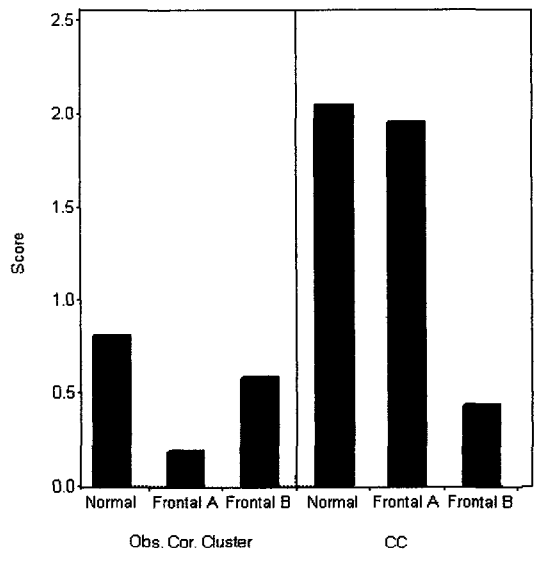

Figure 5.12: *The cluster scores of three networks: a 'normal' network, a 'lesioned' model with weights clamped to zero (Frontal A), and a 'lesioned' network with no memory for past actions and events (Frontal B). Obs. Cor. Cluster is the observed correct clustering score: the tendency to cluster correctly recalled words. CC is the ratio of observed correct clustering to expected clustering. It indicates clustering performance above chance levels. Data is averaged over 20 experiments.*

but later analysis showed that this was due to those patterns having a slightly higher correlation to the studied list patterns than did any of the non-list category patterns.

Although the probability of choosing a wrong 'spice' or a wrong 'fruit' (64%) was greater than the chance of choosing a wrong 'tool' or 'clothing' (43%), this did not affect either Lesioned Model B's or the normal model's probability of recalling a word from any of the correct 4 categories. However, it did affect Lesioned Model A's tendency to recall words from among the different correct categories. This suggested that without the ability to maintain a memory of past actions, selecting the correct Word Unit is still a random process despite narrowing the correct categories down to the appropriate set. Even being able to narrow the set of available categories to the correct ones was difficult for Lesioned Model A, as illustrated by the percentage of times that each correct category was hit over 20 experiments. The normal network nearly always recalled at least 1 word from all the correct categories, while frontal model B recalled a word (at least) from these categories roughly 75% of the time. However, Lesioned Model A only did so about 50-70% of of the time.

As expected, the normal network was capable of recalling more correct list words than either one of the lesioned models. Interestingly, Frontal Model B recalled more correct list words than Frontal Model A. This indicates that even a time-limited Hebbian type learning rule is better than no Cue Unit layer at all.

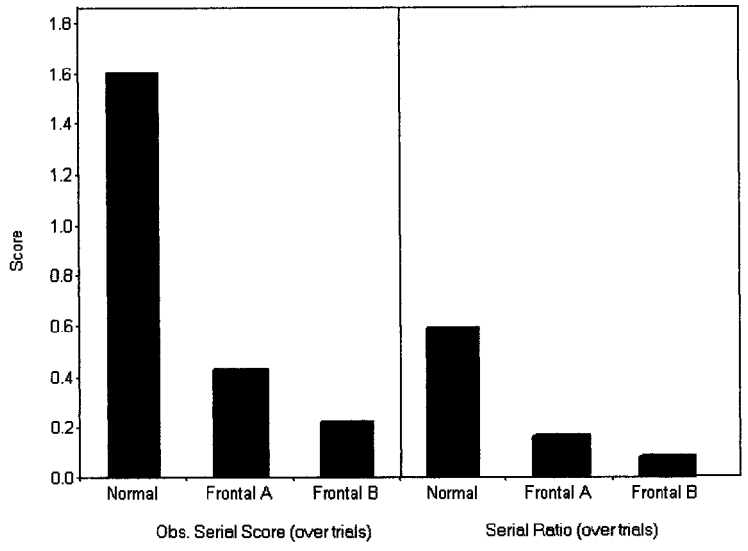

Figure 5.13: *The tendency of the network to recall words in the order in which it recalled words on the previous trial. Observed serial score is the number of times it recalls two words in succession that alsp appear in the same succession on the previous recall. Serial Ratio is the observed serial score over the expected serial clustering. Data is averaged over 20 experiments for 3 networks: a 'normal' network, a 'lesioned' model with weights clamped to zero (Frontal A), and a 'lesioned' network with no memory for past actions and events (Frontal B).*

Cluster Analysis

Normal human subjects typically use a semantic clustering strategy (consecutive recall of words from the same category) to recall the list. However, frontal lobe damaged patients fail to cluster semantically related words and typically show poor recall performance [73].

As expected, simulations of normal subjects showed a tendency to cluster words during recall over the 5 trials whereas the frontal-damaged models did not.

Serial Recall

One danger in repeatedly presenting a series of patterns in non-varying order is that the network may end up memorizing the exact order of the patterns in which they were presented. Despite this, the network which simulated normal subjects exhibited no tendency towards this type of serial recall.

Reinforcement learning was originally designed for agents to learn optimal state-action sequences. Therefore, it is expected that the network designed to simulate a normal subject would show a learning trend over the trials on serial recall based on previous recalls. Although it did not show serial recall above chance levels on the fifth trial, we would expect that it would do so given additional trials. This may be analogous to when a subject is asked to repeat a list over and over again. Eventually, subjects would be able to recite the

list in their own unique orders they created to aid in remembering the list. Low recall consistency reflects haphazard or disorganized styles of learning and may indicate that the subject has difficulty formulating or maintaining a learning plan [14]. Neither frontal lesion model showed this type of serial recall and showed no learning trend in this direction.

In short, the results showed that a 'normal' network showed higher recall consistency as compared to the lesioned networks. That is, words that were recalled on one trial were often recalled on subsequent trials. The 'frontal lesioned' networks did not show this. It has been noted that inconsistent recall occurs when a patient with limited learning capacity abandons one recall strategy for another [14]. Inconsistent recall is more prominent in patients with frontal-system pathology. This is attributed to an 'inability to retain the plan of memorizing' [14].

5.6 Further Discussion

We obtained a large improvement in performance by allowing the network to administer self-punishment when it detected a lack of appropriate responses with the present cue. This was detected by comparing the activity of one Word Unit over the activities of other Word Units and taking the softmaxed probabilities of activating each one during recall. This signalled the network when to switch to a new cue.

In the course of developing this model, other designs and learning algorithms not described here were tried and some interesting discoveries were made in the process. Two factors that determined whether the network perseverated were the level of 'blurred activity' in Frontal Cue Units and how greedy the network was allowed to become; that is, how often the network chose the most active Frontal Cue Units. In the current model, Frontal Cue Units are probabilisticaly chosen during the study phase based on their net input - the dot product between weights and patterns of incoming activity - from the Hippocampus. However, during recall Frontal Cue Units are chosen on the basis of the sustained past history of activity levels: this is information retained in the unit regardless of the pattern of activity currently displayed in the Hippocampus. In other words, one might consider this as a top-down/bottom-up process during the two phases. Study is a bottom-up process in which a pattern is stored in the Hippocampus and the 'best matching' Frontal Cue Unit is chosen to represent that pattern. The more often that particular Frontal Cue Unit is chosen, the higher its activity is due to feedback - a mechanism that allows the unit to retain a 'memory' of past activity. During recall, because the word units are left unclamped, the process is more top-down: there are several Frontal Cue units that compete to fire. When one fires, it sends patterns of activity to the Hippocampus, which leads to the recall of a word. If it is not punished, it will continue to fire and its activity level remains high.

However, upon being negatively rewarded, a strong punishment is given to it which weakens its probability of firing at the next time step. This allows other competing Frontal Cue Units to fire and send signals to the Hippocampus. Even though the temporarily weakened Frontal Cue Unit is eligible to fire again at the next time step, the new Frontal Cue Unit has already built its activity levels and will likely remain firing until punished.

What happens if a Frontal Cue Unit is chosen based on its blurred activity (the top-down process) during both the study and test phases? During the study phase, one unit will 'grab' all the information and increase its activity at the expense of the other units so that during recall it will be the only unit to fire.

The recurrent feedback link to a Frontal Cue Unit also affects the network's tendency to perseverate. The influence that a unit's past activity has on its own current activity is controlled by the parameter δ in Equation 2.15. As δ approaches zero, then its past activity is more salient than its current activity. When $\delta = 1$, past activity is ignored and only its current weighted net input from the Hippocampus is used. If its past activity is too high, then the Frontal Cue Unit will continue to fire even when it is inappropriate to do so.

Similarly, if the network is allowed to be very greedy and always chooses the most active Frontal Cue Unit, then that particular unit will become the only unit to fire throughout all the trials. These conditions will encourage the network to perseverate. Conversely, if the network is allowed to constantly 'explore' alternative Frontal Cues (have a very low greedy setting), then not only will the network cease to perseverate but it will not learn to cluster. Vying Frontal Cue Units will be allowed to fire at random and the overall observed effect would be a recall of words that 'jump' from one category to another.

Another interesting phenomenon is observed when the λ parameter is too high. The network does not perseverate in the sense that it recalls the word 'vest' repeatedly in immediate consecutive order [2]. However, the network would perseverate in a more long-term cyclic manner. For instance, on one trial it would recall words in the order of: 'parsley - *grapes* - *apricots* - *tangerines* - *plums* - *paprika* - *vest* - *sweater* - slacks - *grapes* - *apricots* - *tangerines* - *plums* - *paprika* - *vest* - *sweater*'. The words, categories and order of the cycles varied between 'subjects' and it did not correspond with the order in which the study list was presented. We attribute this cyclic ordering effect to the nature of the reinforcement learning algorithm itself. These learning rules were initially designed to be control-feedback mechanisms that allowed the system to choose the best *sequence of states and actions* that would lead to optimal performance and the highest reward. Our network was not blatantly rewarded for clustering words. It was only punished if it attempted to immediately repeat words or if there was a possibility that it would choose an improbable word (e.g., a

[2]This was due to our short-term memory suppression which prevented the six most recently recalled items

Frontal Cue Unit	State Chosen
2	parsley
1	*grapes*
1	*apricots*
1	*tangerines*
1	*plums*
1	*paprika*
1	*vest*
1	*sweater*
2	slacks
1	*grapes*
1	*apricots*
1	*tangerines*
1	*plums*
1	*paprika*
1	*vest*
1	*sweater*

Table 5.9: Example of network's output when λ is too high. Frontal Cue Units act 'greedily' but due to the nature of the reinforcement learning algorithm, a limited sequence of states and actions are learned.

non-list word). Thus, there was no reason for the network to not engage in this cyclic behaviour. Even if only one or two Frontal Cue Units greedily 'grabbed' all the activity (which was likely the case since λ was high) the network could still learn a limited set of states and actions (see Table 5.6). This type of cyclic perseverative behaviour may be commonly seen in frontal lobe damaged patients [80]. When initially started upon a sequence of actions or 'line of thought' it is difficult for the patient to switch to different sequences of actions or strategies. Cyclic behaviour is not without value: learning and performing repeated sequences of actions is important in a person's normal, daily life and in performing well on memory tasks [80]. Although our model of the hippocampal region was simple and did not have any sequencing abilities, researchers have suggested that the hippocampus is capable of at least simple sequential encoding and retrieval [17]. We would expect that by adding a temporal-sequencing mechanism into our hippocampal model, interesting, long-term, complex sequencing effects would be observed. With the prefrontal cortex intermittently cueing the hippocampal system with information appropriate for the task at hand, the hippocampus may produce its own set of sub-sequences in response to those cues. A 'hierarchy' of sequential actions and sub-sequences of actions would occur.

5.7 Experiment 5: AB-AC List Simulations

5.7.1 Simulation Methods

To simulate the AB-ACS and AB-ACD experiments, we created a set of 32 patterns. Like the patterns used in the previous experiments, these patterns were continuous-valued, vectors centered around a zero mean. Our vectors had 250 elements (features) and a sparseness of 0.10. Each of the 8 CS patterns had a 0.36 correlation with the corresponding B patterns. The correlation between pairs of words within a list was 0.10. Eight vectors of the same size and sparseness were randomly generated and composed the list of CD words. During study, input patterns to the network consisted of the normalized sum of corresponding A and B vectors or A and C vectors depending on the phase of the study.

Before studying either list, the network was pretrained on each pattern. That is, the Hippocampus' and the Word Unit-Hippocampal weights were trained on each of the 32 individual word patterns using the learning rules described in the network for the CVLT experiments.

During the List 1 Study phase, each normalized A-B combined pattern input was presented to the Hippocampus and the appropriate Localist Word Unit weights were updated after the Hippocampus settled. A 'best' Frontal Cue Unit was also probabilistically selected in a manner similar to the previous experiments. After each pass through all the A-B patterns, the network entered the recall phase. This process repeated for 10 trials.

At recall, the normalized sum of the recently most active Cue Unit and each 'A' cue pattern was presented as the input to the Hippocampus. From this cue, the network probabilistically selected the best Word Unit as its output.

The List 2 Study phase was similar to the List 1 Study phase except the inputs used were the summed and normalized A-CS or A-CD combined patterns. After each pass through all the patterns, the network entered a special recall phase. In this special recall phase, the network was required to recall the appropriate target word from both studied lists. The cues that were provided to the network to recall each target word from the second list were the normalized sum of each 'A' cue pattern, for all patterns in the second list, and the recently most active Cue Unit. The cues that were provided to the network to recall target words from the first list were the normalized sum of each 'A' cue pattern and the Cue Unit that had been most active during the previous recall phase.

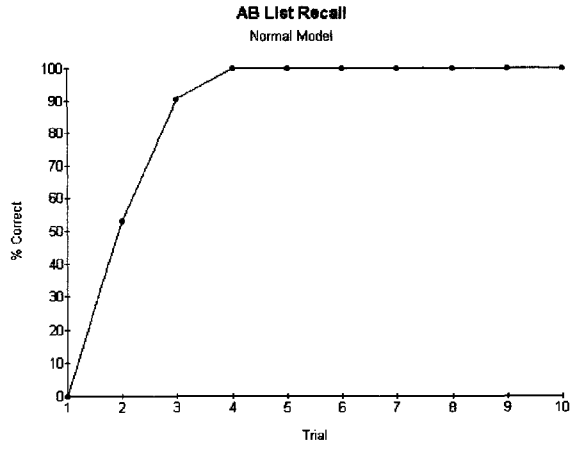

Figure 5.14: Recall performance of a normal subject simulation (A-B list).

5.7.2 Results

The results of our 'normal' model for the first and final recall tests are presented in Figures 5.14 - 5.16. Figure 5.14 shows the network's performance curve on learning List 1. With additional A-B learning trials, recall of the B responses improved from 53% (4.25 out of 8) after one A-B trial to 100% by the third trial. Figures 5.16 and 5.15 show the results of the A-CD and A-CS learning curves respectively. With increasing A-C learning trials, recall of the CD responses increased gradually, from .5 (6.25%) by the seventh trial to 6.25 (78%) by the tenth trial. Learning of the CS responses increased more slowly: from 1 (12.5%) by the ninth trial to 3 (37.5%) by the next trial. At the same time, recall of the B responses steadily declined, from 7.5 (94%) on the seventh trial on the A-CD list to 1.5 (19%) by the tenth trial. Similarly, recall of the B responses on the A-CS list declined, but more gradually: from 7.6 (95%) on the ninth trial to 4.5 (56%) on the last trial. Thus, learning of the C responses interfered with recall of the B responses, and this retroactive interference effect got larger with additional training on the A-C list.

The results of our 'frontal' model for the first and final recall tests are presented in Figures 5.17- 5.19. Figure 5.17 shows the learning and performance curve of List 1. As the number of A-B learning trials increased, recall of the B responses improved more slowly than that of the 'normal' network: from 2 (25%) after one A-B trial to 100% by the fourth trial. Figures 5.19 and 5.18 show the results of the A-CD and A-CS learning curves respectively. As the number of A-C learning trials increased, recall of the CD target words increased gradually, from .7 (8.75%) on the seventh trial to 5.25 (65.63%) by tenth trial. Learning of the CS responses also increased more slowly: from .5 (6.25%) on the sixth trial to 3.4 (42.5%) by the tenth

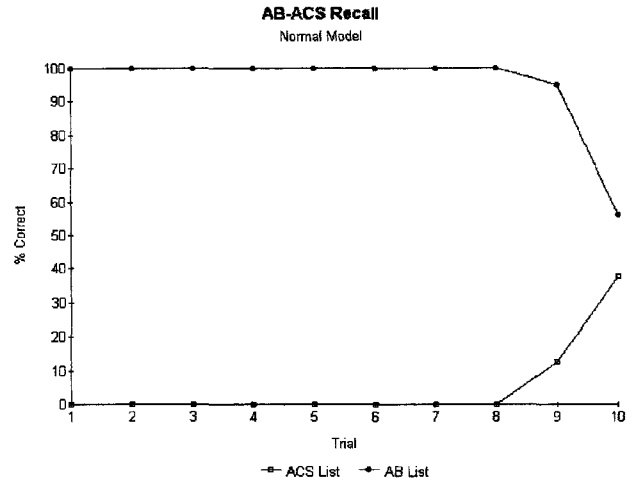

Figure 5.15: Recall performance of a normal subject simulation (A-CS list).

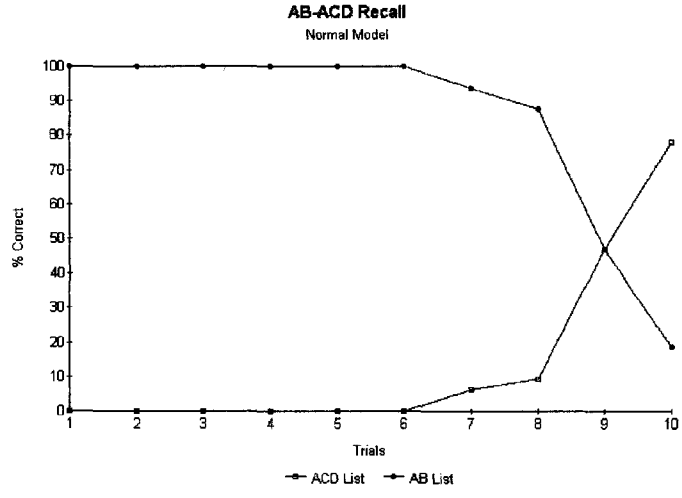

Figure 5.16: Recall performance of a normal subject simulation (A-CD list).

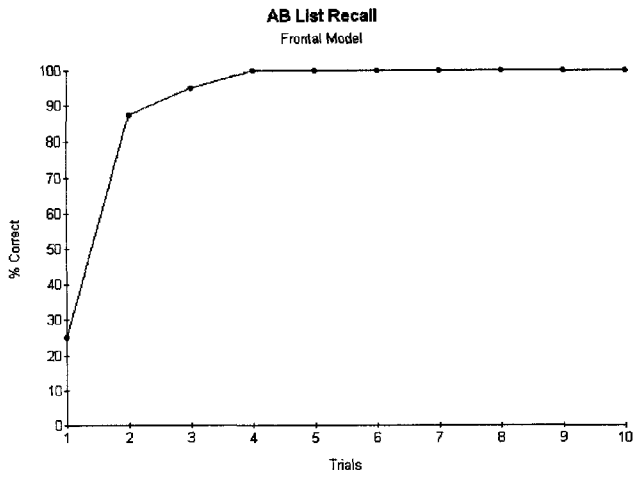

Figure 5.17: Recall performance of a frontal subject simulation (A-B list).

trial. At the same time, recall of the B responses steadily declined in these frontal models, from 7.7 (96.25%) on the seventh trial of the A-CD list to 1.5 (18.75%) by the tenth trial. Similarly, recall of the B responses on the A-CS list declined more gradually: from 7.5 (93.75%) on the seventh trial to 3.75 (46.87%) on the tenth trial. Again, learning of the C responses interfered with recall of the B responses, and this retroactive interference effect got larger with more training on the A-C list.

5.7.3 Discussion

Both network models trained on the A-B list could readily learn all 8 associations. The 'normal' network learned the A-B list to perfection only slightly faster (by 1 trial) than the 'frontal' model. On average it took 3 trials for the network to learn the A-B list completely. After completion of the A-B training, the network was trained either on the A-CS or the A-CD list. For both 'normal' and 'frontal' models, it took more trials to learn the A-CS list than the A-CD list. This was to be expected: experimental evidence on human subjects had suggested that more List 1 intrusions would occur in the learning of List 2 for the A-CS paradigm than for the A-CD paradigm and this was supported by our network's performance.

Comparing our results (Figures 5.14 - 5.16) to the results obtained by Barnes and Underwood (Figures 3.3 - 3.4) shows performance of our model is similar but not accurate. Our model learns the second list more slowly than a human subject, but this problem could be solved by increasing its learning rate parameter. A second difference is that by the last trial, our network has completely forgotten the first association list in order to learn the second one. Normal human subjects, however, manage to retain some memory of the second list even after learning the first one. It turns out that the solution to our second problem isn't so

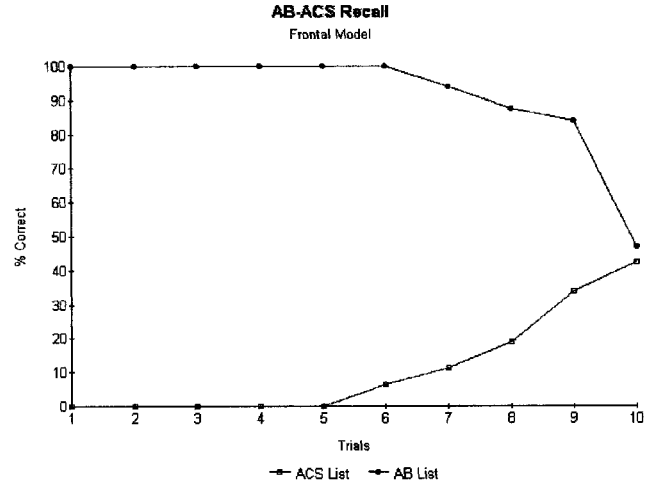

Figure 5.18: Recall performance of a frontal subject simulation (AB-ACS list).

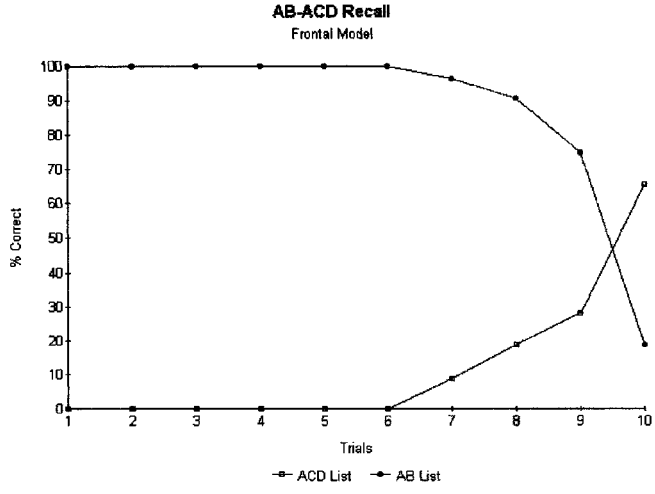

Figure 5.19: Recall performance of a frontal subject simulation (AB-ACD list).

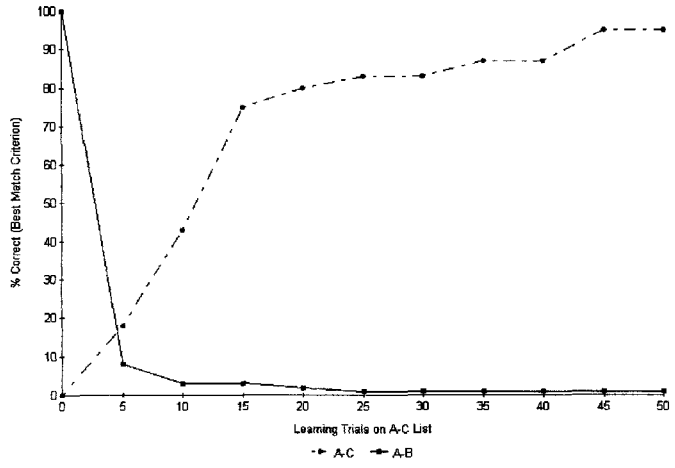

Figure 5.20: McCloskey & Cohen (1989) simulation of a normal subject on the AB-AC test.

simple.

Previous studies have shown that the difficulties neural network models have in performing the AB-AC list learning task lie in the parallel distributed nature of the networks themselves. McCloskey and Cohen (1989) engendered a flurry of research activity among connectionist modelers when they first documented evidence of a distributed memory model's inability to perform the AB-AC list learning test. Using a multi-layered network with a backpropagation learning algorithm to simulate Barnes and Underwood's study, they found that their model exhibited 'catastrophic interference'. This is the complete loss of memory for previously learned items at the expense of learning new information. Comparison of Figure 5.20, which shows the catastrophic phenomenon exhibited by McCloskey and Cohen's model, to Figure 3.3, which is the data obtained by Barnes and Underwood underlines the differences between human and computational model performance. Whereas humans show a gradual loss of ability to retrieve the AB list and are still capable of operating more than 50% correct after the AC list performance has reached asymptote, the network shows virtually complete abolition of AB list performance before AC performance rises above 0% correct.

Although our model does not show severe catastrophic interference per se (it doesn't forget the first list after only the first trial) it does show fairly rapid unlearning. The reason for this is that unlike localized models, all information in the memory of distributed memory models is represented by several features or elements. Thus, a new stimulus is represented by a change in the state of the entire memory system and not by one separate element. One can sidestep the catastrophic interference phenomenon in distributed neural network models by training them concurrently - repeatedly presenting a single set of training items that includes *all* of the items to be learned, or at least incorporates all of the regularities (similarities between

vectors due to common features) to be captured by the network. McClelland et. al. (1995) have referred to this as 'interleaved learning', and suggest that this is how the cortex avoids the problem of catastrophic interference. But not all human learning tasks are concurrent. Humans are more often faced with learning situations that are sequential in nature. However, in computational modeling, disruption of old knowledge by new learning is common when training is sequential, as it is in the AB-AC list learning task. This is because in a connectionist model with distributed representations, each connection weight is involved in responding to many different inputs. Thus, adjustment of weights to encode the desired response to a new input pattern will necessarily alter the network's response to other inputs as well. In many respects this is a desirable feature. It is the basis for 'automatic generalization' - a network, through training on some patterns, comes to respond appropriately to other (untrained) patterns. The disadvantage is that changing weights to encode a new piece of information may alter previously learned responses to other input patterns. Weight adjustments during learning of the second list alters the previously learned responses to the first list.

We made several modifications to the network presented so far in an effort to alleviate this interference:

1. The number of units in the hippocampal module was increased (hence, increasing the word pattern size). This was found to have no effect on the unlearning of the first list.

2. We increased the sparseness levels of the A, B, CS, and CD word patterns. However, this too had no effect in reducing interference. This was because the studied list words still remained correlated (the summed A-B patterns were still correlated with the A-C patterns due to the identical A pattern used in both lists).

3. Direct connections from the Frontal Cue layer to the Word Unit layer were established. This, had no effect in reducing interference. Again, this was because of the distributed sharing of weights between the Hippocampus and Word Units.

In short, none of these modifications had a significant effect on reducing the gradual unlearning of the weights to the first list.

Much research has been done on this problem and other modelers have discovered how some computational features can reduce interference [44, 20, 40, 76]. The presence of an intermediate layer of hidden units that enables the network to build a distinct internal representation of each pattern will make the network less susceptible to interference. In other words, even if the input vectors are not orthogonal (as is the case in our tasks), the two patterns at the hidden layer can be made to be orthogonal, thus interference (at least concerning the updating of the second layer of weights between hidden and output units) is minimized.

There are several methods of achieving this orthogonality at the hidden layer. Kortge (1990) proposed a 'novelty-driven' rule that focuses only on the novel aspects of a to-be-learned pattern. Instead of learning each item 'from scratch', only those components that differ from previously acquired information are learned. Thus, only those weights that have not been previously dedicated to the storage of information are updated. This leads to dissimilar internal representations even in the presence of correlated input.

French (1991) proposed an 'activation sharpening' method. This technique introduces dependence among hidden units by some proportion α while decreasing the activation, a, of all other units such that $a_{new} = a_{old} + \alpha(a - a_{old})$ for the most active units, and $a_{new} = a_{old} - \alpha a_{old}$ for all others. These algorithms lead to the emergence of internal representations that have only a few highly activated units at the hidden layer. Because the identity of these 'sharpened' units likely differs across stimuli, interference is reduced.

Sloman and Rumelhart (1992) proposed a set of units that may be referred to as 'episodic' or 'context' units. Between these context units and the output units was a layer of hidden units. These hidden units acted as gates between what would be the equivalent of our hippocampal - Word Unit weights and some context-Word Unit weights. A hidden unit would turn on ***only*** if a context unit ***and*** its corresponding hippocampal unit were on. Otherwise, the Word Unit would not be activated. By employing these gatekeepers, the system can 'compartmentalize' its knowledge so that any interference between list items occurs mostly within a context.

There may be problems with applying any of these solutions to our particular model. The activation sharpening and novelty rules have been previously applied only to the backpropagation learning algorithm. How these methods can be applied to a multi-layer reinforcement learning algorithm requires further analysis. In Sloman and Rumelhart's model, a direct mapping from input to output units was used. Like an autoencoder, one input unit had exactly one corresponding output unit. However, in our model, we have a distributed vector of units in our hippocampal layer that needs to be mapped to one localist word unit.

One potential solution to our dilemma may be solved by adding more layers to our hippocampal module. With additional layers, the hippocampus may enforce increasing sparsity of the combined A-B and A-C pattern vectors, thereby de-correlating them. The idea that the hippoocampal region is a multi-layered structure with auto-encoding properties of sparse input vectors has been proposed in the past [5] and is suggested by anatomical studies. This possible solution is yet to be tested on our model.

Up to now, we have used a highly simplified hippocampal model in order to focus on the formation of memory cues in the prefrontal cortex and on prefrontal-hippocampal interactions. Whatever method we might try to use to solve this gradual unlearning phenomenon in our network's performance, one thing is clear. The addition of an extra, hidden layer between the hippocampal module and the Word Unit layer

is necessary. Direct connections from the Frontal Unit layer to the Word Unit layer may also be required. In any case, the simplistic architecture of only 3 layers that we describe in this paper is not sufficient to perform the AB-AC task without complete forgetting of previously learned items.

Results from the CVLT simulations suggest that our learning algorithm and the interaction between our hippocampal and frontal modules is enough to from contextual-type cues and clustering to aid in recall. Results from our AB-AC list learning test, however, show that this is not a sufficiently complete model of the prefrontal cortex and hippocampus.

Although anatomical studies have shown evidence of direct connections between the frontal lobes and the hipocampal regions [60, 22, 42], these simulations suggest that a single recurrent layer to represent the hippocampus and a single recurrent layer to model the prefrontal cortex is not enough. Additional hidden layers and direct communication between the prefrontal layer (providing retrieval cues) to other regions is necessary.

Chapter 6

Conclusion

Even though it is believed that multiple brain systems are involved in memory, it may not necessarily follow that multiple regions of the prefrontal cortex are needed to exhibit the functions attributed to it. By using a reinforcing learning algorithm in a connectionist modeling framework, we were able to create a network capable of a wide variety of behaviours: non-perseveration, clustering (chunking) of information, and the organized retrieval of information. Our prefrontal cortex modulated retrieval of memories by:

- Extracting relevant, general features of the items stored in the hippocampus and integrating this information across time.
- Providing constraining cues (the general features of items) to the hippocampal region so that this system could recall specific instances.
- Imposing shifting cues to the hippocampus which prevented perseveration.

Is there a biological basis for our self-reinforcing model? Neurophysiological studies have shown that the prefrontal cortex receives projections from the hypothalamus, the amygdala and several brainstem structures, directly or indirectly through the thalamus [22]. These limbic structures are believed to be involved in the control of behaviour (reviewed in [50]) and can inform the prefontal cortex about drive and reinforcement. Fuster (1988) has also emphasized the importance of a 'drive' signal for the acquisition and execution of goal-directed behaviours. The prefrontal cortex appears to need such a reinforcement signal, which may be available from these limbic structures, to search actively for new, more suitable behaviours.

As noted in the introduction, one of the uses of computational modeling is to examine a theory step by step and avoid the ambiguity of verbal theorizing. In implementing this model, several realizations were made:

1. A model of inhibiting recently recalled word items is more complicated than initially thought.

   In our model, a 'short term memory' system stored the six most recently recalled word items and prevented words from being recalled repeatedly. When a word was output, it was compared to the words kept in this short-term memory store. If it matched any of those words then a negative reward signal was given to the network. In short, the network was punished when it attempted to recall the same word it had already recalled recently. One could view this as a 'filtering' mechanism: a subject's natural inclination may be to recall a recently retrieved word, but conscious and deliberate comparison of this word to the information kept in this short term memory buffer would cause the subject to try to recall a different word.

   An alternative, more lower-level, mechanism to this short term memory suppression would be to temporarily raise the firing threshold of a word unit immediately after it had been activated. This would also prevent a word from being repetitively recalled unless an extremely strong input signal could overcome its activation threshold.

   Yet another alternative would be a compromise between the above two. A short term memory store could exist which has inhibitory connections from each short term memory unit to all the word units. When a word is recalled, the short term memory units inhibit that word with a strength that decreases over time. This may be analogous to a thalamic system implementing an 'inhibition of return', as it does in spatial attention.

   Regardless of which method is used, the end result would be the same: immediate repetitive recall of words is inhibited. These are simple 'fixes', however, and simplistic models of another possibly complex system. A more comprehensive account of such a system would have to take into account the length of time items would be 'kept' in storage, the possible effect of newly stored items on older ones, the mechanisms by which such information can be represented, the transfer of items into longer-term storage, its possible involvement in priming, and its relation to 'working memory' - to name just a few. To model such a system is in and of itself a large task. To keep the complexity of this model in check, our short-term memory system of the brain was simply implemented as an inhibitory mechanism that is capable of suppressing the six most recently recalled word items.

   Even though our network cannot perform our word recall tasks without this inhibition, we are cautious about attributing this mechanism to the prefrontal cortex. It has been suggested that the prefrontal cortex has inhibitory functions. When posterior regions of the frontal lobe (near the sensory-motor strip) are damaged, patients show motor perseveration. When the prefrontal cortex is damaged they

show long-term perseverative behaviour to previously rewarded responses (e.g., as in the Wisconsin Card Sorting task). However, during free recall experiments, they do not cite words in a monotonously repetitive manner. Thus although the frontal cortex may possess inhibitory functions, we can not claim that the particular short term memory suppression used in our model is a component of the prefrontal cortex.

2. The mechanism by which our prefrontal module shifted its cues to the hippocampal region could be further explored.

   Our network detected when a particular cue was exhausted by comparing the difference in activity levels between the 'winning' word unit and the 'runner-up' word unit. A large disparity signaled our prefrontal module to switch to a new cue. However, the biological plausibility of this mechanism is questionable. An alternative, more plausible mechanism would be to use the hippocampus' recognition capabilities. When the prefrontal cortex presents a cue to the hippocampus and a specific word unit is about to be output, the activity that the word generates in the Hopfield network of the hippocampus is first tested. If the calculated energy level is below threshold (indicating a stored attractor state) then the word is output and the prefrontal module keeps using the same cue. However, if the calculated energy level is above threshold then it is a signal for the prefrontal module to switch cues. This mechanism remains to be tested in future work.

Several other avenues of future experiments are open to the modeler with this network. In addition to experimenting with additional layers to reduce catastrophic interference, it would be interesting to create a hippocampal module that is independently capable of temporal sequencing. Complex sequencing behaviours may be observed when it is combined with our prefrontal module's reinforcement-based sequencing behaviour.

Moscovitch (1994) has posited that the prefrontal cortex provides and controls information going to and coming from the hippocampus. However, the prefrontal module of our model merely picks up information encoded by the hippocampus and, in turn, provides it with retrieval cues. In our current model, the prefrontal module neither shapes nor controls environmental information given to the hippocampus. Perhaps an additional component, in the form of attention, needs to be added to make this model more complete.

Based on our initial simulations, we suggest that some experimental predictions may be made. In the California Verbal Learning task, subjects are given five trials where recall immediately follows each study phase. Because our network was allowed to learn not only during the study phase but during the recall phase as well, it was able to create a subjective order during recall of the list independent of the actual word sequence of the list. If, instead, our network were to be presented with five study trials before being able to

recall the list, it is predicted that the order of words recalled would closely resemble the order in which the words were presented and less semantic clustering of words would be observed. This prediction could also be tested in human subjects.

The cerebral cortex possesses five to seven layers of neurons which form inter- and intra-layer connections with other neurons, or recurrently with itself. The hippocampus proper is composed of several distinct fields. This makes the hippocampal region that is believed to be responsible for encoding episodic memories a system of fields connected to the rest of the brain via several layers of neocortex. Obviously, our simple model does not do the complexity of the brain justice and it simulates just a fraction of the possible mechanisms the brain may employ to encode and retrieve memories. However, we had hoped to capture the essence of some key features of the interaction between the hippocampal and prefrontal regions. Although it captured much of what we hoped it would, our model's behaviour on the AB-AC list learning test showed that our network is not yet a complete model of this interaction. Additional layers or sets of 'brain regions' may be needed, or we may find that different learning rules altogether may be needed for a fully functioning model.

This is the first attempt that we know of to simulate in a distributed, computational framework the complex processes underlying the hippocampal and prefrontal interactions that yield strategic organization and retrieval of memories. In analyzing our model's successes as well as its failures, we've been able to take initial steps in refining and revising our ideas on how these processes work.

Bibliography

[1] C. Anderson. Q-learning with hidden-unit restarting. *Advances in Neural Information Processing Systems*, 5:81–88, 1993.

[2] R. Bapi and D. Levine. Modeling the role of frontal lobes in sequential task performance: Basic structure and primacy effects. *Neural Networks*, 7(6/7):1167–1180, 1994.

[3] J. Barnes and B. Underwood. Fate of first-list associations in transfer theory. *Journal of Experimental Psychology*, 58(2):97–104, 1959.

[4] A. Barto, S. Bradtke, and S. Singh. Real-time learning and control using asynchronous dynamic programming. Technical Report 91-57, Department of Computer Science, University of Massachusetts, Amherst, MA, Aug. 1991.

[5] S. Becker. Modelling episodic memory: A global cost function that leads to fast, local, high-capacity learning. *Proceedings of the "Learning" workshop, Snowbird, Utah. Abstract*, (April), 1999.

[6] A. Benton. The prefrontal region: Its early history. In H. Levin, H. Eisenberg, and A. Benton, editors, *Frontal Lobe Function and Dysfunction*, pages 3–33. Oxford University Press, Oxford, 1991.

[7] W. Bousfield. The occurrence of clustering in the recall of randomly arranged associates. *Journal of General Psychology*, 49:229–240, 1953.

[8] P. Burgess, E. Veitch, A. Costello, and T. Shallice. The cognitive and neuroanatomical corelates of multitasking. Submitted.

[9] M. Chappel and M. Humphreys. An auto-associative neural network for sparse representations: Analysis and application to models of recognition and cued recall. *Psychological Review*, 101:103–128, 1994.

[10] J. Cohen and D. Servan-Schreiber. Context, cortex and dopamine: A connectionist approach to behavior and biology in schizophrenia. *Psychological Review*, 99(1):45–77, 1992.

[11] J. Cohen and D. Servan-Schreiber. A theory of dopamine function and its role in cognitive deficits in schizophrenia. *Schizophrenia Bulletin*, 19(1):85–104, 1993.

[12] F. Craik, M. Naveh-Benjamin, R. Govoni, and N. Anderson. The effects of divided attention on encoding and retrieval processes in human memory. *Journal of Experimental Psychology: General*, 125(2):159–180, 1996.

[13] S. Dehaene and J. Changeux. A hierarchical neuronal network for planning behavior. *Proc. Natl, Acad. Sci. USA*, 94(Nov):13293–13298, 1997.

[14] D. Delis, J. Kramer, E. Kaplan, and B. Ober. *The California Verbal Learning Test.* San Antonio, TX: Psychological Corporation, 1987.

[15] A. DellaRocchetta. Classification and recall of pictures after unilateral frontal or temporal lobectomy. *Cortex*, 22:189–211, 1986.

[16] M. Diamond, A. Scheibel, and L. Elson. *The Human Brain Coloring Book.* Harper Perennial, New York, 1985.

[17] J. Dusek and H. Eichenbaum. The hippocampus and memory for orderly stimulus relations. *Procedures of National Academy of Sciences*, pages 7109–7114, 1994.

[18] H. Ebbinghaus. *Memory: A contribution to experimental psychology.* Columbia University, Teachers College, 1913.

[19] L. Fox, J. Olin, F. Erblich, C. Ippen, and L. Schneider. Severity of cognitive impairment in alzheimer's disease affects list learning using the california verbal learning test (cvlt). *Int. J. Geriat. Psychiatry*, 13:544–549, 1997.

[20] R. French. Using semi-distributed representations to overcome catastrophic forgetting in connectionist networks. In *Proceedings of the 13th annual converence of the cognitive science society*, pages 173–178, Hillsdale, NJ. Lawrence Erlbaum Associates.

[21] C. Frith. *Control of Cognitive Processes: Attention and Performance XVIII.* MIT Press, Cambridge, MA, 2000.

[22] J. Fuster. *The Prefrontal Cortex: Anatomy, Physiology and Neuropsychology of the Frontal Lobe.* Raven Press, NY, 1988.

[23] F. Gershberg and A. Shimamura. The role of the frontal lobes in free recall: Interference, organization, and serial position effects. *Society for Neuroscience Abstracts*, 17, 1991.

[24] M. Gluck and C. Myers. Hippocampal mediation of stimulus representation: A computational theory. *Hippocampus*, 1993.

[25] A. Goldman-Rakic and H. Friendman. The circuitry of working memory revealed by anatomy and metabolic imaging. In H. Levin, H. Eisenberg, and A. Benton, editors, *Frontal Lobe Function and Dysfunction*, pages 72–92. Oxford University Press, Oxford, 1991.

[26] M. Griniasty, M. Tsodyks, and D. Amit. Conversion of temporal correlations between stimuli to spatial correlations between attractors. *Neural Computation*, 5:1–17, 1993.

[27] E. Guignon, B. Dorizzi, Y. Burnod, and W. Schultz. Neural correlates of learning in the prefrontal cortex of the monkey: A predictive model. *Cerebral Cortex*, Mar/Apr(2):135–147, 1995.

[28] M. Hasselmo and B. Wyble. Free recall and recognition in a network model of the hippocampus: simulating effects of scopolamine on human memory function. *Behavioural Brain Research*, 89:1–34, 1997.

[29] D. Hintzman. *Relating Theory and Data: Essays on Human Memory in Honor of Bennet B. Murdock*, chapter Why are formal models useful in psychology?, pages 39–56. Lawrence Erlbaum Associates, Hillsdale, NJ, 1991.

[30] J. Hopfield. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences*, 1982.

[31] D. Hultsch. Adult age differences in retrieval: Trace-dependent and cue-dependent forgetting. *Developmental Psychology*, 11:197–201, 1975.

[32] R. Jaffard and M. Meunier. Role of the hippocampal formation in learning and memory. *Hippocampus*, 3:203–218, 1993.

[33] J.S. Janowsky, A.P. Shimamura, M. Kritchevsky, and L. Squire. Cognitive impairment following frontal lobe damage and its relevance to human amnesia. *Behavioral Neuroscience*, 103:548–560, 1989.

[34] W. Jetter, U. Poser, R. Freeman, and H. Markowitsch. A verbal long term memory deficit in frontal lobe damaged patients. *Cortex*, (22):229–242, 1986.

[35] D. Kausler. *Learning and Memory in Normal Aging*, chapter 7, pages 229–235. Academic Press, San Diego, 1994.

[36] M. Kibby, M. Schmitter-Edgecombe, and C. Long. Ecological validity of neuropsychological tests: Focus on the california verbal learning test and the wisconsin card sorting test. *Archives of Clinical Neuropsychology*, 13(6):523–534, 1998.

[37] D. Kimberg, M. D'Esposito, and M. Farah. Cognitive functions in the prefrontal cortex- working memory and executive control. *Current Directions in Psychological Science*, 6(6):185–192, 1998.

[38] D. Kimberg and M. Farah. A unified account of cognitive impairments following frontal lobe damage: the role of working memory in complex organized behavior. *Journal of Experimental Psychology: General*, 122(4):411–428, 1998.

[39] K. Kimberg, M. D'Esposito, and M. Farah. Cognitive functions in the prefrontal cortex - working memory and executive control. *American Psychologist Society*, 6(6):185–192, 1997.

[40] C. Kortge. Episodic memory in connectionist networks. In *Proceedings of the 12th annual converence of the cognitive science society*, pages 764–771, Hillsdale, NJ. Lawrence Erlbaum Associates.

[41] J. Kramer and D. Delis. Interference effects on the california verbal learning test: A construct validation study. *Psychological Asessment*, 3(2):299–302, 1991.

[42] L. Laitinen and K. Livingston, editors. *Surgical Approaches in Psychiatry*, chapter 39, pages 303–314. Medical and Technical Publishing, Great Britain, 1973.

[43] D. Levine and P. Prueitt. Modeling some effects of frontal lobe damage - novelty and perseveration. *Neural Networks*, 2:103–116, 1989.

[44] S. Lewandowsky and S. Li. *Interferece and Inhibition in Cognition*, pages 330–359. Academic Press, New York, 1995.

[45] L. Lin. Self-improving reactive agents based on reinforcement learning, planning and teaching. *Machine Learning*, 8:293–321, 1992.

[46] A. Luria. *The Neuropsychology of Memory*. V.H.Winston, Washington, DC, 1902.

[47] A. Mayes, P. Meudell, and A. Pickering. Is organic amnesia caused by a selective deficit in remembering contextual information? *Cortex*, 21:167–202, 1985.

[48] J. McClelland, B. McNaughton, and R. O'Reilly. 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(3):419–457, 1995.

[49] M. McCloskey and M. Zaragoza. Misleading postevent information and memory for events: Arguments and evidence against memory impairment hypotheses. *Journal of Experimental Psychology: General*, 114:3–18, 1985.

[50] D. McGinty and R. Szymusiak. Neuronal activity patterns in behaving animals: brainstem and limbic system. *Annual Review in Psychology*, 39:135–168, 1988.

[51] B. McNaughton and R. Morris. Hippocampal synaptic enhancement and information storage within a distributed memory system. *Trends in Neurosciences*, 10(10):408–415, 1987.

[52] B. Milner, S. Corkin, and H. Teuber. Further analysis of the hippocampal amnesic syndrome: 14-year followup study of H.M. *Neuropsychologia*, 6:215–234, 1968.

[53] B. Milner, P. Corsi, and G. Leonard. Frontal-lobe contribution to recency judgments. *Neuropsychologia*, 29(6):601–618, 1991.

[54] S. Moody, S. Wise, G. Pellegrino, and D. Zipser. A model that accounts for activity in primate frontal cortex during a delayed matching to sample task. *Journal or Neuroscience*, 18(1):399–410, 1997.

[55] M. Moscovitch. Cognitive resources and dual-task interference effects at retrieval in normal people: the role of the frontal lobes and medial temporal cortex. *Neuropsychology*, 8(4):524–534, 1994.

[56] M. Moscovitch and C. Umilta. Conscious and nonconscious aspects of memory: A neuropsychological framework of modules and central systems. In H.J. Weingartner and R.G. Lister, editors, *Perspectives in Cognitive Neuroscience*. Oxford University Press, Oxford, 1991.

[57] M. Moscovitch and C. Umilta. Modularity and neuropsychology: Implications for the organization of attention and memory in normal and brain-damaged people. In M.E. Schwartz, editor, *Modular Processes in Dementia*. MIT/Bradford, Cambridge, MA, 1991.

[58] M. Moscovitch, E. Vriezen, and J. Gottstein. *Handbook of Neuropsychology*, volume 8, chapter Implicit tests of memory in patients with focal lesions or degenerative brain disorders. Elsevier, 2000.

[59] L. Nadel and M. Moscovitch. Memory consolidation retrograde amnesia and the hippocampal complex. *Current Opinion in Neurobiology*, 7(2):217–227, 1997.

[60] W. Nauta. The problem of the frontal lobe: A reinterpretation. *Journal of Psychiatric Research*, 8:167–187, 1971.

[61] M. Petrides and B. Milner. Deficits on subject-ordered tasks after frontal- and temporal lobe lesions in man. *Neuropsychologia*, 20:249–262, 1982.

[62] A. Robin and D. Macdonald. *Lessons of Leucotomy*, chapter The Anatomy of the Frontal Lobe. Henry Kimpton Publishers, London, 1975.

[63] E. Rolls. A theory of hippocampal function in memory. *Hippocampus*, 6:601–620, 1996.

[64] M. Rugg and K. Allan. Memory retrieval: An electrophysiological perspective. In M. Gazzaniga, editor, *The Cognitive Neurosciences*, pages 755–765. MIT Press, London.

[65] G. Rummery and M. Niranjan. On-line q-learning using connectionist systems. Technical report, Cambridge University Engineering Dept, Cambridge, England, September 1994.

[66] D. Salmon, S. Zola-Morgan, and L. Squire. Retrograde amnesia following combined hippocampus-amygdala lesions in monkeys. *Psychobiology*, 15(1):37–47, 1987.

[67] D. Schacter. *Memory systems of the brain: Animal and human cognitive processes*, chapter Multiple Forms of Memory in Humans and Animals, pages 351–379. Guildford Press, NY, 1985.

[68] D. Schacter. Memory, amnesia and frontal lobe dysfunction. *Psychobiology*, 15:21–36, 1987.

[69] D. Schacter and M. Moscovitch. Infants, amnesics and dissociable memory systems. *Infant Memory (Advances in the Study of Communication and Effect)*, 9:173–217, 1984.

[70] T. Shallice. *The Neuropsychology of Cognitive Function*, chapter Specific impairments of planning, pages 199–209. The Royal Society, London, 1982.

[71] B. Shen and B. McNaughton. Modeling the spontaneous reactivation of experience-specific hippocampal cell assemblies during sleep. *Hippocampus*, 6:685–692, 1996.

[72] R. Shiffrin and J. Cook. Short-term forgetting of item and order information. *Journal of Verbal Learning and Verbal Behavior*, 17:189–218, 1978.

[73] A. Shimamura. Memory and frontal lobe function. In M. Gazzaniga, editor, *The Cognitive Neurosciences*, pages 803–814. MIT Press, London, 1994.

[74] A. Shimamura, P. Jurica, J. Mangels, F. Gershberg, and R. Knight. Susceptibility to memory interference effects following frontal lobe damage: Findings from tests of paired-associate learning. *Journal of Cognitive Neuroscience*, 7(2):144–152, 1995.

[75] A.P. Shimamura, J.S. Janowsky, and L. Squire. Memory for the temporal order of events in patients with frontal lobe lesions and amnesic patients. *Neuropsychologia*, 28:803–813, 1990.

[76] S. Sloman and D. Rumelhart. *From Learning Theory to Connectionist Theory*, chapter Reducing Interference in Distributed Memories through Episodic Gating. Lawrence Erlbaum Assoc, Hillsdale, NJ, 1992.

[77] L. Squire. Memory and hippocampus: a synthesis from findings with rats, monkeys and humans. *Psychological Review*, 99(2):195–231, 1992.

[78] L. Squire and B. Knowlton. Memory, hippocampus and brain systems. In M. Gazzaniga, editor, *The Cognitive Neurosciences*, pages 825–837. MIT Press, London, 1994.

[79] D. Stuss, M. Alexander, C. Palumbo, L. Buckle, L. Sayer, and J. Pogue. Organizational strategies of patients with unilateral or bilateral frontal lobe injury in word list learning tasks. *Neuropsychology*, 8(3):355–373, 1994.

[80] Donald Stuss and Frank Benson. *The Frontal Lobes*. Raven Press, New York, 1986.

[81] R. Sutherland and J. Rudy. Configural association theory: The role of the hippocampal formation in learning, memory and amnesia. *Psychobiology*, 17(2):129–144, 1989.

[82] R. Sutton. Learning to predict by the methods of temporal differences. *Machine Learning*, 3:9–44, 1988.

[83] A. Treves and E. Rolls. Computational constraints suggest the need for two distinct input systems to the hippocampal ca3 network. *Hippocampus*, 2(2):189–200, 1992.

[84] A. Treves, W. Skaggs, and C. Barnes. How much of the hippocampus can be explained by functional constraints? *Hippocampus*, 6:666–674, 1996.

[85] M. Tsodyks. Associative memory in asymmetric diluted network with low level of activity. *Europhysics Letters*, 7(3):203–208, 1988.

[86] M. Tsodyks and M. Feigelman. The enhanced storage capacity in neural networks with low activity level. *Europhysics Letters*, 6(2):101–105, 1988.

[87] E. Tulving. Subjective organization in free recall of 'unrelated' words. *Psychological Review*, 69(4):344–354, 1962.

[88] E. Tulving. *Elements of Episodic Memory*, chapter Inchoate Distinction, pages 17–31. Oxford University Press, New York, 1983.

[89] E. Tulving, D. Schacter, and H. Stark. Priming effects in word-fragment completion are independent of recognition memory. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 8:336–342, 1982.

[90] F. Vargha-Khadem, D. Gadian, K. Watkins, A. Connelly, W. Paesschen, and M. Mishkin. Differential effects of early hippocampal pathology on episodic and semantic memory. *Science*, 277:376–380, July 1997.

[91] C. Watkins. *Learning from Delayed Rewards*. PhD thesis, King's College, Cambridge University, UK, 1989.

[92] D. Weinberger, K. Berman, and D. Daniel. *Frontal Lobe Function and Dysfunction*, chapter Prefrontal Cortex Dysfunction in Schizophrenia, pages 275–288. Oxford University Press, Oxford, 1991.

[93] I. Wickelgren. Getting a grasp on working memory. *Science*, 275(Mar 14):1580–1582, 1997.

[94] D. Willshaw and J. Buckingham. An assessment of marr's theory of the hippocampus as a temporary memory store. *Phil. Trans. R. Soc. Lond.*, 329:205–215, 1990.

[95] G. Winocur. Functional dissociation of the hippocampus and prefrontal cortex in learning and memory. *Psychobiology*, 19(1):11–20, 1991.

[96] G. Winocur. A comparison of normal old rats and young adult rats with lesions to the hippocampus or prefrontal cortex on a test of matching-to-sample. *Neuropsychologia*, 30(9):769–781, 1992.

[97] S. Zola-Morgan and L. Squire. The neuropsychology of memory: Parallel findings in humans and nonhuman primates. In A. Diamond, editor, *The Development of Memory*, pages 434–456. New York Academy of Sciences, NY, 1990.