A MULTI-TIME SCALE LEARNING MECHANISM FOR NEUROMIMIC PROCESSING

DISSERTATION

Presented to the Graduate Council of the

University of North Texas in Partial

Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

By

George E. Mobus, B.A., M.B.A.

Denton, Texas

August, 1994

A MULTI-TIME SCALE LEARNING MECHANISM FOR NEUROMIMIC PROCESSING

DISSERTATION

Presented to the Graduate Council of the

University of North Texas in Partial

Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

By

George E. Mobus, B.A., M.B.A.

Denton, Texas

August, 1994

Mobus, George E. A Multi-time Scale Learning Mechanism for Neuromimic Processing. Doctor of Philosophy (Computer Science), August, 1994, 127 pp., 2 tables, 32 illustrations, bibliography, 72 titles.

Learning and representing and reasoning about temporal relations, particularly causal relations, is a deep problem in artificial intelligence (AI). Learning such representations in the real world is complicated by the fact that phenomena are subject to multiple time scale influences and may operate with a strange attractor dynamic. This dissertation proposes a new computational learning mechanism, the adaptrode, which, used in a neuromimic processing architecture may help to solve some of these problems. The adaptrode is shown to emulate the dynamics of real biological synapses and represents a significant departure from the classical weighted input scheme of conventional artificial neural networks. Indeed the adaptrode is shown, by analysis of the deep structure of real synapses, to have a strong structural correspondence with the latter in terms of multi-time scale biophysical processes.

Simulations of an adaptrode-based neuron and a small network of neurons are shown to have the same learning capabilities as invertebrate animals in classical conditioning. Classical conditioning is considered a fundamental learning task in animals. Furthermore, it is subject to temporal ordering constraints that fulfill the criteria of causal relations in natural systems. It may offer clues to the learning of causal relations and mechanisms for causal reasoning.

The adaptrode is shown to solve an advanced problem in classical conditioning that addresses the problem of real world dynamics. A network is able to learn multiple, contrary associations that separate in time domains, that is a long-term memory can co-exist with a short-term contrary memory without destroying the former. This solves the problem of how to deal with meaningful transients while maintaining long-term memories.

Possible applications of adaptrode-based neural networks are explored and suggestions for future research are made.

ACKNOWLEDGEMENTS

I would like to thank my major professor, Dr. Paul S. Fisher and the members of my dissertation committee who have provided me with guidance, useful insights and most of all, stimulation of ideas.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

CHAPTER I

INTRODUCTION

1.1 The Nature and Significance of the Problem

Autonomous intelligence is characterized by the capacity of an agent to adapt to the contingencies of its environment. That is, the agent must be capable of incorporating new knowledge about the objects with which it interacts in its environment into internal representations. It must also learn the dynamic relationships between objects in its world in order to reason about how those objects behave so that it might effectively exploit opportunities and avoid threats. The purpose of knowledge might therefore be characterized as enabling the agent to predict the future state of the environment given the current and past states.

Operating in natural environments presents a particularly challenging problem to constructing autonomous agents. These environments do not have fixed boundary conditions, they are open to the influences of outside forces so that local conditions will, in general, change over time. As such, these environments are nonstationary and the agent's knowledge representations must, themselves, be able to change over time. Knowledge which is useful at one time may become useless at another time. Learning representations in such circumstances requires a real-time, on-line mechanism.

A consequence of the problem of nonstationarity in natural environments is that relationships between objects, and hence between those objects and the agent, may change over several time scales. For example, the diurnal cycle will change lighting and temperature conditions over the 24 hour period. Over a year changes in the seasons modulate these daily fluctuations. Over a yet longer period, that of, say, solar "sun spot" activity (roughly 11 years), the seasonal averages for light and temperature are themselves varied. And over the scale of geological time, shifts in tectonic plates and even the tilt of the earth relative to the orbital plane contribute yet longer time-scale modulations.

While an agent might not be required to maintain knowledge over geological time scales, it is certainly likely that it will be reasonably expected to do so over tens of years; as in the case of mobile, autonomous robots sent to Mars. Within this time scale the agent should be able to adjust its behavior to slowly changing patterns in the relationships of objects. A useful agent may be faced with pattern changes that occur over years or weeks down to minutes, seconds and even milliseconds.

Another complicating factor in natural environments is that there is reason to believe that many, perhaps most, processes in nature are operating in a strange attractor dynamic (West, B., 1993) or with inherent flicker noise (Bak, P., 1992). Interacting systems of this type are subject to episodic and possibly catastrophic change for which there is no known prediction method. One of the defining features of chaotic systems is that they may be highly organized in the sense of an attractor dynamic, while

remaining unpredictable due to their extreme sensitivity to initial conditions.

Against this backdrop of multi-time scale and chaotic changes in real environments we are still intent on devising machines which can acquire knowledge, make decisions based on their current situation and act autonomously. They are expected to survive and accomplish their assigned missions without explicit programming or *a prior* knowledge of all objects and relationships in their environments. As will be argued below, the majority of current approaches to machine learning, both in the traditional artificial intelligence (AI) paradigm of symbolic representation and in the neural network paradigm of distributed, subsymbolic representation, do not adequately address the problems of real-time, on-line learning in natural environments. The work to be presented here addresses these problems directly and the results indicate that an efficient, computational learning mechanism is realizable.

The economic impetus behind the above mentioned intent is the on-going desire to have machines do the unpleasant or dangerous work that would otherwise have to be done by humans. There is a broad applications potential for military, space and commercial use of autonomous agents. As an example, over the last several years many Small Business Innovation Research (SBIR) calls for proposals have detailed the need for robots to handle munitions, autonomous vehicles for transporting munitions and toxic substances, and explorer robots for planetary missions and underwater surveillance. Robots are needed to clean up toxic waste sites and nuclear facilities; jobs that are considered dangerous for humans but

that require considerable skill and on-site judgment with respect to navigation and negotiation of the local terrain under highly uncertain and changeable conditions.

Applications are important and certainly provide the economic incentive for research in autonomous agent technology. However, there is an equally important intellectual basis for this research. We want machines that can perform tasks which require considerable intelligence. The exploration of how to go about constructing machines that display intelligent behavior may provide insight into the computational properties of biological intelligence as well, giving important clues as to the nature of how intelligence might be achieved in animals and humans.

Machine Intelligence and Learning

Can we build machines that mimic natural intelligence in animals? What constitutes natural intelligence is a subject of ongoing debate in AI and cognitive psychology. There seems not to be a generally agreed upon definition of intelligence. Brooks (1991) has argued that a definition, per se, is not the issue. Rather, the focus of research should be on behaviors that agents can perform. Furthermore, Brooks argues that looking at the way in which evolution of natural intelligences built cognitive capacity atop lower-level sensory, perceptual, and motor systems provides the right view of how machine intelligence might be achieved. Brooks calls this approach the construction of "behavior-based agents."

Following Brooks' suggestion, if biology should be scrutinized more deeply for clues to the computational properties of intelligence then one might ask what properties of behaviors are common across phylogenetic lines on the presumption that such properties are fundamental to the emergence of intelligence. In the spectrum of behaviors, from primitive animals foraging for food up through human cognition, there seems to be at least one critical property shared by all: All organisms have the capacity, within qualified limits, to adapt to changes in their environments and/or improve their abilities to interact with their environments over time. In other words, a cornerstone of intelligence in biological systems is the ability to learn from experience. It is natural, therefore, to suspect that learning should be important for machine intelligence as well.

Even without a clear cut definition of intelligence we can describe machines that operate in a manner that produces what we would call intelligent behavior if it were done by a human or animal (Turing, 1950). Such a description offers an operational definition. From the pragmatic perspective of constructing machines that act as if they were intelligent, this may be sufficient. The key question then becomes, how can we construct machines that behave intelligently, and, in particular, that adapt their behavior to meet the conditions of a dynamic environment. The question will be approached from the basis of behavior-based systems.

A *behavior* is an action taken by an agent in response to an input signal or stimulus from the agent's environment. The behavior is *reactive* if its time course is in consonance with that of the stimulus signal (Brooks, 1991). Reactive systems are often characterized as operating in real-time.

The mapping of stimuli to behaviors, called an action map, (Kaelbling, 1993) in the general case, is a many-to-many mapping. Furthermore, stimuli signals may be real valued but will be considered as bounded for this discussion.

An agent may be characterized as having a set of primitive behaviors, $B = \{b_1, b_2, b_3, \ldots b_n\}$,*n* a fixed integer and a pairwise mapping $r(b_i, b_j)$, $i \neq j$, which maps the relationship between any two behaviors into the set, $R = \{b_i > b_j, b_i < b_j, b_i = b_j, b_i \neq b_j\}$. The relational symbols indicate competitive priority if two, or more, behaviors are initiated at the same time. The "not equal" symbol in this case means no priority - both behaviors may be performed simultaneously. Each primitive behavior is initiated by at least one stimulus in a set, $S = \{s_1, s_2, s_3, \ldots s_m\}$, *m* a fixed integer. An action mapping, then, is a set of functions relating the evoked behavior, b_k, to a stimulus, s_l, under the restrictions of *r* (). At any time step, *t*, (in a discrete analysis) some arbitrary subset of S may be non-zero leading to the activation of a subset of B. Such a subset is called a *composite behavior.*

It is assumed that some subset of the action map is fixed in the sense that a given stimulus will always result in a specific behavior unless over-ridden by the *r* () function. This corresponds to the idea of reflex responses in animals. It is likewise assumed that the compliment of fixed mappings are malleable in the sense that they may be formed as needed and can be zero until so formed. The action map can, therefore, be viewed as a sparse matrix of functions.

Intelligent behavior, by the above suggested operational definition, has the following characteristic: Given a novel situation, that is a subset of stimuli whose members have not been observed in this particular combination by the agent previously, the agent will respond after some reasonably short time interval with a composite behavior. If the result of the behavior is the acquisition of a reward, or the avoidance of a punishment, then, over several episodes the non-fixed elements of the action map will be altered so as to maximize the reward or minimize the punishment for the effort expended in behaving. The modification of the non-fixed action mappings is what will be meant by *learning* in the context of this work.

In the above, a tendency to optimize means that the objective function will show some improvement from trial to trial, perhaps leveling off after some number of trials. Many animal learning experiments show this pattern empirically (Alkon, 1987; Alkon, et al., 1991; Baxter, et al., 1991; Klopf, 1988; Mackintosh, 1983).

An Animal Learning Model and its Implications

Examination of a real-life animal learning model will help to frame the concept of leaning in intelligent behavior and serve to underscore some problems in current approaches to machine learning that will be addressed below. It will also provide a target performance criteria for the model to be developed in Chapter V.

Following a regimen of classical conditioning (Pavlov, 1927), an animal might learn to associate a given stimulus cue (e.g., a tone) with a

reward (e.g., a food pellet.) The animal learns from experience that the sound of a tone precedes the availability of a food pellet and orients itself so as to obtain the pellet as quickly as it can. If this regimen is repeated for a short time, say over the course of a day or two, the animal will show an initial association. After a time, this association will fade and the animal is no longer responsive to the tone alone. If, however, the regimen is repeated over a very long time, on the order of weeks say, the animal forms what can be shown to be a long-term response-to-tone association. In other words, the response to the tone, in orientation behavior, persists over a long time scale even if the reinforcement is not given on every event of the tone (a reduced reinforcement schedule).

Consider what happens if on several sequential occasions, after the tone sounds, and the animal orients according to its expectation for food, a mild shock is delivered instead of food. That is, a new relationship, contrary to that previously learned, has arisen. The animal responds by trying to escape the shock. But more subtly the animal's behavior changes over the course of these trials with respect to the 'meaning' of the tone. After not too many such pairings, it can be shown that the animal comes to expect not food, but a shock and so on sounding of the tone, it will respond with aversion.

If this treatment is followed by a sequence of tones without any subsequent shocks the animal's behavior undergoes a marked and interesting change. Initially the animal averts indicating that it has formed a short-term association between tone and shock. Over the course of several trials where shock does not follow the tone, however, the animal

may actually show signs of orienting as if it expected food. What appears to have happened is that the animal's long-term memory of the tone-food association re-emerges when the short-term memory of tone-shock extinguishes. What this model demonstrates is the subtle but important role of multiple time scales in the learning of associations. It can be seen that some association which is valid in one time scale may be completely contrary in a longer (or shorter) time scale.

There is a second, more subtle but very important aspect of the above animal learning model. The association encoding that took place in the animal, the conditioning of a non-semantic (that is non-meaningful) stimulus to a semantic stimulus-response pair, encodes what can be called a *causal relationship*. That is, a causal relationship exists between two events, A and B, iff A nearly always precedes B in occurrence by not more than some (usually small) finite time difference and B never precedes A. B may occur without the occurrence of A preceding it, or A may occur without B always following but the two must co-occur some statistically significant number of times. A causal relation, so defined, is simply a correlation with an imposed temporal order. By this definition it is possible to derive the inference that either A causes B or A and B are jointly caused by some third, unseen event - in the case of classical conditioning whatever caused the food pellet to appear also caused the tone. In either case the temporal ordering is important for the following reason: for the animal, the non-semantic cue, the tone, became a predictor of the semantic stimulus, food or shock. In the more cognitive arena of AI causal encoding is a significant factor in both reasoning from the present state to the future state and

in abductive reasoning, from the present state to prior states. Causal relation encoding is a deep problem in knowledge representation and knowledge-based systems in AI (Pazzani, 1990).

1.2 Purpose of the Research - Steps Toward Machine Intelligence

From the proposition that a critical component of machine intelligence must be the capacity to learn from experience and adapt behavior accordingly, the objective of the research reported here is to obtain a computational model of a learning mechanism that addresses the problems of multiple time scales and chaotic dynamics as embodied in the contrary association memory task outlined above. Given the intention that this model should be realizable in a working agent, it will be, ideally, computationally efficient in both time and space.

The thesis advanced here is that the construction of autonomous, intelligent agents will depend on the availability of a mechanism for learning representations of causal relations in the environment and for associating those representations with semantic (that is meaningful to the agent) stimulus-behavior pairs pre-existing in the agent's action map. The mechanism should be capable of encoding relations that vary over a wide range of time scales and it must be capable of retaining associations that separate in time, even when those associations would be mutually exclusive in a single time scale as in the animal learning model above.

Machine intelligence has been pursued by two, quite different routes. Traditional AI has assumed that intelligence is a property independent of

the underlying machinery. Under this assumption the mechanisms of biological intelligence are unimportant. What AI has sought to do is find algorithms that simulate the performance of tasks that would, ordinarily, require intelligence. The second approach considers the underlying mechanisms of biological intelligence as possessing properties important to the emergence of intelligence in general. This route, using computational models of brain-inspired processors called neural networks, has been explored in solving problems in pattern recognition and classification and adaptive control. What makes these models potentially very powerful is that they provide a means for naturally integrating some method of learning directly into the processing component. In one of the most studied models, the multi-layered perceptron, the network learns a mapping of inputs to outputs, given an exemplar set of mappings, even for nonlinear maps.

At the same time that successes are accruing to neural network learning research, the mainstream AI community is placing increasing importance on learning as a necessary component of any model of intelligence. After an initial combative stance (Fodor & Pylyshyn, 1988) that put the symbolic school (traditional AI) at odds with the subsymbolic school (neural networks), in more recent years there has been a growing amelioration between symbolic AI models and neural network models of knowledge representation and reasoning (c.f. Aparicio. & Levine, 1994; Shastri & Ajjanagadde, 1993).

Neither the performance paradigm of conventional AI, nor the brain paradigm of neural networks has yet generated truly autonomous

intelligent agents. Such an agent is expected to act in a real environment where events are sporadic and episodic. Most attempts to model intelligent behavior have been done in abstract, simulated environments such as Blocks World for SHRDLU (Winograd & Flores, 1986) or in problem domains selected specifically to avoid the open systems problem. These worlds are convenient for formal analysis but have failed to capture the essence of real world dynamics. Brooks has argued that real autonomous intelligence cannot be built into agents that are not embodied and situated in a real world (Brooks, 1991). Agents must have the capacity to deal with stochastic processes of real environments while having observational access to a small sample of events in those environments.

In this work is presented an computational model of a learning mechanism, called an adaptrode, that will be shown to efficiently encode multi-time scale correlations of both associative (cross correlations) and non-associatiative (autocorrelations) signals. Adaptrode-based neural networks are used to implement an action map as outlined above. It will be shown that such a network can address the problem of learning in potentially chaotic environments, thus providing a step in the direction of machine intelligence.

1.3 Methodology

Of late, a growing number of researchers have become interested in what may be called biologically-inspired computation. Artificial neural networks, as mentioned above, are examples. The guiding principal in this program is that biological systems, in particular animals, solve some

interesting and hard computational problems - at least approximately - and if one could discover the mechanisms underlying the computation, one could design emulators that could mimic this ability. Rather than the 'black-box' approach of AI, the new approach seeks to reverse engineer the brains of animals in order to construct computational counterparts. It shares philosophical underpinnings with the behavior-based, subsumptive architecture agent espoused by Brooks (1991).

The effort reported here follows such a program. The steps are: 1) consider a problem that at least some animals solve that cannot now be solved as readily algorithmically, e.g., pattern recognition; 2) characterize the method by which an animal solves the problem - a combination of biological research and mathematical modelling; 3) infer a mechanism that emulates the problem solving behavior; 4) simulate it and see if it captures the essential qualitative properties of the biological system.

This work presents a computational model of an adaptive element that, used in the context of a neuromimic synaptic processing element, fits the animal learning model above. In doing so, the mechanism employed in the computation provides a means for constructing neural network-based agents with at least the primitive level of intelligence seen in simple animals that exploit causal associations across multiple time scales to improve their chances of survival in a complex, nonstationary environment.

The model is obtained from consideration of memory trace dynamics in living synapses. To test the model, it is incorporated into a neuromimic architecture that, in turn, is shown to generate behavior which has the essential properties of a very fundamental form of learning in animals -

classical conditioning. It will be shown that the model meets the important criteria of classical conditioning reflected in animal learning studies. Several classical conditioning phenomena will be demonstrated following methods developed by Sutton and Barto (1987) and Klopf (1988). In addition, this model goes beyond prior studies of these basic criteria to show that adaptrode-based neuromimes have the capacity to encode memory traces that separate in time scale, even when those traces encode contrary associations.

1.4 Organization of the Thesis

The thesis is comprised of six chapters. This chapter has served to introduce the argument that intelligent behavior in machines will depend on the capacity to learn from experience. The associated problems of multiple time scales and nonstationary environments were also introduced. Chapter II will provide a review of related work in neural network learning, particularly the area of reinforcement and real-time learning rules as models of classical conditioning in animals. In Chapter III the adaptrode model is obtained from consideration of the dynamical behavior of living synapses as adaptive filters. Whereas the adaptrode model as derived in Chapter III is based on the dynamics of observed behavior of synapses, Chapter IV will provide a more detailed account of the correspondence between the adaptrode model and the biophysical model of processes underlying synaptic plasticity.

Chapter V will present the simulation models that demonstrate the behavioral efficacy of a network comprised of neuromimic processors

using adaptrode-based synapses in the context of classical conditioning. Additionally, Chapter V will show how the adaptrode model extends the capability of learning into multiple time scales with non-interference between mutually-exclusive stimulus-response associations that separate in time. This demonstration is the basis for the claim that the adaptrode can address the issue of learning in a nonstationary environment.

The simulation results together with the demonstration of structural correspondence between the adaptrode model and biological synapses in Chapter IV constitute a compelling argument for considering the adaptrode as a viable alternative to learning rules now employed in neuromimic processing.

Finally, Chapter VI will summarize the conclusions of this research, give examples of the applicability of the adaptrode model and suggest avenues for future exploration.

CHAPTER II

A REVIEW OF RELATED WORK

2.1 A Neurobiological/Artificial Neural Network Connection

As argued in Chapter I, the hallmark of intelligence in animals is the capacity to adapt behavior in response to environmental change. Adaptability is based on learning spatio-temporal patterns that, themselves, may change over time. An important feature of learning is that an animal can predict the near-future state of the environment based on prior states (Sutton & Barto, 1981). Prediction should allow the animal to anticipate a future stimulus that has physiological consequences (meaning or semantic content) and emit appropriate behavior in anticipation of that stimulus. Computation of expected states must rely on some form of time series analysis, particularly for non-associative adaptation (Koshland, 1980), and clearly involves correlational analysis for associative adaptation (Sejnowski, Chattarji & Stanton, 1989; Hebb, 1949). Furthermore, the computation of a future expected state wherein the occurrence of an event (or pattern) in one channel predicts the occurrence of an event (or pattern) in another channel, as will be seen is the case in classical conditioning, must involve aspects of both of these techniques.

There is reasonably general agreement in neurobiology that the principal site of learning, in the sense of encoding memory traces, is the synaptic junction. Both pre-synaptic and post-synaptic mechanisms have been elucidated (Alkon, 1987; Kandel & Schwartz, 1982; Small, Kandel & Hawkins, 1989). Each patch of post-synaptic membrane contributes to the overall depolarization (or hyperpolarization, in the case of inhibitory synapses) of the cell body membrane with each arrival of a signal pulse called an action potential. A sufficient depolarization of the cell membrane can result in the initiation of an outgoing action potential along the neuron's axon. Action potentials are fixed-amplitude and fixed-width pulses arriving at the pre-synaptic bouton (ending of the axonal branch). Synapses are known to produce a variable amplitude response, the post-synaptic excitatory potential (EPSP), for a fixed-amplitude stimulus depending on the frequency of the stimulus. Increases in the capacity of a synapse to contribute to the depolarization (or hyperpolarization) of the membrane is called the synaptic efficacy.

Artificial neural network models, currently under investigation, are idealizations of biological neurons based on large-scale known features of the latter and some presumed mechanisms of information processing. The computational model of a formal neuron can be traced back to that proposed by McCulloch and Pitts (1943). This model was inspired by what was known of biological nerve nets at the time, which was basically the gross morphological view of neurons and a limited anatomical model of a few neurological preparations (c.f. Kandel & Schwartz, 1981, Chapter 1, pp 3-13). The prevailing view of the time was the "cellular connectionism"

model of Ramon y Cajal and Karl Wernicke. In this model, neurons are the signalling elements of the brain. Very little was known of the details of the neurophysiology of individual cells and even less about the synaptic connections between cells.

Against this background, and motivated to find some commonality between the functioning of nerve nets and computing machinery, McCulloch & Pitts considered the obvious features of neural tissue that might be a basis for such a commonality. Neurons were known to have a large fan-in of incoming signals through their dendrites. The existence of a discrete action potential argued for a boolean-type logical operation. There was a single output channel, the axon, which might, however, fan-out to many destination cells. Very little was known at the time about synaptic efficacy changes as the basis of learning. Additional prominent features of neural tissue seemed to be the substantial number of neurons involved, and even more substantial number of interconnections made between them.

Without a basis in actual knowledge of what neurons did, internally, to process information, McCulloch and Pitts proposed their model of a simple threshold-logic element. In the milieu of "brain-as-computer" paradigm that was emerging in the period, the idea of neurons being very simple logic gates appealed to many other researchers (c.f. von Neumann, 1958). McCulloch and Pitts showed that their model could compute arbitrary logic functions. Hence, they provided a tacit reinforcement to the notion that real neurons were indeed simple computing elements.

The output for a McCulloch-Pitts formal neuron, x_j, is given by:

$$x_j = f(\sum_i x_i w_{ij} - \theta_j) \qquad (2.1)$$

where w_{ij} are the weights associated with each input line from neuron (or external source) i to neuron j, and θ_j is a threshold value. The function $f(\bullet)$ in the original formulation is a simple binary output, i.e., $x_j = 1$ if $f(\bullet) > 0$, 0 otherwise. This outer product interpretation of synaptic efficacy has largely been accepted without question in the ANN community but has been questioned by a number of neuroscientists (Crick, 1989; Shepherd, 1992). The actual mechanisms for synaptic contribution to neuron firing is drastically different from the formal model (Sejnowski & Qian, 1992).

McCulloch and Pitts used fixed weights for their outer product formulation. Neurons did not adapt or learn. In 1949 Donald Hebb published his book, The *Organization of Behavior*. In it he provided a relatively simple model of how one neuron might encode the correlation of activity between itself and another neuron. The "Hebb" rule of connection weight modification became the basis for applying learning mechanisms to McCulloch-Pitts neurons. The idea is simple enough. If a neuron that is providing input to another neuron is found to be firing at the same time the second neuron is firing then there is an increase in the weight on the link between them. The simple Hebb rule implies:

$$w_{ij}(t+1) = w_{ij}(t) + \Delta w \qquad (2.2)$$

and

$$\Delta w = g(w_{ij}(t), x_i(t), x_j(t)) \qquad (2.3)$$

where $(t+1)$ is the next time step, (t) is the current time step and Δw is the change in weight from t to $t+1$. The function $\mathcal{g}(\bullet)$ is determined by the specifics of a learning law that is being implemented.

What is cogent in this brief review is the strong influence this model has had on neural network models in general. Specifically, there has been a generally held belief, sometimes voiced, but often lurking as a hidden assumption, that neurons are *just* simple processing elements. This has been joined and reinforced by another view that holds that it is the network level that does the interesting processing (Hopfield, 1982; Grossberg, 1991a). Emergent behavior from simple interacting elements has been the credo of the neural network field for most of the past decade. Until very recently this has had a stifling effect on the investigation of more biologically realistic models of neurons as potentially more powerful computing elements.

Learning in these formal neurons is the modification of the link weights according to a learning rule. As more detailed information has emerged regarding the molecular mechanisms underlying synaptic plasticity, coupled with an understanding of the behavioral aspects of learning, particularly in invertebrates, more biologically-plausible learning rules have been proposed and studied. Alkon, *et al.*, (1990 and 1991) have proposed a model of reinforcement learning called DYSTAL (DYnamically STAble Associative Learning) based on their findings of local interaction potentiation in *Hermissenda crassicornis*, a marine nudibranch, and also found in the hippocampus of the rabbit. Similarly, Gingrich & Byrne, (1987); Byrne, & Gingrich (1989); Byrne, Gingrich, &

Baxter, (1990); Buonomano, Baxter & Byrne, (1990); and Baxter, *et al.*, (1991) have done extensive modelling of learning mechanisms based on their work with *Aplysia*, another marine snail. A land snail, Limax, was the inspiration for the LIMAX model of taste aversion learning of Gelperin, Hopfield & Tank, (1985). Another model that sought to capture biological realism in the form of the temporal displacement between input and output signals, Drive-Reinforcement, was proposed by Klopf, (1988). Additional models that attempt to adhere to biologically plausible considerations may be found in (Levy & Desmond, 1985; Lynch, Granger & Larson, 1989; Moore & Blazis, 1989; Sejnowski, Chattarji & Stanton, 1989; Shepherd, 1992; Sutton & Bartow, 1981).

Though these models have been based on biologically plausible considerations and, in many instances have generated results that fit the empirical data, two common aspects of the proposed learning rules are the primacy of associativity and singular time scale dynamics. The first aspect refers to the fact that the cause of a change in a weight value always depends on the correlation between two signals. The most widely cited example is the Hebb (1949) rule as given above. The second aspect refers to the dynamics of weight change which proceed only in the time scale of changes in the signals. Kosko (1991) has shown that simple associative rules compute the exponential weighted average of changes in the correlated signals such that the a memory trace would be washed out by subsequent processing. The trace would be lost exponentially fast when the conditions change. Efforts to preserve a trace in this context have to disallow forgetting. Thus is produced a conundrum for continuous, on-

going learning based on simple associative and single time scale encoding. Either you remember none of the past or you must remember all of it. Clearly, biologically plausible learning rules must consider long-term memory phenomena as well.

There are some additional biologically relevant problems with these two aspects in simple associative-based, single time scale learning laws. First, with respect to required associativity, there are a number of examples of non-associative memory formation and retention in excitable cells (Alkon, 1987; Alkon, 1989; Byrne & Gingrich, 1989; Byrne, et al., 1990; Kandel & Schwartz, 1982; Small, et al., 1989; Staddon, 1993). Habituation, sensitization, and muscle development with training are several examples of purely activity-dependent encoding mechanisms. What is learned is a temporal association or expectation of future activation based on past and current activation. Post-tetanic potential (PTP), a transient elevation in synaptic efficacy following a brief burst of afferent action potentials (the stimulus trace of Sutton & Bartow, 1981), and long-term potentiation (LTP) in mossy fiber afferent synapses of the CA1 pyramidal cells of the mammalian hippocampus are other relevant examples of plasticity that is nonassociative.

The second problem that current rules cannot adequately address is the role of time domains in memory formation and recall. Current real-time models do address the need to differentiate between the activity memory trace (Klopf, 1988; Sutton & Bartow, 1981) or a short-term memory (STM) through recurrent self-stimulation (Grossberg, 1987 and 1991) and a synaptic efficacy change representing long-term memory

(LTM). Little has been done, however, in artificial neural networks, to investigate the dynamics of memory over several time scales of biological importance (see Alkon, 1987; Alkon, et al., 1990; Alkon, et al., 1991; Staddon, 1993 for exceptions).

2.2 Classical Conditioning: A Model Paradigm for Causal Knowledge Encoding

To begin understanding how an animal perceives and reasons about the world, how it captures and uses the regularity of nature, we might ask how it is that causality, or more precisely, causal relations, are encoded in the brain. How does the brain represent causal relations and how are these representations learned? To make the task approachable we can note first that all animal life is faced with the same problem of discovering and using regularity in the quest for survival and propagation. Thus by studying primitive (that is phylogenetically simpler) brains we may discern some mechanisms for such encoding which will be found to be invariant across the phylogenetic spectrum. What has been sought is a direct connection between neurophysiological processes of information encoding and the outward behavior of the animal that shows adaptation to changing environments.

During the past decade some real progress has been made in understanding how simple neural systems function with regard to encoding information in synaptic efficacy (Alkon, 1987, 1989; Getting, 1980; Jaffe & Johnston, 1990; Kandel & Schwartz, 1982; Small, Kandel & Hawkins, 1989; Starmer, 1987). Additional insights into the details of neural

representation, at the network level, learning and processing have been gained by building formal models of neural networks that simulate aspects of behavior in animals using these rules for synaptic encoding (Alkon, *et al.*, 1990; Buonomano, Baxter & Byrne, 1990; Byrne & Gingrich, 1989; Byrne, Gingrich & Baxter, 1990; Gelperin, Hopfield & Tank, 1985; Klopf, 1988; Klopf & Morgan, 1990; Koch & Segev, 1989; Morgan, Patterson & Klopf, 1990; Rumelhart & McClelland, 1986; Sejnowski, Chattarji & Stanton, 1989). The empirical studies in neurobiology along with these formal models demonstrate the potential connection between synaptic efficacy and overt animal behavior due to adaptive response to changes in causal relations in the environment.

Specifically, the learning model of classical conditioning, demonstrable in some very primitive animals, and considered as a fundamental basis of higher-order learning mechanisms (Dretske, 1988; Mackintosh, 1983), is an operational version of learning and representing a causal relation. This follows from the fact that the success of conditioning has been shown to depend on the temporal ordering and contiguity of the conditionable signal with the unconditionable signal. Therefore, the general approach taken is to find correlates between neural events and the acquisition of conditioned response learning so that the former may be posited as contributory to the latter.

The outline of neural and even molecular substrates of conditioned learning has begun to emerge from the laboratories of neuroscience (Alkon, 1987; Kandel & Schwartz, 1982; Small, *et al.*, 1989). Connectionist views of classical and operant conditioning, likewise, have

provided computational models which may aid in the discovery of the invariants mentioned above (Carpenter & Grossberg, 1987a, 1987b; Grossberg, 1987; Sutton, 1988; Sutton & Barto, 1987, 1991). Such models, constrained by the wealth of neurophysiological and psycho-physical data regarding conditioning, provide valuable insights into the principles of learning and may produce further hypothesis for investigation. Additionally, one hopes to find clues that will aid in the construction of machine-based mechanisms with animal-like learning competence (Anderson, Merrill & Port, 1989; Elman & Zipser, 1988; Morgan, *et al.*, 1990; Rumelhart, Hinton & Williams, 1985).

How does classical conditioning relate to the notion of a causal relation? The answer to this question provides a final link between learning models and higher-order intelligence in autonomous agents. As mentioned above, it is the temporal ordering and contiguity as well as contingiency constraints that earmark conditioned response learning. These are exactly the characteristics of a causal relation.

A causal relation between two events, A and B, denoted $A \Rightarrow B$, has several important properties which impose temporal constraints on the form of any representation and derived inference as described in Chapter I. As noted already, an event A must precede an event B in order for us to say that A is a cause of B. We would never say that the breaking window caused the ball to be thrown. More precisely we require that the onset of A precede the onset of B by some $\Delta t > 0$. We can relax this constraint somewhat by requiring (or allowing) that A almost always precedes B, where 'almost always' refers to some statistically defined frequency. There

may be other causes of B and the occurrence of A may not always result in B. The constraint, however, strongly requires that B never precedes A, at least within some defined latency period, which is to say the temporal relation is one-way.

Another, related but technically different, constraint is that of temporal contiguity. The events must occur within a contiguous temporal window of opportunity. This window is defined by the context of events and the memory trace retention of event A. Thus it is not that B must occur shortly after A, but rather, B must occur after A but before the memory of A fades and no intervening event changes the context established by A. This is a subtle aspect of contiguity not often fully appreciated. Memory plays an important role in inference of causality. If B occurs too long after A has occurred, then the linkage between them is weakened.

A third constraint has to do with computing the correlation of A and B over time or contingency. In probabilistic causality we allow that the occurrence of A may increase the probability of the occurrence of B. The event B can have other, unobserved causes. Thus is inferred a causal relation between A and B only if the frequency of co-occurrence is sufficient to the purpose (note this need not be a majority of the times). The temporal constraint requires that a sufficient period of time has to pass in which multiple occurrences of A and B can be experienced. Note again that memory is involved in keeping a record of co-occurrences and that memory must persist over the time scale of the 'sufficient' observation

period. This is usually considerably longer than the short-term memory trace of the occurrence of A mentioned above.

These constraints, implicit in the prototypical causality rules, underscore the importance of the role of time in learning, representing and processing (reasoning) causal inferences. Furthermore, we can see from the above discussion that temporal information extends across several time scales. It is not sufficient to deal with just the time scale of the real-time events (A and B). Therefore the nervous system must employ mechanisms which encode temporal information in multiple time domains.

The conditioned learning task embodies the rules of causality and is well documented in a wide variety of vertebrate and invertebrate animals. As such it has become a standard testbed for computational models of reinforcement learning (Alkon, 1987; Alkon *et al.*, 1991; Aparicio, 1988; Dretske, 1988; Gelperin, *et al.*, 1985; Grossberg, 1991; Klopf, 1988; Mackintosh, 1983; Pavlov, 1927; Sutton & Barto, 1987, 1991; Tesauro, 1986). In this paradigm (of which there are two flavors) the animal (model) learns to associate an event such as an environmental cue with a behavior (or with a consequence). Though the notion behind conditioning has its origin in laboratory experiments, the general idea of conditioned learning and how it benefits the animal in its natural environment is fairly straightforward. If an environmental event (cue) which is neutral with respect to the animal's survival is found by experience to precede another event which has direct survival impact, such as the availability of food or pain caused by a wound, then the animal forms a lasting association between those events and can use the prior event to predict the occurrence

of the second event. Such predictive ability allows the animal to respond more quickly to the impending meaningful event. This can be seen in the experimental paradigm called classical or Pavlovian conditioning (Pavlov, 1927). The animal learns to associate a conditioned (or conditionable) stimulus (CS) with an unconditioned (hard wired) stimulus-response pair (US/UR). The association results in a conditioned response (CR), which is either the same as or similar to the UR, being elicited upon presentation of the CS alone. The animal has learned to respond to the CS as if it were the US.

The second form of this type of learning involves the stochastic emission of a behavior which reliably produces some beneficial result (for the animal) in the environment. The animal learns to associate the behavior with the beneficial result and can, in principle, emit the behavior by choice in order to elicit the result. This latter form of learning, termed operant (or instrumental) conditioning, depends on a contextual situation, for example a physiological drive, which creates the condition in which the result will be beneficial. In the laboratory setting, animals are kept hungry so that they are "motivated" to press levers or buttons so as to receive a food pellet.

How are these conditioned learning tasks to be viewed as simple causal representations? There is the philosophical side to this question wherein we can speculate over the animal's perceptions and beliefs about cause and effects (Dretske, 1988). However, I am more interested in an operational view in which the animal behaves as if a cause and effect relation has been learned. Whether A (the CS) actually causes B (the US) (which in turn

causes C, the UR) in an objective sense, or is the perception of the animal is of little concern just now. The point is that the animal responds to A (with C) *as if* a causal chain had been established. From the standpoint of behavior, there exists the inference of a causal chain.

The balance of this chapter will focus on the way in which neural networks can encode the temporal and associative rules of causality. This will be examined at the level of conditioned response learning where it will be shown that a neural architecture can be built in which the prototypical rules, in particular the temporal constraints, of causality can be instantiated in representations of events which are causes (predictors) and events which are effects (consequences or reinforcers). If it is true that conditioned learning is a fundamental basis for higher forms of learning as has been suggested (Dretske, 1988; Mackintosh, 1983), then this approach may fulfill the promise of directing us toward a theory of causal inference learning and reasoning in higher cognitive processes.

2.3 Representing Time in Neural Networks

As has been argued above, the representation of time plays a central and crucial role in causal inference. In order to meet the temporal constraints of causality it is necessary to show how a neural network can encode temporal knowledge, integrated with spatial associative knowledge, such that information processing produces the correct inferential result. In this section some methods which have been employed to incorporate temporal representation in neural networks are briefly reviewed.

The Primacy of Associativity

The vast majority of learning rules that are used in neural network architectures have at their base the assumption of associativity as the driving influence in changing the edge weights associated with processing element inputs. What this means, simply, is that in order for a weight to be modified, there must be some kind of correlation between two or more independent signals in the processing element. These can be, for example, an input correlated with the output of the element such as the Hebb rule (Hebb, 1949; Hertz, Krogh & Palmer, 1991), input correlated with an error (difference between the output and a desired output or delta rule, Widrow & Hoff, 1960), or two inputs (local interaction rule, Alkon, *et al.*, 1990). Such rules can be further modified to take into account the time derivative of the signals (Klopf, 1988; Kosko, 1986) or the error derivative in the case of the generalized delta rule (Rumelhart, *et al.*, 1985). However, the change in a weight is still dependent on the activity of two or more signals.

This form of associativity is a spatial encoding mechanism. A large number of neural network applications have addressed the issues of pattern learning, classification and recognition. Their successes have led to what seems to be a general consensus that learning rules must, at their base, be associative. As a result this has led to an interesting problem: How to represent temporal knowledge when the basic rule is spatial (Elman, 1990).

Adding temporal representation to associative-based networks

A typical solution to this problem has been an attempt to add temporal representation to an otherwise associative learning scheme. The earliest efforts attempted to construct architectures in which time could be represented by a spatial analogue. For example an avalanche circuit could be used to represent time steps in a sequence (Grossberg, 1982).

Another method for representing temporal information at the level of the network is through recurrent loops and/or time delay units (see Hertz, *et al.*, 1991 esp. Section 7.3). In the simplest version a neuron excites itself through a loop with inherent decay. Grossberg and others have used this method to instantiate a short-term memory (STM) function at the level of individual neurons (see Grossberg, 1991a for an excellent review). An associative rule is used to update the link weights in order to form long-term memory (LTM) encoding on the other inputs to the neuron.

Recurrency can be used on a network-wide basis as well. One neuron can excite (through a non-learning connection) another neuron which can, in turn, excite the former directly or through a chain of neurons. This method has been used by a number of workers using the backpropagation learning method. Context units which feed back historical states to earlier layers in the network act as a memory of prior input and processing, thus having an impact on the current processing (Elman, 1990). Systems of this kind have been shown to be able to learn and recognize and/or generate sequences of spatially extended patterns.

Adding associativity to temporal representation - an alternative

In the types of systems described above some form of temporal representation is added to a network architecture where learning is based on an associative learning rule. An alternative approach would be to add associative encoding to a temporally-based learning rule. I had been interested in the biological phenomenon of adaptive response in which an animal modifies its response to a stimulus based on the time course behavior of that stimulus. This is a fundamentally nonassociative phenomena which can become associative through modulation processes.

The problem of temporal representation is more complex than simply representing sequences. Perhaps the most important aspect of temporally extended phenomena is that modulation may occur over multiple time scales. The impact of real-time signals, in terms of the synaptic contribution to the neuron's decision to fire, can be modulated by the longer time course activity of signals in that channel. The following simple example of adaptive response may help to illustrate: Muscles undergo adaptive response to athletic training.

The real-time response of the body to athletic demand (performance) changes with time and training. If a rigorous training schedule is maintained the athlete's muscles strengthen and increase in bulk as new tissue is created. This will occur only if the demand schedule is maintained over an extended time period. The muscles come to expect an increase in workload as a result of past experience. They adapt to the expected level of demand. This is a form of autocorrelation learning which is based

primarily on activity and time. There are short-term effects, intermediate-term effects and long-term effects that come into play.

Muscle growth will, however, be constrained or modulated by associative factors such as nutrition - adequate protein and vitamins in the diet. Adaptation in the form of muscle growth will take place only if these factors are satisfied. In this example associative factors (nutritional elements) act as modulators to an otherwise temporal encoding process (training). There are numerous other examples of activity-dependent adaptation in biological systems.

Could adaptive response be a biologically plausible basis for learning in neural networks? More to the point, could this phenomenon be used as the basis for efficacy modification of synapses. If so, could the temporal representation meet the criteria of causal inference, which from prior arguments means conditioned learning? In Chapter III a model of synaptic contribution and plasticity based on adaptive response is obtained and will be shown to have the desired temporal properties. An immediate advantage of this approach is that it will provide a passive means for encoding short-term memory traces at the synapse rather than maintaining activation of the neuron.

Multiple Time Scale Encoding

Recently, multiple time scale dynamics in biological synaptic plasticity and adaptive response has gained attention (Alkon, 1987). Chapter IV provides a more in-depth description of this concept, however, the basic notion is outlined here. In the post-synaptic compartment, the interior

portion of the neuron underlying the synaptic junction, a sequence or cascade of biophysical processes interact. The initial process is driven by the change in membrane potential. Each process in the cascade has a longer time course behavior than its predecessor, as measured by the accumulation of some chemical constituent. In turn, that constituent, acting as input to the next process in the sequence, drives the next process in the forward direction. Conversely, through a set of unique positive feedback loops, a subsequent process either reduces the rate of the reverse direction of the prior process or actually increases the forward rate of that process. These cascades constitute a series of ever longer-term memory traces in the compartment. Figure 1. shows a conceptual representation of such a cascade of processes.

Multiple Time Scale Process Cascade

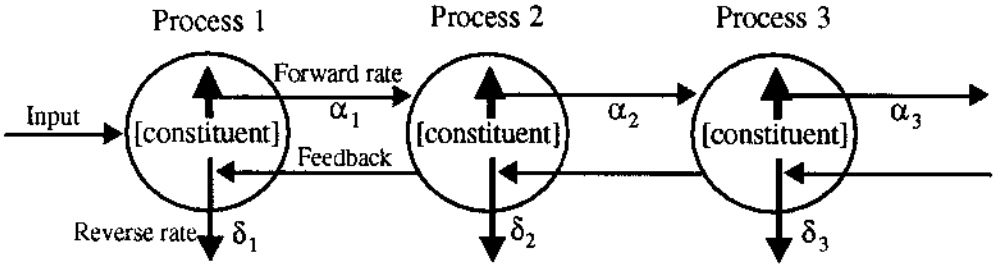

Fig. 1. In a cascade of interacting processes, each process accumulates some constituent that drives the forward rate of the next process by time constant α. The process has a reverse or decay direction with time constant δ. Each subsequent process operates over a longer time scale so that the constituent of Process 2 accumulates much more slowly than Process 1, and that of Process 3 even more slowly. Feedback from a later process to a former one reduces the rate of decay of the former creating a "memory" trace in the former process.

The long-term effect of this series of processes with their feedforward and feedback interactions is that the synapse displays varying degrees of efficacy in contributing to the neuron's firing depending on the time course behavior of the input signal (associative factors aside for the moment). The synapse might, thus, be viewed as an adaptive filter since its response, in the form of the local membrane potential, varies according to its input history. It is cogent to ask, then, what is the behavior of an artificial neural system in which the simple link weights are substituted by an adaptive filter with the characteristics of biological synapses.

A recent attempt to model a multi-time scale non-associative memory phenomenon, habituation, has been carried out by Staddon (1993). He uses a two domain (time scale) model to show how such a cascade produces model output that better fits the empirical data. Somewhat similar to the model to be presented in the next chapter, Staddon linked two leaky integrators (exponential averagers) in series. What is important to note, here, is that the time constants of integration of the two stages are set so that the second stage operates over a significantly longer time scale. Staddon shows that the depth of habituation and the recovery rate from it in the model match the rate sensitive variations seen in living animals. A simple single stage integrator cannot accommodate the rate sensitive data.

Staddon's model is important in showing that a by incorporating consideration of multiple time scale encoding, even as simple two-stage model, produces significant gains in modelling the behavior or real memory systems in nature.

2.4 Concluding Remarks

This chapter has provided a background review of work related to the research and has attempted to show the relationship between three main disciplines that contribute to the work reported here. These are neurobiology, artificial neural network modelling and animal learning theory. Over the next several chapters I will develop the adaptrode model and examine it, and its consequences for machine learning, from all three perspectives. The integration of these perspectives into a realizable mechanism for real-time, on-line learning is central to effort to produce a working approach to artificial intelligence.

CHAPTER III

A BIOLOGICALLY-REALISTIC MODEL NEURON

3.1 Overview

In this chapter the adaptrode model is introduced as a middle ground between the simple weighted input synapses used in neural networks, as described in the previous chapter, and the much more complex models of membrane dynamics used in neurobiology. In the adaptrode model the synapse becomes a much more powerful computing element as is posited for biological neurons (Shepherd, 1992). This is as contrasted with artificial neural networks in which the neuron is considered the principal processing element. It will be shown that this increase in power allows for the emergence, in an artificial neuron, of important capabilities one would want such a neuron to possess and that are not achieved by the simpler models. It will also be shown that this increase in capability is bought at a relatively cheap price in terms of computational complexity relative to the more complex neurobiologically-derived membrane models. That is, the computational requirements for this model are not much greater than those of simple neural networks, but the processing power achieved is significantly greater in the sense that the model produces system-level memory acquisition and recall performance that approaches simple biological models.

All models seek to condense the essence of the modelled system in order to capture its behavior while reducing unnecessary detail. The key insight in the approach taken here is that the relevant phenomena giving rise to memory trace acquisition and recall at the synapse appear to segregate by time domain. Therefore, in the development presented below the state variables, representing memory traces, have been collapsed by categories of time scales, or domains. The resulting model is greatly simplified, compared to the membrane models, yet appears, by the results obtained here, to retain the essential features of the dynamical properties of real synapses. That is, the response of the adaptrode adapts to changes in the input signal over time scales covering several orders of magnitude. Thus the behavior of an adaptrode is a closer analogue of biological synapses and the artificial neuron resulting from use of the adaptrode synaptic processor will become a closer analogue of biological neurons as well.

3.2 An Adaptive Filter Synaptic Model

The adaptrode seeks to model certain key aspects of the qualitative behavior of biological synapses, namely the way in which a synapse acts as an adaptive filter. The next chapter will be devoted to a discussion of the correspondence between the adaptrode model and the multi-time domain model of biological synapses presented by Alkon (1987). The present chapter will only focus on the overt behavior of a prototypical biological synapse and provide a motivation for the development of the adaptrode mechanism.

Biological synapses are not simple multipliers with fixed weights, relative to the time scale of the input signal (Kandel and Schwartz, 1981, esp. Chapter 7, 1982; Alkon, 1987; Shepard, 1992). Most classical neural network models treat synaptic inputs as either time-averaged firing frequency of the presynaptic cell (real-valued numbers) or as binary (0,1) or bipolar (-1,1) representing discrete action potentials. These inputs are multiplied by a weight representing the 'efficacy' of that synapse. There is no time-dependence in the process, except through the longer time scale of learning as represented by something like the Hebb rule.

The change in efficacy of biological synapses is quite different. The postsynaptic membrane undergoes an activity-dependent change in membrane potential that may bring it closer to or farther away from the triggering potential for an action potential to fire. An excitatory postsynaptic potential (EPSP) contribution is due to an excitatory synapse while an inhibitory postsynaptic potential (IPSP) is due to an inhibitory one. Figure 2. shows a stereotypic trace of a unitary EPSP following a single spike arrival. The horizontal axis is time in milliseconds, while the vertical axis is membrane potential with respect to the exterior of the cell which is more positive. The spatial summation of these local transients gives rise to the final potential arriving at the axonal hillock. It is this summation that gives rise to an action potential if the summed potential is at or above the triggering level. The development of the adaptrode model will focus on the local synaptic input and subsequent changes in the synaptic response.

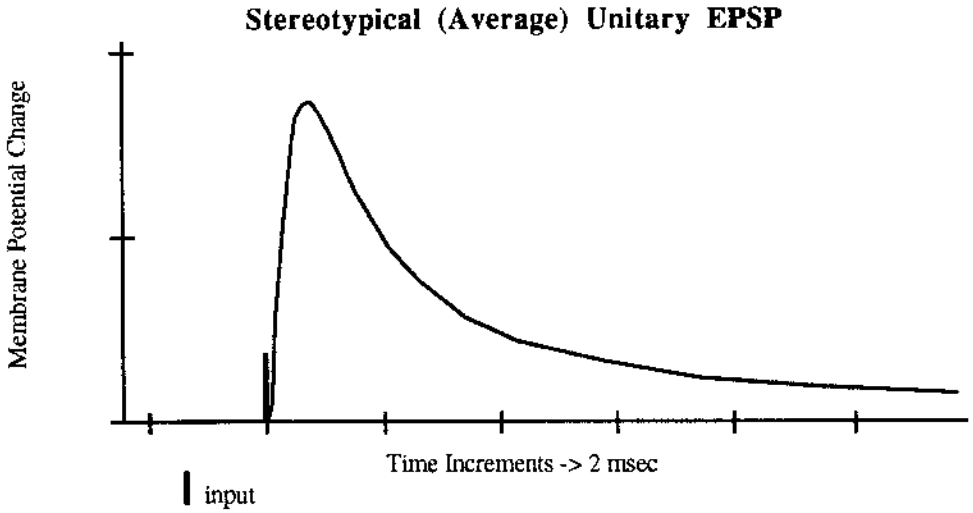

Fig. 2. The time course trace of a single-source (unitary) EPSP with arrival of a single action potential at the synapse. Scales of membrane potential, time and input intensity are approximate in this figure. (After Aidley, 1989, Chapter 9).

The time-course of the transient value of the postsynaptic membrane potential depends on several factors. The arrival of an action potential at the presynaptic bouton causes the release of neurotransmitter into the synaptic cleft. This diffuses across the cleft and combines with specific receptors on the postsynaptic membrane. In turn, these receptors cause specific ion channels to open allowing ion currents to flow due to strong gradients across the membrane. These gradients are developed by active processes in the cell and constitute the so-called resting potential of the membrane. The induction of currents cause the depolarization of the membrane (in the case of excitatory input).

The relevant details of the subsequent actions will be covered in Chapter V. For now it is sufficient to point out that the depolarization of the membrane starts from a given resting state -- nominally -80 mV with

respect to the outside of the cell -- and moves toward the triggering potential -- nominally -60 mV. If the synapse has not been excited for some extended period, then a single action potential input will not produce sufficient depolarization to reach the triggering level. If however, subsequent action potentials arrive in a short period, then a temporal integration takes place, so that the local potential is elevated and the new increment of depolarization brings it closer to the triggering level threshold.

Furthermore, after such a burst of activity, the membrane potential remains somewhat elevated, not immediately returning to the normal resting value, for many seconds. In Figure 3, several spikes have arrived over a short time. The figure shows a stereotypical EPSP resulting from such a burst. If the burst was of a long duration and/or if it is accompanied by correlated factors in the postsynaptic compartment, the level may stay elevated for relatively long periods, minutes, hours, even days. The exact time-course behavior of the postsynaptic potential varies from species to species and among cell types, however the basic properties of transient response, short-term and longer-term efficacy changes seem to be ubiquitous. In the following development the focus will be on those general features of synaptic plasticity which seem to have been conserved both across cell types within a single species and across phylogenetic lines. Alkon (1987) reports that some of the key mechanisms involved in invertebrate systems such as *Hermissenda* have been found in certain rabbit neurons. These same mechanisms are directly involved in the dynamics of plasticity.

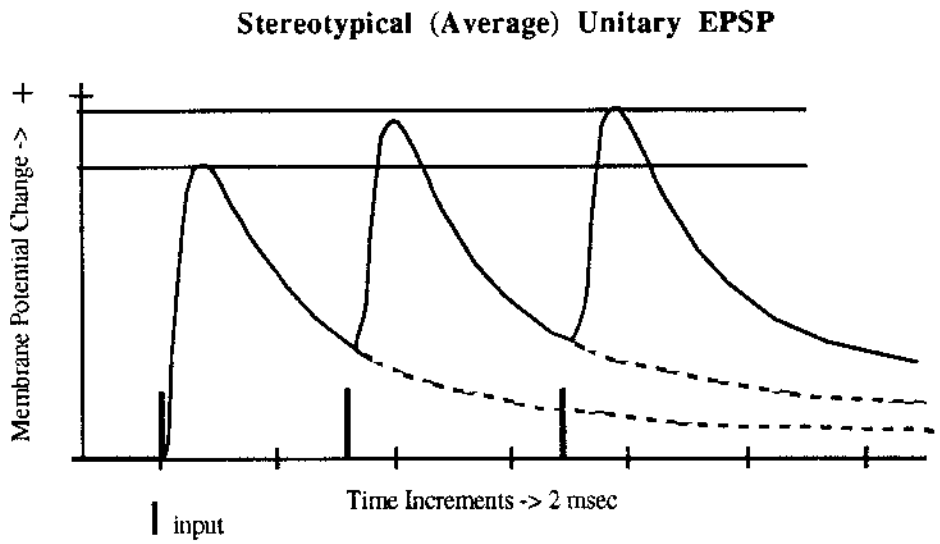

Fig. 3. The time course trace of a unitary EPSP with a series of spikes arriving over an interval of time. Dashed lines show approximate membrane potential level that would have been obtained in the absence of subsequent input.

What is a synapse doing? Its main job is to respond to an impulse signal by transiently depolarizing (or hyperpolarizing in the case of an inhibitory synapse) the local membrane. Depolarizations from a number of synapses are integrated, spatially, to produce a resultant EPSP at the neuron soma. As can be seen in Figure 3, however, the unitary response appears to depend on the immediate prior history of the signal. In the case of a sequence of relatively high frequency input pulses as shown in Figure 3, each subsequent response is marginally stronger than the prior one. In fact it is well known that if action potentials arrive in rapid succession, that the signals are integrated over time with the result that the synapse may generate a larger depolarization than can be obtained from a single action potential. Thus the synapse is acting as a kind of adaptive filter, generating a 'significant' response only if the frequency of the input signal is above

some minimum. It is assumed that the input to a synapse is a constant-width, constant-amplitude pulse-density coded signal so that high frequencies correspond to high intensity transduction from a source analog process. Such a process could be, for example, the light falling on a light-sensitive cell such as the type-B cell in *Hermissenda* (Alkon, 1987) that eventually generates a burst of action potentials, the frequency of which corresponds to the intensity of light.

Figure 3 shows the basis for the temporal integration. Dashed lines are used to depict approximately what the membrane potential trace would have been had the subsequent action potential not arrived. Due to the extended decay of the trace a subsequent arrival generates a response from a more depolarized state than would have occurred had the membrane completely returned to its resting potential. Each subsequent response starts from a more depolarized state relative to the resting potential with the result that the peak of the response is marginally greater.

The trace of a unitary EPSP might be considered a form of expectation for future signal rates. Thus, a synapse could be viewed as an adaptive filter extracting information from the message to anticipate the contents of future messages for the purpose of responding to those messages more quickly. In the design of adaptive filters in nonstationary environments one is faced with several difficult problems with respect to real-time, on-line computation of the filter response. Typically, the statistical properties of the relevant signal parameters (in this case frequency) are computed from the sample space of the signal itself when the ensemble statistic is not known. Moving window averaging is

commonly used to compute this on-line, however, the method requires the discrete storage of samples over the course of the sample window and relatively sophisticated computation with each sample. Synapses do not appear to have either the storage space (particularly for discretely stored real numbers), nor the computational machinery to compute something as sophisticated as a moving window average.

The general solution to these problems when designing certain kinds of adaptive filters is to use exponential smoothing (Sutton and Barto, 1981) which is an on-line, real-time computation that can be shown to be formally equivilent to the moving window averaging technique. A single stage averager is given by the recurrence formula:

$$w(t+1) = \alpha x(t) + (1 - \alpha)w(t) \tag{3.1}$$

where:

w	is the storage term
x	is the input term
α	is the smoothing constant, $0 \leq \alpha \leq 1$
t	is the time increment

which is a particularly easy value to compute on-line and in real-time. The adaptrode development begins with this formulation. Rearranging terms, we have:

$$w(t+1) = w(t) + \alpha x(t) - \alpha w(t) \tag{3.2}$$

In this form it is easier to see the impact of α on the result. The storage term, w, will decay as fast as it rose when x goes to zero. Figure 4 shows a graph of the formula for $\alpha = .4$. The choice of α is somewhat arbitrary but has been chosen so that the peak value of w after a single spike is less

than the trigger value of the neuron. This follows from the filtering capacity of the synapse as discussed above. In the graph the upper bound on w represents the maximum depolarization that could be achieved after a long train of spikes arriving at the maximum frequency.

The trace represented in the graph shows some similarities with that of the stereotypic EPSP. However, there are also clear differences. Note in particular that the value of w falls asymptotically toward zero and does not have the desirable slow downslope shown in Figure 2. The solution is to change the second alpha to a new term, δ, which is set to some value less than α. The result of doing so is to slow the rate of decay so that a longer trace of the storage term is maintained as in Figure 5. This effectively lengthens the trace of memory in the system.

$$w(t+1) = w(t) + \alpha x(t) - \delta w(t) \tag{3.3}$$

where:

δ is a new decay term, $0 \leq \delta \leq \alpha \leq 1$

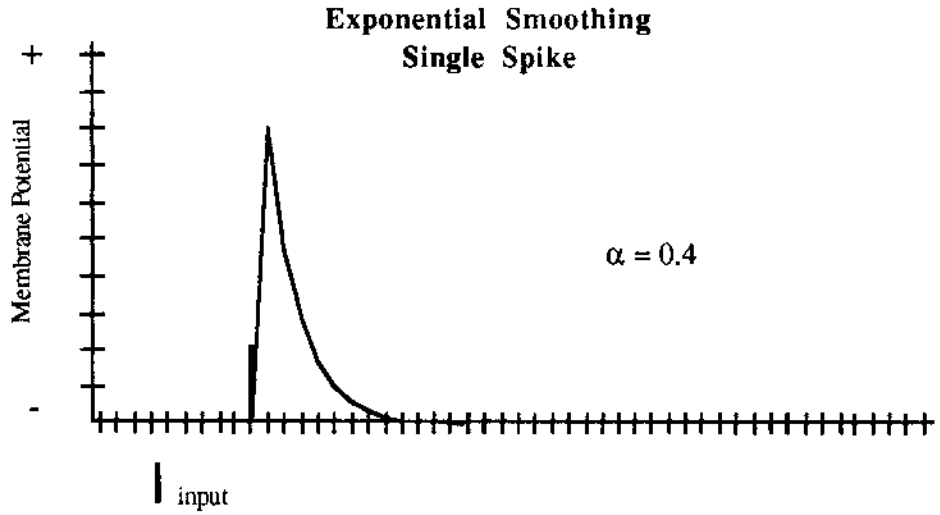

Fig. 4. Standard exponential smoothing for a single input pulse.

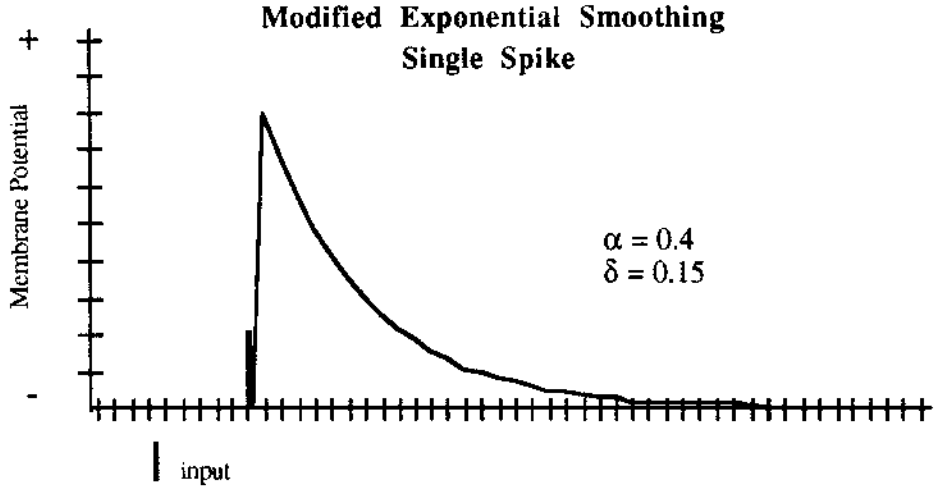

Fig. 5. Modified exponential smoothing for a single input pulse with a decay constant different from α.

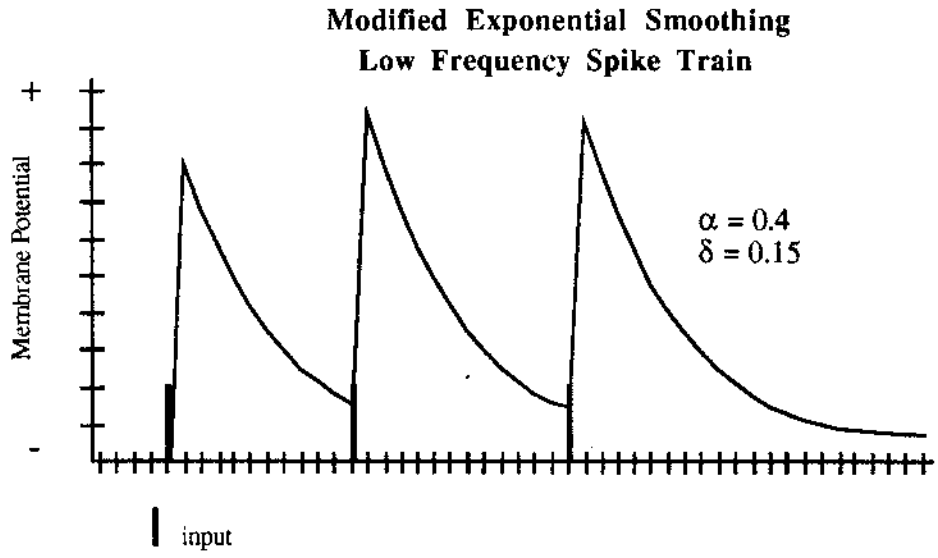

Fig. 6. Modified exponential smoothing for a sequence of input pulses.

The formulation as in (3.3) does produce the desired longer downslope but produces an undesirable result when computing the trace after a rapid sequence. The value of $w(t+1)$ is no longer bounded in the interval (0,1). This is shown in Figure 6 on the previous page.

It is clearly undesirable to allow w to increase without bound. Therefore a shunting term is introduced into the formulation.

$$w(t+1) = w(t) + \alpha x(t)(w_{max} - w(t)) - \delta w(t) \qquad (3.4)$$

where w_{max} is the upper bound on the system, here set to 1.
This has the effect of maintaining the upper bound on w, since as w approaches the boundary the term approaches zero and the effect of further inputs tend toward zero as well. Figure 7 shows the trace of w from (3.4) given a high frequency input signal.

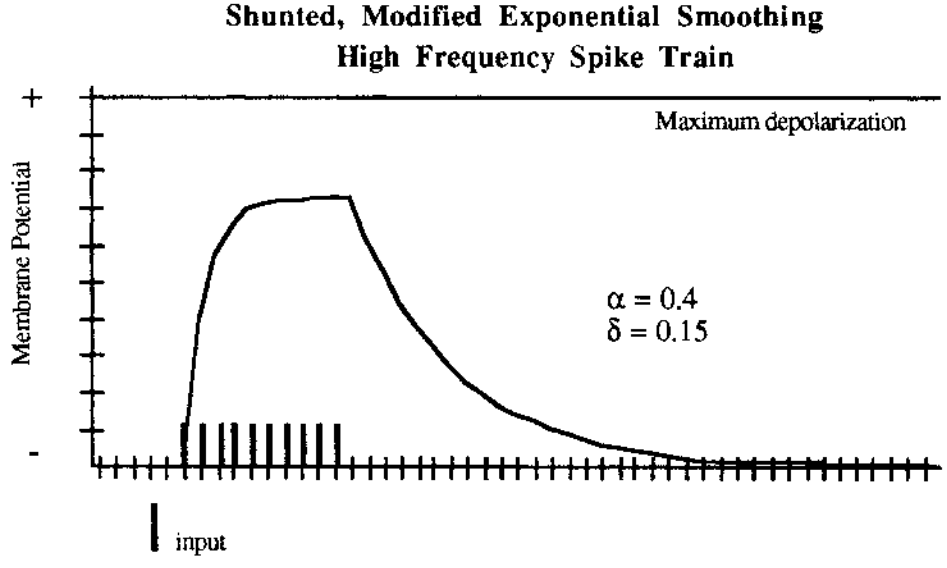

Fig. 7. Modified exponential smoothing for a burst of input pulses with shunting term.

Comparing the traces of Figure 3 to Figures 6 and 7 shows that even this modified exponential smoothing mechanism lacks a longer-term memory component. The trace in the former figure shows that the floor level or equilibrium toward which the membrane potential returns is higher than before the series of spikes was received. This phenomenon in biological synapses is called potentiation, of which there are several types (Lynch and Baudry, 1985). In fact there is a range of potentiation time domains (Alkon, 1987) from very short-term to very long-term effects dependent on the input signal and correlated secondary factors. It is this complex dynamic which can be used to explain a variety of memory-dependent behavioral phenomena (Alkon, 1987).

I extend the now-modified exponential smoothing mechanism yet further by introducing a floor factor, f, into the last term in (3.4).

$$w(t+1) = w(t) + \alpha x(t)(w_{max} - w(t)) - \delta(w(t) - f) \qquad (3.5)$$

If f is set to zero, (3.5) reduces to (3.4). On the other hand, if f takes on some value greater than zero, but less than w, then it is clear that w will be bounded from below by this factor. That is, it will not be able to decay below the value of f when x is clamped to zero. What values shall f be allowed to take on? And what should act to change those values?

Taking inspiration from biology again, f must evolve slowly compared with w, which represents the real-time memory trace. Furthermore, f, like w itself is somehow dependent on the input activity, but over a longer time scale. Could f, in fact, represent the longer-term time average of w?

Let $d \in (0, 1, 2, ..., D)$ be the index on a set of time domains chosen such that the time constants, α_d and δ_d are progressively larger as d goes from 0 to D, some upper limit. Replacing f in (3.5) with a new storage factor, w_{d+1}, the adaptrode equations are then given by:

$$w_d(t+1) = w_d(t) + \alpha_d\, x_d(t)(w_{d-1}(t) - w_d(t)) - \delta_d\,(w_d(t) - w_{d+1}(t)) \qquad (3.6)$$

where:

x_d is an input signal as described later. x_0 is called the primary input to the adaptrode corresponding to the presynaptic input.

$w_{d-1} = w_{max}$ if $d = 0$

$w_{d+1} = w_{equil}$ if $d = D$

That is, if $d = 0$, then the shunting term in the second expression uses w_{max}, a constant, in place of w_{d-1} and if $d = D$, then the floor factor in the third expression is w_{equil}, another constant less than w_{max}, in place of w_{d+1}.

The time-course behavior of storage terms, w_0 and w_1 of a two-domain adaptrode is shown in Figure 8 for the case of a single spike input. Adding the adaptive floor represented by w_1 clearly increases the duration of a trace of w_0 without significantly changing the shape of w_0's trace. Compare this figure with Figure 2 (decay curve). Clearly, and qualitatively, the former mirrors the latter.

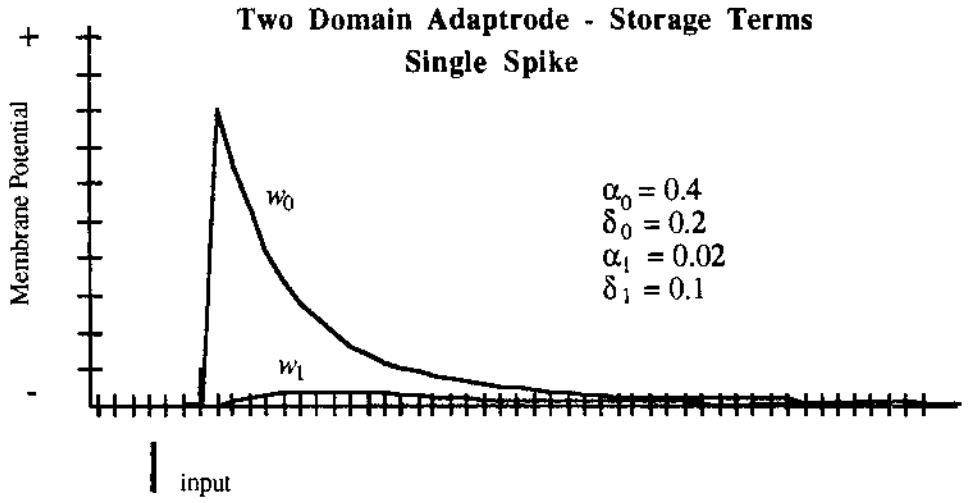

Figure 8. The memory of a two-domain adaptrode.

Similarly, Figure 9 mirrors the behavior of Figure 3 in that each subsequent peak is marginally higher than the previous peak, and the trace of depolarization is potentiated such that it will remain elevated for an extended period.

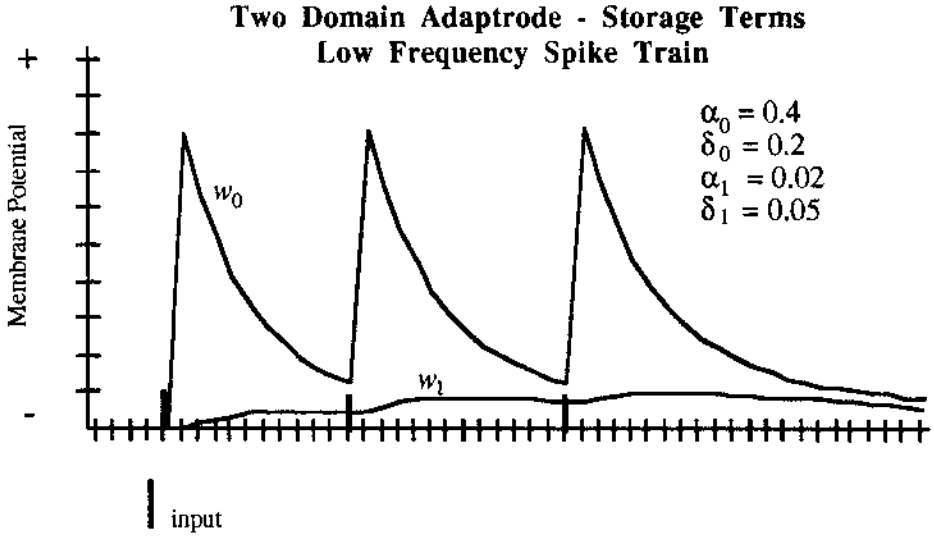

Fig. 9. Memory traces for the two-domain adaptrode receiving a low frequency burst.

3.3 Construction of a Functioning Adaptrode

Equation (3.6) provides a computationally efficient mechanism for integrating a single signal over time. Its qualitative behavior approximates that of biological synapses. With these results it is possible to construct an artificial synapse which will be embedded in an artificial neuron.

In the following (and all previous) difference equations, an increment of time has been chosen so as to represent an interval from the onset of an action potential to the onset of the next possible action potential. This choice reflects the argument, made previously, for a model between the detailed Hodgkin-Huxely type and the typical neural network type. Thus in any given Δt I represent the presence (1) or absence (0) of an action potential.

Response Unit

While the value of w_0 might be used as the response output of an adaptrode unit, for reasons of flexibility I have incorporated a separate response unit, the output of which is given by:

$$r(t+1) = \kappa \begin{cases} w_0(t), \text{ if } x(t) = 1 \\ r(t) - \delta_r r(t), \text{ otherwise} \end{cases} \tag{3.7}$$

where:

κ is a preweighting constant $\in \Re$, in general.

$x_0(t)$ is the action potential input to the adaptrode at time t, $x_0 \in \{1,0\}$.

δ_r is a decay constant for the response.

The constant, κ, can be any negative or positive real number which can be used to: 1) preweight the synapse to model electrotonic distance from the neuron soma; 2) if positive, represent an excitatory synapse or if negative, represent an inhibitory one. In the simulations reported in the next chapter, κ is either -1 or +1. Electrotonic distance refers to passive electrophysical characteristics of the cellular membrane that are affected by the physical distance from the cell soma (body) of a patch of membrane - in this case a post-synaptic patch.

Equation (3.7) describes a response that uses the trace of (3.6) where $d = 0$, if an action potential has arrived at the synapse. The adaptrode will produce an excitatory or inhibitory contribution (depending on κ) with each action potential. Figure 10 is a graph of the response, r, along with the storage terms, w_0 and w_1, as in Figure 9. As can be seen in the figure, r can decay at an independent rate from w_0 and generate no contribution during periods of quiescence.

Associative Encoding

The adaptrode as given above, computes a temporal, multi-resolution average of the input signal as a basis for predicting the appropriate response level. That is it computes a form of autocorrelation over increasing time scales as a basis for response. Associative encoding addresses the issue of cross-correlation among multiple signals.

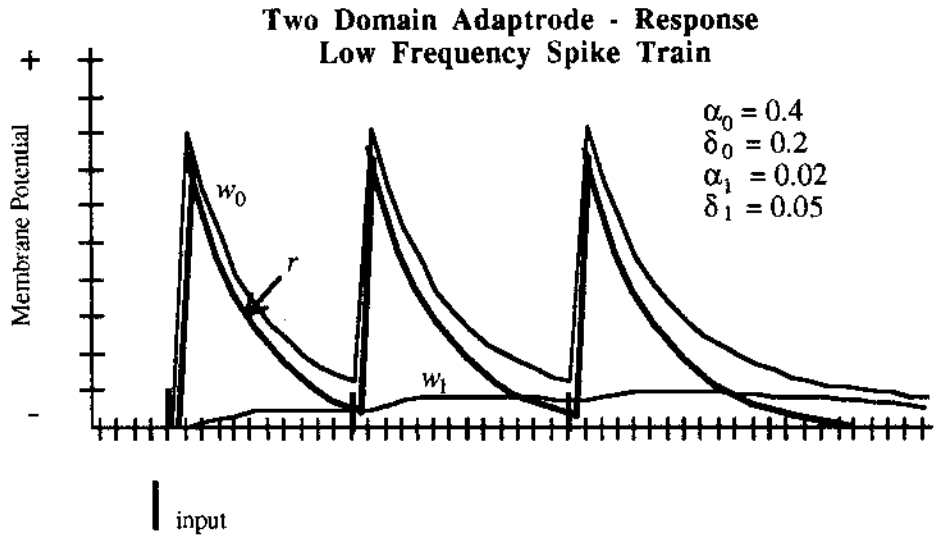

Fig. 10. A two-domain adaptrode with a response unit producing trace *r*.

Associative encoding is accomplished in a rather straightforward manner. The principal idea is to introduce a threshold gate in the coupling between any two time domain levels, d and d+1. Given a gate, γ_d ($d \in \{1, 2, 3, ..., D\}$, is an index of the level), and $\gamma_d \in \Re$, let h_d denote the sum of response signals, r_{ij}, from a proscribed set $H_d = \{r_{ij} \mid i \in \{1, 2, 3, ...N\}$, the set of neurons in the network, including sensor inputs and $j \in \{0, 1, 2, ..., n_i\}$ the set of adaptrodes in neuron $i\}$. That is:

$$h_d = \sum_{H_d} r_{ij} \tag{3.8}$$

The set H_d is given in its most general form where the response signal from any adaptrode (including the current one) can be designated in the set. For simplicity (as well as in practice), I will assume a single signal source rather than a set.

A particular temporal ordering constraint is imposed by the nature of associative learning phenomena such as classical and operant conditioning as covered in Chapter II. That is, the order of arrival of signals must be strictly enforced, where the arrival of the hurdling signal must come after the arrival of the primary excitation, x_0, of the learning adaptrode. The former takes on the role of a reinforcement signal in the terminology of conditioning. The current response, $r_k(t)$, of the adaptrode is used to gate the computation of x_d. If r_k exceeds a threshold, ρ_d, before the summation of the hurdle sources exceeds γ_d, then the computation is performed. Otherwise it is locked out until the circuit resets. Mathematically for the input variable, x_d, for a specific adaptrode, k, of Equation (3.6) can now be defined as:

$$x_{dk}(t) = \begin{cases} 1, \text{ if } r_k(t\text{-}1) > \rho_d \text{ and } h_{dk}(t) > \gamma_{dk} \\ 0, \text{ otherwise} \end{cases} \tag{3.9}$$

where:

d indexes the time domain as above

k is a new index for the current adaptrode in a population of adaptrodes.

ρ_d is a constant threshold to be described below.

If the response, r, of the adaptrode is greater than a given constant, ρ, and the sum of the associated signals, h, is greater than the gate, g, then x_d, is set to 1 and for any $d > 0$ in Equation (3.6), w_d will be increased iff, $w_{d\text{-}1}$ is greater than w_d by the equation. By setting ρ appropriately, one can control the time required to pass before the w_d level is permitted to potentiate. As with Equation (3.8), (3.9) is given in its most general form.

However, in practice, again, I have used this mechanism only in the w_1 level with good results (Chapter IV).

3.4 Artificial Neurons

At this point we are ready to construct a working artificial neuron that emulates some of the important computational properties of real neurons. The unit constructed here may be viewed as a very simple complete neuron as might be the case in the nervous system of an invertebrate, or it can be be construed as a compartment model for constructing much more complicated neurons such as pyramidal neurons of the mammalian cortex. This is quite different from the classical neural network caveat that "units" may represent individual neurons or *groups* of neurons.

The conventional formal neuron has a set of input edge weights which are representative of the efficacy of a synapse in contributing to the generation of an action potential by the neuron. Model neurons based on adaptrodes are not much different except that the single weight (per input) is replaced by an adaptrode with its internal set of weight signals and its external response. It is the latter which constitutes the input to the spatial integration process leading to an overall activation value for the neuron. Integration is performed once in each time step, the same as the input sample rate, and the resulting activation is compared to a threshold. If the activation exceeds the threshold, then the neuron fires an action potential, otherwise it does not. It is this clocking at the maximum frequency rate

which squashes the output signals of neurons to be in the same range as the input signals at the synapses.

Figure 11 compares a "conventional" artificial neuron (a) with an adaptrode-based neuron (b). Focusing on the adaptrode-based figure, the response signals from adaptrodes, labeled A_1 through A_n, are summed and compared to the threshold, θ, as shown. The output signal, labeled y , will be a 1 (one) if the threshold is exceeded. In Figure 11b the output, y, serves a second purpose - as input to a special adaptrode labeled A_0. This adaptrode may be optionally used for several purposes. One important use is to compute a variable threshold based on output signal. This amounts to an activity-dependent learning process at the neuronal level which will be demonstrated more fully in Chapter IV.

A second use is to provide a graded response output from the neuron. This signal may be used to establish cross-neuronal associations when the output of one neuron can act as a hurdle source for adaptrodes in another neuron.

One major advantage gained in constructing neurons with adaptrodes is the ability to build a wide variety of neuronal types. This is possible because different types of synapses, with different dynamical properties, can be built by designating specific numbers of levels and values of α_d and δ_d for all levels d in that type of adaptrode. This is an advantage from the standpoint of flexibility in creating models that emulate real biological neurons. It is, however, problematic from the standpoint of specification. There is, as yet, no theory that would guide the selection of these parameters and it may turn out that they be selected only empirically.

Comparison of Conventional Formal Neuron with Adaptrode-based Neuron

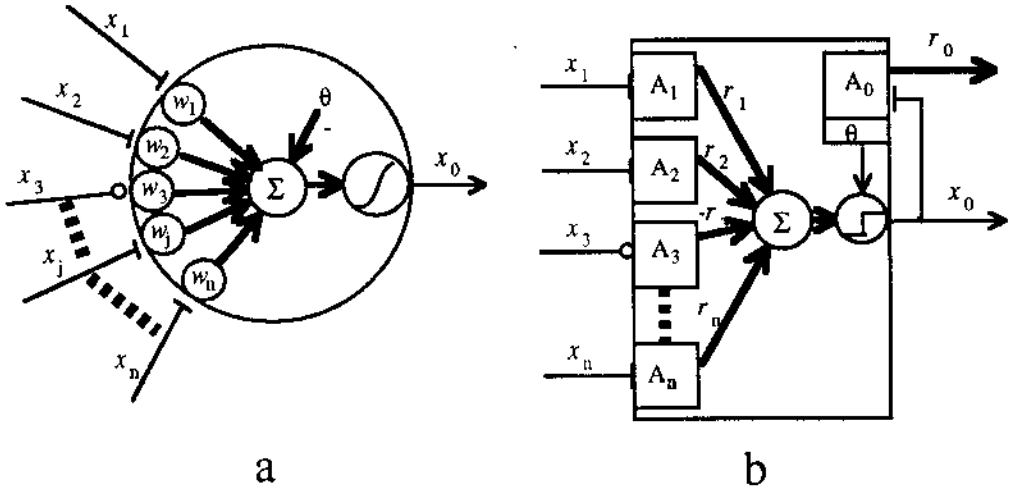

Fig. 11. Conventional formal neuron (a). Adaptrode-based neuron (b). Thin arrows with flat termini represent excitatory inputs, those with circular termini are inhibitory. Thick arrows represent weighted input (a) or response (b) values. The sigma processors (Σ) sum the graded values and produce an activation value which is fed into either a squashing function (a) or a simple threshold comparator (b). A unique aspect of the adaptrode-based neuron is the use of an adaptrode to monitor the output of the neuron. The response of the adaptrode (A_0) can be used to modify the threshold (θ).

To formalize the above description, let *a* be the activation of the neuron from the sum of current responses of all but "silent" adaptrodes. That is:

$$a(t) = \sum_{j=1}^{n} \sigma_j r_j(t) \tag{3.10}$$

where *n* is the number of adaptrodes in the neuron and $\sigma_j \in (0,1)$ is set for each adaptrode that is not silent. An example of a silent adaptrode, typically, is the output adaptrode.

Then output, $x_0(t)$, of the neuron is defined to be:

$$x_0(t) = \begin{cases} 1, \text{ if } a(t) > \theta \\ 0, \text{ otherwise} \end{cases} \tag{3.11}$$

Recalling that each $\Delta t = (t+1) - t$, is the the time interval of an action potential plus refractory period, the interpretation of (3.11) is that an action potential either fired, or it did not, during the time increment depending on the summed activation (EPSP) of the unit.

3.5 Concluding Remarks

The basic model of a biologically-realistic artificial neuron has been presented in this chapter. The basis of the claim for biological realism is that the adaptrode synapses show dynamical behavior in consonance with what is currently known about real synapses with respect to the post-synaptic potential. The result is that the spatial integration of synaptic contributions and the decision to fire an action potential in any given Δt is much more like a biological neuron than current models derived from the McCulloch-Pitts formal neuron.

As mentioned in Chapter I, one of the motivating principles followed in the development of the model has been to retain the computational tractability of more conventional ANNs while increasing the overall power of the artificial neuron. Clearly the equations presented in this chapter introduce no significant problems with respect to this goal. Typically, the computational complexity of ANNs is given in terms of the number, n, of connections or synapses that must be processed and in the majority of cases

this is found to be linear in n (Hecht-Nielsen, 1986). The adaptrode formulation increases the computational complexity of a neural network by no more than a constant amount.

In this chapter the adaptrode was developed from consideration of the synapse as an adaptive filter. That is, only the input/output characteristics of a biological synapse were used to infer an internal mechanism by which adaptive response might be achieved in a computationally tractable way. In the next chapter, the adaptrode formulation is compared structurally with aspects of the post-synaptic compartment that give rise to the multi-time scale memory trace processing given in Alkon (1987). As will be seen, there is a striking correspondence between the model and the biology which lends weight to the argument that adaptrode processing captures the essential qualities of biological synaptic plasticity.

CHAPTER IV

CORRESPONDENCE OF THE ADAPTRODE WITH BIOLOGY

4.1 The Biological Basis for an Adaptrode-like Model

The adaptrode model was developed in Chapter III from a consideration of the signal filtering role of synapses in biological neurons. In this chapter I will examine the correspondence between the adaptrode model and what I believe to be the important aspects of the neurophysiology of biological synapses that explains adaptive response.

The plan of this chapter is to give a summary of the Alkon model and then to compare it with the adaptrode in terms of structural correspondence relating to the abstract view presented in Chapter 2 (Figure 12). Limitations of the adaptrode in emulating a biological synapse will also be examined.

As shown in Chapter III, the adaptrode's dynamics emulate, at least qualitatively, the time-course behavior of the post-synaptic membrane reactivity. Alkon (1987) has developed a model of synaptic efficacy which depends not only on real-time modulation of the post-synaptic membrane patch (specifically the conductance of potassium ion channels), but also on intermediate-term and long-term molecular processes which operate deeper in the compartment cytosol. Changes in protein and mRNA synthesis are implicated and nuclear processes such as DNA activation have been suggested as well.

The general model is summarized in Figure 12 and described here. The arrival of an action potential at a synaptic bouton (x_{10} and x_{20} in the figure) initiates a rapid depolarization of the post-synaptic membrane. The nature of a unitary excitatory (inhibitory) post-synaptic potential, or EPSP (IPSP), was reviewed in Chapter III. As the EPSP increases rapidly, somewhat slower acting processes, such as the sodium (Na^+) ion pump, work to restore the membrane polarization to its resting state. As has been seen in the trace of the EPSP in Chapter III, this decrement process is negative exponential.

Heavy black arrows in the figure represent slow decrement processes, while the thin, solid arrows represent fast forward driving processes. Dashed arrows represent feedback loops that down-modulate decrement rates. Up-pointing open arrows represent accumulation or state variables such as the EPSP or calcium concentration in the cytosol ($[Ca^{2+}]$). Throughout the diagram as and ds represent kinetic rate constants that will be corresponded with the adaptrode model below. Specific references to neurophysiological components are made for the purpose of letting the interested reader be guided in referring to Alkon (1987). No attempt will be made here to illucidate the details of this model.

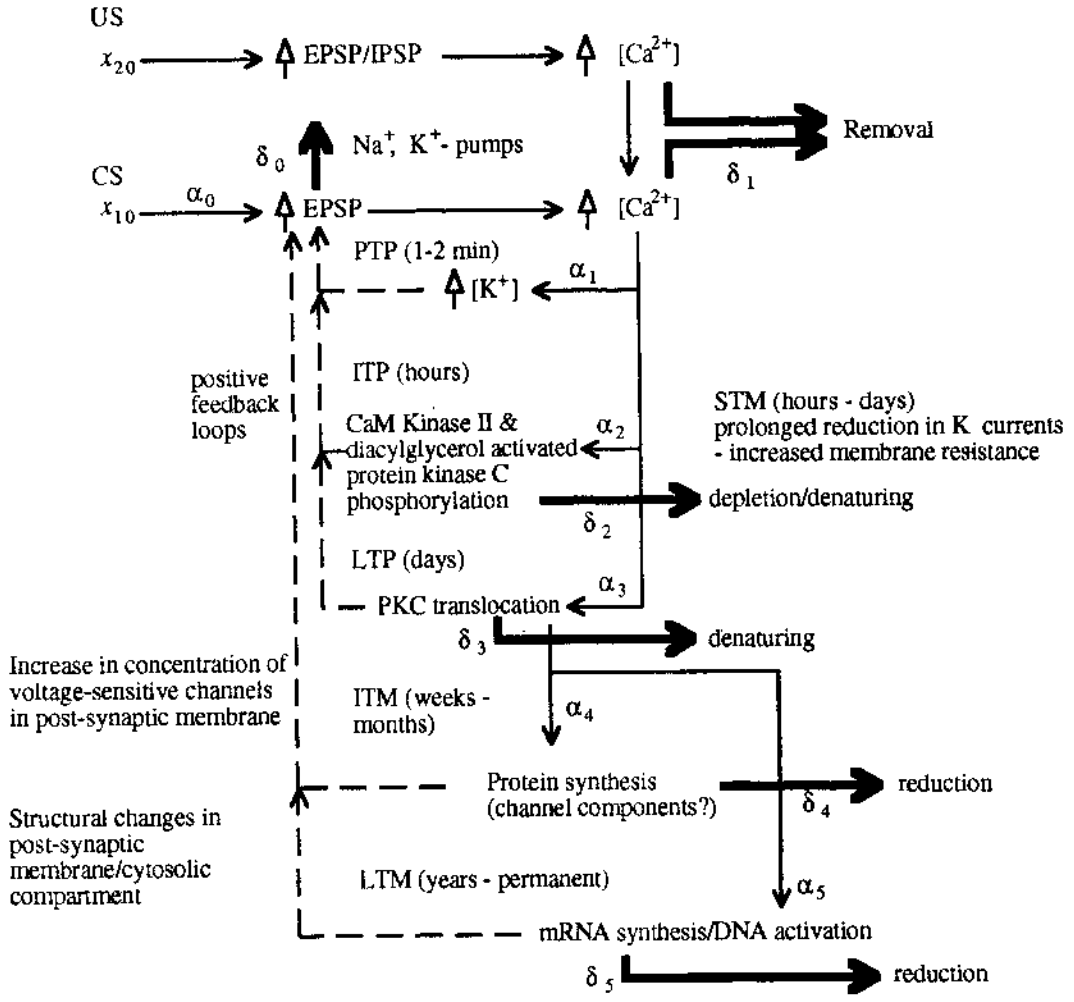

Fig. 12. A summary of the major biophysical processes that transpire in the post-synaptic cytosolic compartment. See the text for an explanation of the elements of the diagram.

Following the track of events from the CS input (x_{10}), through its EPSP, an arrow representing a forward, fast reaction points to the accumulation of calcium ions in the post-synaptic cytosol. Calcium accumulates due to the opening of ion-specific channels that allows the Ca^{2+} ions to enter the cell following a gradient. The rate of accumulation is

dependent on the magnitude of the EPSP. Over a slightly longer time scale (δ_1 in the diagram) Ca^{2+} is removed from the cytosol by a number of mechanisms that work to restore the cytosolic concentration to pre-excitation levels. It is important to note that these Ca^{2+} removal processes operate over a slightly longer time scale than the possible excitation rate of the EPSP. If several action potentials arrive in rapid succession, the calcium channels will be reactivated and there will be a net greater accumulation of Ca^{2+}. Thus, Ca is acting as an excitation integrator (leaky) providing a rough approximation of a time averaged excitation.

The accumulation of Ca^{2+} in the compartment triggers another process, the calcium-mediated phosphorylation of certain proteins associated with channels which provide for a potassium ion (K^+) outflow (counter to the Na^+ and Ca^{2+} inflows). When K^+ outflow in response to depolarization is so down-modulated, that is K^+ concentrations remain elevated in the cytosol, the membrane potential is kept more positive with respect to the exterior. Effectively the post-synaptic membrane is left in a state of polarization that is closer to the triggering potential. Through this feedback loop, any subsequent EPSP starts from a higher level. This is thought to be the basis for the temporal integration phenomenon reported in Chapter III.

The time course for the EPSP as has been mentioned in on the order of milliseconds (time constants in the microsecond range). In the event of a rapid series of action potentials, a tetanic input, the local potential remains elevated due to the above mechanism for several seconds to minutes (PTP or post-tetanic potential). Thus the general form of the

dynamics is captured in this one loop. A fast activating process with a somewhat slower recovery rate drives the slower accumulation of a variable such as the Ca concentration. This, in turn drives a yet slower process (Ca^{2+} mediated phosphorylation). All of these processes have recovery rates that are longer than the activation rates. Eventually, the accumulation of one of these variables provides a unique kind of positive feedback to the original driving process. The feedback diminishes the decay rate of the fast driving process. Yet due to the decrement of the intermediate variable (e.g., removal of Ca^{2+} from the cytosol) this does not, ordinarily lead to a runaway process.

The accumulation of Ca^{2+} triggers other processes shown in the diagram of Figure 12. These processes are characterized as having a yet longer time course (order of hours) and lead to an intermediate-term potentiation (ITP) of the post-synaptic membrane. If the Ca^{2+} concentration remains high enough for a long enough period of time, as would happen with either strong tetanic input or weaker, but sustained input, or, as I will discuss below, associated sources from cytosolic inputs, still slower acting and longer lasting processes accrue leading to a long-term potentiation (LTP) of the post-synaptic membrane. These processes, involving the translocation of protein kinase-C from the cytosol to the membrane where its impact on additional protein phosphorylation is more efficacious, is thought to trigger (or drive) yet longer term processes such as the synthesis of protein, perhaps increasing the availability of new channel components for insertion into the membrane. Alkon (1987) reports evidence of increased messenger RNA synthesis and possibly DNA

activation that could lead to specific structural changes in the post-synaptic membrane and compartment. Such morphological changes after learning have been noted.

The model presented so far might well be common to both associative and nonassociative learning mechanisms. Associativity requires the convergence of signals from two or more sources. Alkon (1987) points to the Ca^{2+} concentration in the post-synaptic compartment as a possible candidate (there are, in fact, many other candidates and in all likelihood many of them may operate in the same system - for the purposes of this discussion it is only necessary to focus on one such candidate). If the removal rate of calcium ions is increased by the presence of additional mechanisms (calmodulin, etc.) then a synaptic compartment might not undergo the kinds of processes outlined above without extra influxes of Ca^{2+}. Such sources might be neighboring synapses. In Figure 12 an input at US (x_{20}) would initiate the same kinds of reactions in the post-synaptic membrane. If the synapse type (and there are numerous types) is such that it accumulates Ca^{2+}, and that Ca^{2+} diffuses to its neighbors, then such a synapse would provide a correlated source of Ca^{2+} that could participate in initiating the sequence of events described above.

4.2 Correspondence of the Adaptrode Model

The Alkon model describes a time-domain cascade of processes (some of which may operate in parallel) which, collectively, modulates the post-synaptic membrane reactivity. A set of processes which have similar time-course behavior operate together, defining the dynamics of a characteristic time domain. For example, processes of neurotransmitter- and voltage-activated ion channel openings, triggered by the arrival of an action potential at the pre-synaptic bouton, act together to reduce the membrane potential (in the case of excitatory synapses). These forward-driving processes are followed and countered by the somewhat slower processes of active ion pumps working against the ion concentration gradients to restore the resting potential of the membrane. As a group, they operate in the microsecond time constant range with the time-course trajectory of post-synaptic potential in the millisecond scale. The adaptrode emulates this set by collapsing several component variables into a single time-domain variable, the w_0 level, where an action potential, x_0, triggers the depolarization, $\alpha_0(w_{max} - w_0)$, offset by the slower recovery, $\delta_0(w_0 - w_1)$, due to the actions of the sodium and potassium pumps. This last expression, $(w_0 - w_1)$ provides the feedback coupling. The Ca^{2+} concentration, as qualitatively represented in the adaptrode, would be at the w_1 level. The forward drive, and its intrinsic rate, is coupled with the degree of depolarization of the membrane by $\alpha_1(w_0 - w_1)$. Similarly, the removal of Ca^{2+} , a much slower process than the action of ion pumps mentioned previously, is approximated as $\delta_1(w_1 - w_2)$.

Which brings us back to the feedback coupling noted above. In the Alkon model, the presence of Ca^{2+} in the cytosol has an immediate damping effect on the potassium ion, K^+, outflow. The effect, measured as the post-tetanic potentiation (PTP), is highly transitory, lasting only a few hundred milliseconds to perhaps several minutes depending on the preparation being studied. This marginally reduced potential, which is shown to have exponential decay, is a trace recording which is dependent on the intensity and/or duration of the input signal. As Alkon (1987) points out this translates into an increased, if transient, efficacy for the synapse. Hence, in the adaptrode, we make the decay of w_0 (depolarization state) dependent not on just its own magnitude but on the difference between its magnitude and that of w_1 (Ca^{2+} concentration).

The w_2 level in the adaptrode might be considered as a qualitative model of the phosphorylation of proteins, presumably components of the K^+ channels or channel blockers that constitutes another, longer time domain. This activity is driven by the concentration of Ca^{2+} and is reflected in the adaptrode model in the form of $\alpha_2(w_1 - w_2)$. Though the restoration mechanism for the K^+ channels was not directly addressed in Alkon (1987), I suspect that some process restores them to their former conductivity in the absence of Ca^{2+}. I believe it is reasonable to expect this restoration to follow an essentially exponential form.

Finally, the gate hurdling mechanism in the adaptrode fills the role of the cytosolic convergence site for correlated signals. An interesting phenomenon in neural potentials involves what is known as a reversal potential. If the post-synaptic membrane is already too positive with

respect to the exterior, instead of a strong EPSP, excitation, the membrane undergoes a repolarizing shift. The characteristic potential at which this occurs is termed the reversal potential. It is intriguing to see that if a cytosolic process, such as the diffusion of Ca^{2+} into the compartment occurs prior to the input of an action potential actually prevents, or at least diminishes the possibility of the action potential setting off the chain of events outlined above. As I have shown in both in the derivation of the adaptrode model in Chapter III and in the simulations of Chapter IV, the hurdle mechanism, as in Equation (3.10), specifically the role of the ρ constant, acts to prevent potentiation from occurring in the event that the hurdle signal (from the US input) arrives before the input from the CS signal.

4.3 Limitations of the Model with Respect to Biology

The adaptrode as currently given is a linear, deterministic device. There are known to be a number of nonlinear processes at work in the feedback mechanisms of real synapses. Therefore, as an emulation of biological learning, the adaptrode is likely to be, at best, a first approximation.

Another point to be reemphasized is that the adaptrode is not meant to be a mathematical model of the synapse. The time steps used in adaptrode processing are far to course to represent the fine structure of synaptic dynamics which would ordinarily be represented by a system of coupled, ordinary differential equations. It may be possible to build approximate models of neural circuits with adaptrode-based neurons that,

to some degree, demonstrate the large-scale features of network dynamics in living systems. However, the adaptrode, though derived from the multi-time domain model of those dynamics can in no way describe the level of detail that a neurophysiologist would require.

In terms of the behavior-producing capacity of the model, as will be shown in the next chapter, several key features of classical conditioning in simple invertebrates are found in the simulated adaptrode-based neuron. This does not translate directly into the production of animal-like behavior (animatics) in a fully embodied robot, but it is certainly suggestive of the possibility.

4.4 Concluding Remarks

In this chapter it has been shown that the adaptrode model, to a first approximation, corresponds quite well with now known or implicated biological mechanisms in post-synaptic placticity. This aspect, taken together with the results of the simulations reported next is compelling evidence for giving this model serious consideration as an alternative to simple weighted input connections in neural networks.

CHAPTER V

CONDITIONED RESPONSE LEARNING - SIMULATION RESULTS

5.1 Classical Conditioning as an Experimental Model

The animal learning paradigm of classical conditioning is considered to be a good starting point for the study of learning phenomena, both in neurobiology, to link behavior to neural substrates, and in artificial neural systems, to show performance (Alkon, 1987; Alkon, 1989; Alkon, et al., 1990; Alkon, et al., 1991; Buonomano, et al., 1990; Byrne & Gingrich, 1989; Byrne, et al., 1990; Gingrich & Byrne, 1987; Klopf, 1988; Grossberg, 1991b; Klopf & Morgan, 1990; Moore & Blazis, 1989; Morgan, et al., 1990; Staddon, 1993; Sutton, 1988; Sutton & Bartow, 1981, 1987, 1991; Tesauro, 1986). In this chapter the adaptive capabilities of adaptrode-based neurons in the framework of the empirical work on classical conditioning will be demonstrated. This framework has a rich history which provides a widely accepted testbed for neural learning algorithms and neural network architectures (see esp. Klopf, 1988).

Traditionally, researchers have demonstrated how their models match the animal performance data reported in the animal learning literature. It is customary in this framework to show that a model can account for basic conditioning phenomena and then show that the model performance measures approximate, in the sense of curve fitting, the animal performance data. Usually the modeler attempts to show how their

model addresses some more subtle aspects of learning and memory behavior which extends the usefulness of the model. This chapter will follow that formula. The first task is to show how the performance of an adaptrode-based neuron matches that of classical conditioning in animals, in particular, invertebrates. Then it will be shown how this model addresses the problem of contrary association learning without memory washing out. This demonstration advances the capabilities of artificial neural systems further into the domain of animal-like learning.

In many of the models where artificial neural systems have shown some of the characteristics of classical conditioning there have been several difficulties with respect to the setup and interpretation of the data. It is frequently the case that the learning regimens or 'protocols' that are used in the model are highly abstract versions of those employed in actual animal studies. It must be remembered that the networks that are being simulated are not representing whole brains or whole animals. They do not, for example, live out or experience inter-trial periods as an actual animal does. Therefore it is hard to asses the role of forgetting or consolidation in these models. Also, the measurements taken from the simulation are often not necessarily those that are related to whole-animal performance in terms of evoked behavior. For example, Klopf (1988), who helped to pioneer this methodology, took the weight values of his drive-reinforcement neurons as a measure of memory acquisition over a sequence of trials. Similarly, Sutton and Barto (1981, 1987) used the connection weight as an analogue of performance. Both of these approaches assume that the connection weight is equivalent to what has

been learned. Yet what animal learning researchers measure is not synaptic efficacy, but the propensity (i.e., probability) for an animal to perform a specified behavior after receiving the appropriate stimulus. Whole animals, particularly mammals, have very complex networks that mediate behavior and memory. It is completely unclear how the measured performance of such a complex animal, even on simple protocols, is directly correlated with specific memory encoding sites within its brain.

Finally, none of the models mentioned above address the role of motivation, neuromodulator effects, or other complicating factors in present in real animal studies. For example, animals must generally be hungry in order for a reward-based (food) protocol to show the unconditioned stimulus/response required for conditioning. The adaptrode model to be presented below does not address this issue directly. However, the means by which associative traces converge in the hurdle computation between time domain in an adaptrode offer a mechanism for use in exploring such issues in the future.

For these reasons, the simulatións that are described below are related to conditioning of very primitive animals - invertebrates - such as sea slugs (Alkon, 1987 - *Hermissenda*) and snails (Gelperin, et al., 1985 - *Limax*). As was discussed in Chapter IV, Alkon (1987) and others have developed considerable evidence for the correlation between memory traces in the primitive brains of such animals and the behavior of the animals in classical conditioning trials (Kandel & Schwartz, 1982; Small, Kandel, & Hawkins, 1989). In the simulations reported here both trace phenomena and performance (i.e., network output) will be shown.

In addition to staying with animal models where memory trace phenomena are better understood and correlated with performance, the simulations below attempt to more closely approximate the real-time quality of actual training protocols used for these animals. They include long inter-trial periods in which the network is allowed to 'equilibrate' or experience purely (pseudo) random input fluctuations as would be the case for a live animal.

5.2 Local-Interaction vs. Hebbian Learning

The adaptrode mechanism produces a memory trace of a signal either as an uncorrelated time average or, in combination with a correlated hurdle signal, as an associative time average. This addresses the dynamics of signal encoding (trace formation) but does not speak directly to any specific learning rule. In this work I will focus on associative learning since it is a requisite for the conditioning paradigm. In that arena there are a number of correlation learning rules that have been proposed. It is possible, with adaptrodes, to construct neurons which emulate several of those rules. I will use two reinforcement models which relate to known biological mechanisms. One is a form of the Hebb postulate (1949) in which neuron output as recorded by the output adaptrode is used as feedback to the intermeditate-term memory trace to provide the potentiating hurdle signal in the learning synapse. The other is related to the local-interaction postulate of Alkon, et. al. (1991) in which the unconditioned stimulus input is directly used to potentiate the learning synapse.

Figures 13 and 14 show the schematic diagrams of these two configurations. In the case of the Hebbian neuron (Figure 13), the response output from adaptrode A_0 is used as the hurdle signal for level w_1 in the learning adaptrode (A_2). The input to this latter adaptrode is labeled CS to indicate that it is the conditionable stimulus that will come to be associated with the US or unconditionable stimulus, arriving through adaptrode A_1. For the Hebbian neuron this occurs when the neuron's output is sufficient to cause the output adaptrode to become potentiated. The latter is set so that the response will 'follow', that is use the same value as, the w_0 value in each time step.

For the local-interaction (LI) rule, the A_0 adaptrode is not used. The response from adaptrode A_1, the US input, is used as the hurdle signal to adaptrode A_2.

In both figures an additional input, $x_3(t)$, to adaptrode A_3 is shown. This adaptrode will be used in network simulations where cross inhibition will be employed. Also, the response output of adaptrode A_0 is shown as input to the threshold processor, a small rectangle labeled θ. This indicates the capacity for the output adaptrode to be used to modify the threshold of the neuron if desired. This too, will be used later.

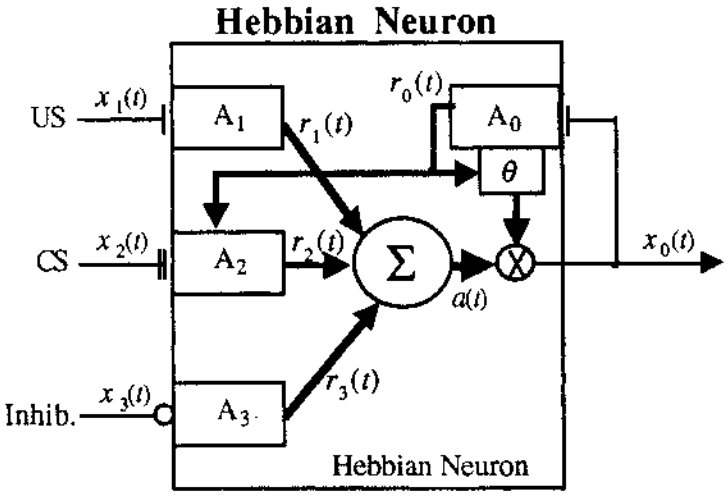

Fig. 13. The response of adaptrode A_0 is used as a hurdle signal to the w_1 level of the learning adaptrode A_2. This arrangement constitutes a Hebbian-type feedback from the output so that correlations captured by the A_2 adaptrode are between the input at x_2 and the output of the neuron.

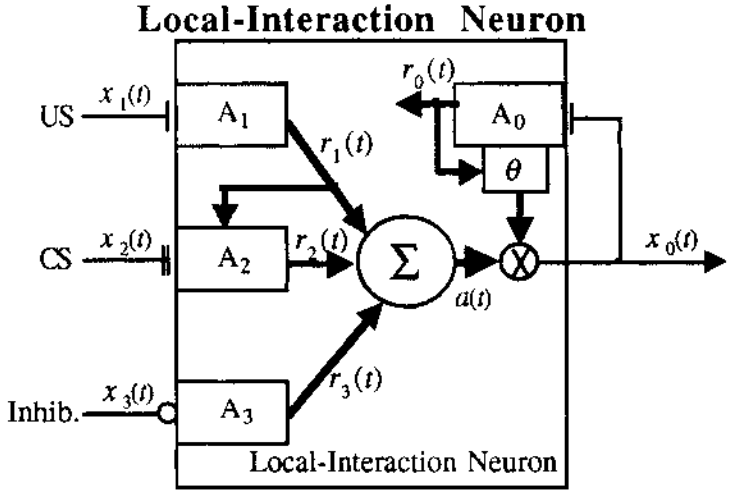

Fig. 14. In the local interaction configuration, the response of the US receiving adaptrode, A_1, is used as the hurdle source for the learning adaptrode. Thus the correlation captured is between two inputs.

5.3 Input-Output Relations - Experimental Protocols

An experimental protocol is defined along several dimensions. The first involves the phase relation between the offset of the CS signal and the onset of the US signal. If the US onset occurs before the offset of the CS the conditioning method is called *delay* conditioning. Alternatively, if the US starts sometime after the offset of the CS then the method is called *trace* conditioning. In the latter instance the focus is on exploring the efficacy of a memory trace, (Hull, 1952). Figure 15 shows the phasic and amplitude relationships which are used in the following simulations. Arrows indicate the onset of the signal. It is customary to show these as step functions and, indeed the interpretation is that the sensory source for these signals is either present or absent. In animal experiments such as the classical Pavlovian work (Pavlov, 1927) with dogs, the CS is the non-semantic (that is, meaningless, *a priori*) stimulus such as the ringing of a bell sometime prior to the presentation of food, which is the US or reinforcer (having semantic quality). In delay conditioning, the bell would still be ringing when the food is presented, while in trace conditioning, the bell would be silent for some, usually short, period prior to food presentation.

The CS signal for simple invertebrates such as snails could be the presence of light (Alkon, 1987) to which the animal is mildly attracted. Alkon showed that the snail learns through a set of conditioning trials on shaking (the US) to avoid the light or contract its foot in a characteristic 'clinging' mechanism (the UR/CR). He conjectured that this is the type of response that would benefit the snail in its natural marine habitat where its food source (algae) would ordinarily be associated with the presence of

light which is near the surface. If, however, there is wave turbulence near the surface then the snail might get knocked off the algae and fall to the bottom, possibly becoming food for some other creature.

Studies of the gill withdrawal reflex (UR) in another snail, *Aplysia* (Carew, Walters & Kandel, 1981), and taste aversion response in *Limax* (Gelperin, Hopfield, & Tank, 1985), have shown similar results in demonstrating classical conditioning in animals possessing relatively simple nervous systems.

Delay and Trace Conditioning Protocols

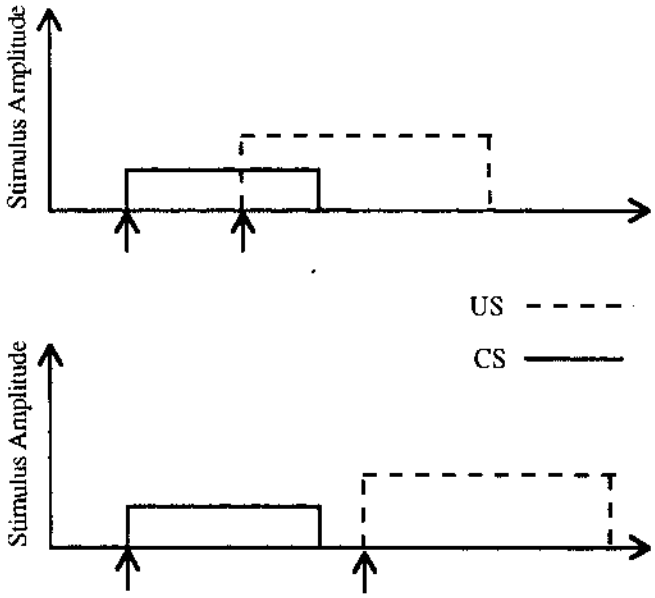

Fig. 15. Signal traces for the CS and US signals under the two protocols, delay (upper) and trace (lower) conditioning.

Within either of, trace or delay conditioning, protocols several variations on specified parameters are used to test for the signatures of conditioning phenomena. The experiments are conducted so as to measure the acquisition of the CR, conditioned response, which is the same as or similar to the UR or unconditioned response in terms of performance (amplitude) and overt behavior (shape). The UR is the output behavior evoked by the US alone. Several in-protocol variants have been found to reveal specific characteristics of classical conditioning. Experimenters will typically vary one of these while holding the others at some nominal (generally empirically determined optimal) value.

The values assigned to these variables constitutes a protocol regimen which is applied to several animals, over several sessions to obtain data for that regimen. The regimen schedule used in the simulations reported here are:

1. Interstimulus Interval (ISI), defined as the time period between onset of CS and onset of US signals.

2. Inter-episode Interval (IEI), defined as the time period between the offset of the US of one trial and the onset of the CS for the next trial. Also called the 'between-trial period.'

3. CS Duration, defined as the time period during which the CS signal is on.

4. US Duration, defined as the time period during which the US signal is on.

Figure 16 summarizes these variables in a typical delay conditioning protocol. The figure also shows a 'Sample Output' period after several trials. This represents a test probe in which the CS alone is presented and the CR is measured to determine the degree of acquisition. In addition to the above variables, the number of sequential trials prior to a test probe can be varied as can the amplitude of the signals. In order to keep the task manageable, within the scope of this work, these were fixed for the simulations run.

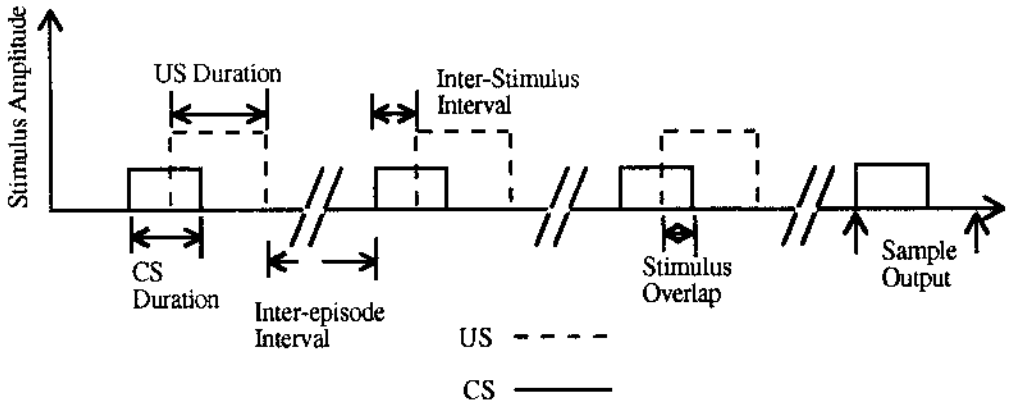

Fig. 16. The factors that are varied in the course of learning experiments in classical conditioning. All of the factors shown are time dependent. Not shown, but also of importance are amplitude values and relationships of the two signal types. As the figure depicts, after some number of pairings of the CS and US, with long intervals in between, the subject is sampled with a CS-alone test probe. This protocol is used in the simulation experiments presented here.

Experimental Protocol Setup

Protocol:	Local Interaction (LI)					Hebbian (HE)				
	ISI (.6 sec)	CS-dur (4 sec)	US-dur (60 sec)	IEI (5 min)	Extinct (25 min)	ISI (.6 sec)	CS-dur (4 sec)	US-dur (60 sec)	IEI (5 min)	Extinct (25 min)
Delay	0.0	1	5	1	5	0.0	1	5	1	5
	.2	2	10	2	10	.2	2	10	2	10
	.4	3	20	3	15	.4	3	20	3	15
	.6	4	30	4	20	.6	4	30	4	20
	.8	5	40	5	25	.8	5	40	5	25
	1.0	6	50	6	30	1.0	6	50	6	30
	1.2	7	60	7	35	1.2	7	60	7	35
	1.4	8	70	8	40	1.4	8	70	8	40
	1.6	9	80	9	45	1.6	9	80	9	45
	1.8	10	90	10	50	1.8	10	90	10	50
Trace	(.4 sec)	(4 sec)	(60 sec)	(5 min)	(25 min)	(.4 sec)	(4 sec)	(60 sec)	(5 min)	(25 min)
	.2	.5	5	1	5	.2	.5	5	1	5
	.4	1	10	2	10	.4	1	10	2	10
	.6	2	20	3	15	.6	2	20	3	15
	.8	3	30	4	20	.8	3	30	4	20
	1.0	4	40	5	25	1.0	4	40	5	25
	1.2	5	50	6	30	1.2	5	50	6	30
	1.4	6	60	7	35	1.4	6	60	7	35
	1.6	7	70	8	40	1.6	7	70	8	40
	1.8	8	80	9	45	1.8	8	80	9	45
		9	90	10	50		9	90	10	50
		10					10			

Table 1. Nominal and variation values used in the simulation experiments.

Table1, above, summarizes the values that were used to test the model neurons of Figures 13 and 14. Numbers in parenthesis show the nominal values used for simulations when a different variable was being tested. There are small differences in some values (particularly the ISI values) between the delay and trace protocols, reflecting the temporal

differences between these two protocols. Values used to test the Hebbian vs. Local Interaction models were the same.

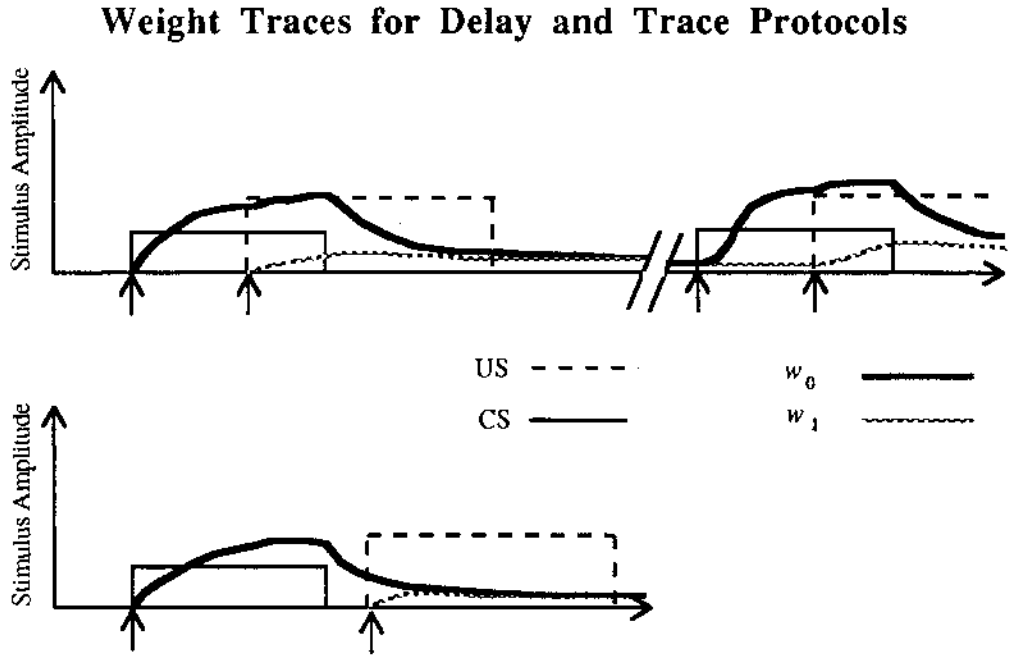

Fig. 17. Traces of the w_0 and w_1 weights as the CS and US signals are applied. The top panel shows an initial trial and a subsequent trial under the delay conditioning protocol. The bottom panel shows an initial trace conditioning result.

5.4 Single Neuron Performance

As mentioned previously, most of the simulation work has focused on synaptic weight evolution over the above described protocols. In this work I will focus on the output of the neuron which I believe more appropriately represents performance. However, it is instructive to observe the relationship that weight change has relative to the signal traces of Figure 15. Accordingly, Figure 17, above, shows traces of the w_0 and

w_1 values for the CS receiving (A_2) adaptrode, superimposed over traces of the CS and US inputs. In the case of delay conditioning several points are worth noting. First, if the ISI period is long enough to allow w_0 to approach its asymptotic value, we can see that the onset of the US causes w_1 to start to rise which, in turn, allows w_0 to marginally rise.

In the top trace (Delay Conditioning) is seen a second trial after an inter-episode interval. The values of w_0 and w_1 have continued to decay gradually, but remain slightly elevated due to the potentiating effect of w_1 on w_0. The impact is that at the onset of the subsequent episode (CS), w_0 rises slightly faster, with a steeper slope, and achieves a slightly higher asymptotic value. Over many trials, the slope of the rise of w_0 becomes significantly steeper, and, as will be seen in the simulation of a simple network below, becomes an important factor in the situation where two adaptrodes are competing through negative feedback.

The bottom trace shows that trace conditioning produces essentially the same relationship as does delay conditioning but at a lower asymptotic value. Also, since the US does not overlap the CS, there is no marginal rise in the w_0 trace. Rather, the late rise of w_1 simply changes the shape of the exponential decay curve following the offset of the CS.

5.5 The Simulation Environment

The simulations reported herein were done using an Intel Corporation, iA-486 platform running the MS-DOS, version 5, operating system. The program was written in C and accommodates two versions of the adaptrode equations. Since the objective of this line of investigation is

to eventually produce a real-time, on-line processing capability, the adaptrode formulations were converted to integer representation. All divisions were converted to right shift operations using the mathematical equivalence of:

$$fract = \frac{1}{2^f \tau}$$

where:

$f \in (0, 1, 2, \ldots F)$, F the number of right shift operations

τ is the number of time ticks between evaluations of *fract*.

Every rate constant (e.g., α or δ) in the adaptrode equations are converted to equivalent fractions by designating the f and τ values. For example, α_0 for the learning adaptrode in the simulations is set to 0.125 which is equivalent to $f = 3$ and $\tau = 1$. Similarly, δ_0 for the same adaptrode is 0.025 which is $f = 3$ and $\tau = 5$. In the latter case, the difference between w_d and w_{d+1} is taken every 5th time step and the result is shifted right 3 times. The resulting quantization error for time steps 2 through 4 is not judged to be significant in terms of overall performance of the adaptrode. A rule of thumb has been used in selecting values for f and τ that says for any given integer scale, keep f small so that the quantization error at the zero or w_{equil} asymptote is minimized. For example, it is found that for values of w_{max} of 200 - 255 (8 bits resolution) $f = 3$ produces good general results. In this case the quantized asymptote is 9 (w_d will never fall below $w_{d+1}+9$) which is mitigated by a secondary error minimizing method of decrementing the w_d value by one as long as it is

greater than the w_{d+1} value every ten time steps. The same approach is used at w_{d-1} asymptote.

The two versions of the adaptrode that have been simulated are the standard form as given in equations (3.6) through (3.11), which uses a pulse coded input (1 or 0), and a time-averaged version in which the number of pulses over ten time ticks are lumped and multiplied by 1/10th the maximum signal frequency is used as input. This second version affords additional speedup of computation without, apparently, sacrificing the quality of results. Both versions have been tested on the models described in this work and have been found to produce equivalent results. Thus, the time-averaged version is preferable for long simulation runs.

5.6 Simulations and Results

A model neuron was simulated in the above environment under the various protocols to be described in the text, and as summarized in Table 1. Table 2, below, gives the parameter values for the learning adaptrode in the model. The neuron consists of three adaptrodes, two for input and one receiving its input from the output of the neuron. One of the input adaptrodes is designated as the 'learning' adaptrode or the unit that encodes the memory traces of the CS and, through the hurdle mechanism at its w_1 level, encodes the association between CS and US or between CS and UR for the Local Interaction and Hebbian rules, respectively. The other input adaptrode is the US input. This adaptrode does not learn in the sense that it encodes an association. It is a single-domain unit that undergoes w_0

dynamics but does not potentiate. The output adaptrode, similarly, does not learn in the sense of encoding an association.

In the simulated neuron, the threshold value, q, is held constant at 100 for all runs.

Learning Adaptrode Parameters

Time Domain	α	f_α	τ_α	δ	f_δ	τ_δ
0	6.250E2	4	1	2.500E2	3	5
1	3.125E3	4	20	1.250E3	3	100
2	6.250E4	4	100	3.125E5	5	1000

w_{max}	γ_1	ρ_1
200	50	50

Table 2. Parameter values used for the simulations of the learning adatrode in the environment described in section 5.6. The non-learning adaptrodes had the same values for α_0, δ_0 and w_{max}.

Performance Acquisition

A conditioned response (CR) is acquired after repeated presentation of a conditioned stimulus (CS) followed shortly by an unconditioned stimulus (US) acting as a reinforcer. The CR resembles, or may be the same as the unconditioned response (UR). After so many trials where the CS-US pairing has occurred, the experimenter subjects the animal to a CS-alone episode having the same magnitude as the paired occurrence of CS. This is a test probe, and the experimenter measures the CR that obtains in

response to the CS alone probe. Over many series of trials and probes, the experimenter measures the evoked response.

The CR may have an all-or-none character or be a graded response. In the former case, the experimenter measures the increase in the probability of response and plots this as a function of the number of trials. In the latter case, the CR magnitude may be plotted against number of trials to show the acquisition of the association.

Much has been made of the characteristic S-shaped curve often found with these acquisition procedures (c.f. Klopf, 1988, esp. Figure 4). However not all acquisition measurements have proven to be S-shaped (Baxter, et al, 1991, esp. Figure 2.4). Coupled with the characteristic shape of the EPSP shown in the previous chapter, it is not at all clear how the synaptic weight change and the whole-animal acquisition curves are related. Klopf's model (1988) contains an explicit nonlinearity in the weight update function that accounts for the S-shaped curve he obtained. Other, more neurobiologically-motivated modelers have not found such a relationship in the physiological data (Baxter, et al., 1991).

For the simulations reported here the neuron output, a graded response, is used as an indicator of the acquisition of a CR. It is easy to picture the output of such a neuron coupled to an actuator device (muscle) that produces the UR/CR behavior. Therefore, in what follows, the measure of CR refers to the relative strength of the neuron output over some stated number of trial-test cycles. In all of the simulations reported below, the number of paired trials was 6 per set, followed by a single CS-

only test probe. The amplitude of the CS and US signals, while on, was held constant for all of the simulations.

Figure 18 shows the acquisition curves for delay and trace conditioning protocols for both the local interaction rule and Hebbian neurons These are the curves obtained from models using the nominal values for the variable factors (see Table1).

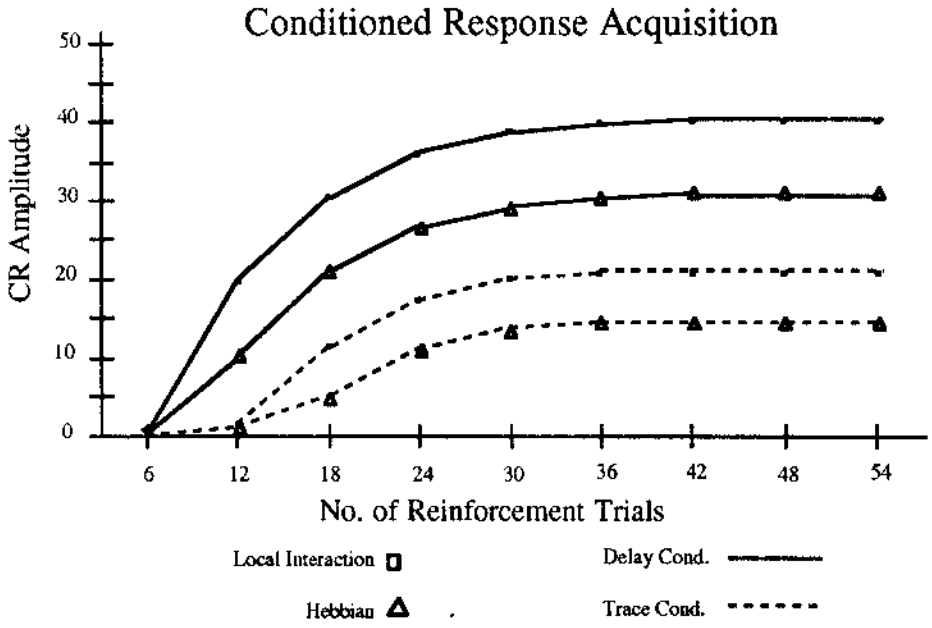

Fig. 18. Comparison of LI/Hebbian learning in delay and trace protocols.

At the end of each set of six reinforcement trials, a test probe is used and the neuron response is measured. The measurement is actually the average response over the period of the CS stimulus plus eight time steps after CS shutoff. This is because the neuron continues to fire for some period after the removal of the signal.

As can be seen by the graph of Figure 18, the rate of increase of the response negatively accelerates and the curves approach some asymptotic values at the fixed amplitude input. As can be seen from the curves there are no appreciable differences between the delay and trace protocols except that the maximum CR values are lower. The principle difference between the LI and Hebbian learning rules is that the latter is shifted to the right due to the time delay for feedback in the Hebbian architecture. The apparent S-shape to the curve here is an aberration of the graphing method and results from the fact that the resolution of the test probe is every 6th trial. Since the LI and Hebb neurons appear to perform essentially the same, except for the time delay effect, I will show the results primarily for the LI neuron, except in the sections on second-order conditioning and blocking where only the Hebb rule was used.

This acquisition performance alone demonstrates that the adaptrode neuron is encoding an associative relationship between the two signals in the sense that after acquisition the neuron responds to the CS as if 'expecting' the US. Furthermore since the two signals have a specific phasic relationship defined by the causal ordering constraint, the encoding captures the causal association if the US follows the CS. The encoding is completely causal iff it captures the association of a CS-before-US temporal ordering, and no encoding occurs for a US-before-or-at-CS ordering. In order to establish this test the interstimulus interval relationship was investigated.

Interstimulus Interval (ISI)

The ISI effect has been described as an inverted U curve (Baxter, et al., 1991; Klopf, 1988; Grossberg, 1991; Mackintosh, 1983). Its significance in delay conditioning is that it establishes a temporal contiguity window within which conditioning can occur. If a US arrives too soon after the onset of the CS, then the CS is not valuable as a predictor of the onset of the US. The behavioral interpretation is that CS-US associations represent a predictive binding of a non-semantic stimulus to a semantic one so that the occurrence of the CS can be used to generate an early onset of the CR (some version of the UR) so as to increase the probability of gain of a positive reinforcer (food) or avoidance of a negative reinforcer (pain). A CS that starts shortly before the onset of the US will not provide an adequate temporal advantage in the form of allowing time for an adequate response. There would be no purpose in encoding such an association. Typically, no encoding (as measured by the acquisition rate as a function of ISI) occurs if the US arrives before the CS (Alkon, 1987, Baxter, et al., 1991, page 24; Klopf, 1988, Mackintosh, 1983). The amount of encoding increases rapidly as the ISI increases, reaching a peak, in many animal models, at about 500 milliseconds, after which it falls off exponentially.

Figure 19 shows the results of simulating the adaptrode neuron over a range of ISI values. The vertical scale is a percentage of the total number of trials required to reach asymptote (42). At 800 milliseconds it required 50 trials to reach asymptote so that the percentage of maximum is .84. For the set of constants that were used (Table 2) the peak in this model appears at about 600 milliseconds and does not begin to fall exponentially

until about 800 milliseconds. Nevertheless, the characteristic inverted U-shaped curve with an exponential fall-off is evident. In several informal simulations using 'exploratory' values for several of the parameters, the shape of the curve, and its peak value varied somewhat. However, the overall inverted U form was found to be robust.

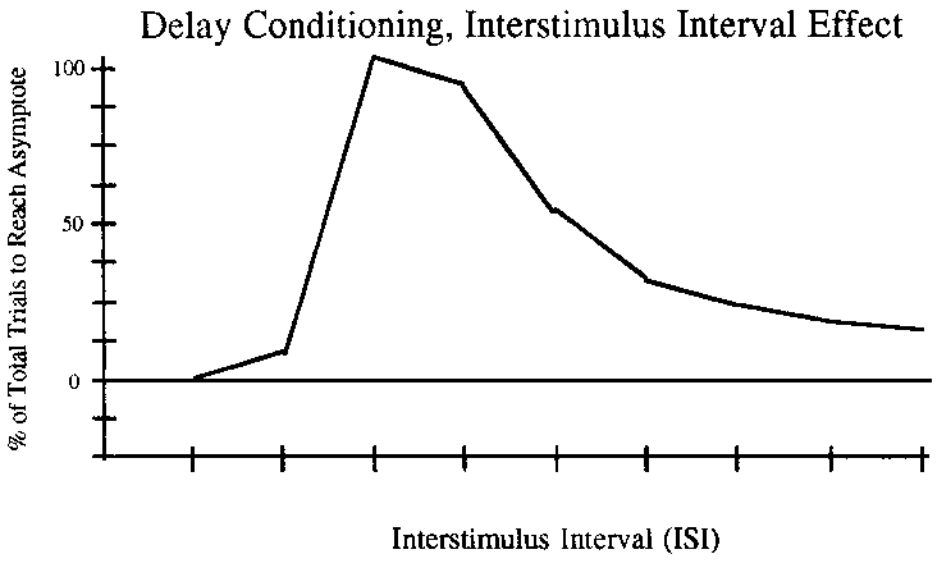

Fig. 19. The effects on CR acquisition of increasing the interval between onset of CS and onset of US.

The sharp rise in acquisition as the ISI period goes from .4 to .6 seconds, in Figure 19, is explained by the role of the hurdle computation of Equation (3.10). Specifically, the value of ρ is found to play a role in the timing of the onset of acquisition relative to the ISI. If ρ is set to zero, then encoding starts as soon as the US arrives (ISI = 0). That is, w_1 of the learning adaptrode can rise as soon as the hurdle signal exceeds whatever gate has been set.

Note that since the w_0 of the learning adaptrode cannot start rising and, hence, neither can the w_1 weight, until a CS signal is received, the adaptrode cannot encode anti-causal relations, that is situations in which the US may precede the CS in time. These types of relations are implicated in several encoding mechanisms used for pattern recognition (Hopfield, 1982) and abductive reasoning (Pazzani, 1990). Due to the inherent flexibility of the adaptrode implementation, it is possible that two, or more, adaptrodes could be set up to mutually provide the other's hurdle signal. Thus a set of adaptrodes would collectively code either forward or anti-causal associations and be useful in such situations.

Extinction

Memories, if not reinforced, fade with time. This is true of nonassociative as well as associative memories, requiring an associated reinforcer. Technically there are two kinds of memory decrement noted in the animal learning literature (see Mackintosh, 1983 for a summary). Standard forgetting is a slow loss of memory after a stimulus pair is no longer presented. *Extinction* may be the loss of memory of an association due to the occurrence of a CS without the reinforcement of the US. In the classical conditioning model, extinction of a memory arises from the presentation of the CS without the reinforcement generated by a following US. Presumably, the CS is no longer acting as a good predictor or has come uncoupled causally from the US and so should not be retained in association with the latter. A rather large topic of debate has centered around the exact form which extinction takes. That is, it could be due to

simple passive decay (or forgetting), it could be due to active decay (selective forgetting - Grossberg, 1991b), or as was originally suggested by Pavlov (1927), it could be due to active inhibition. Different opinions on the exact nature of the underlying mechanism abound. Confounding the picture is the fact that a new series of co-presentations of CS and US will illicit the CR in fewer trials than were required for the original training. This phenomenon is called *savings* and is considered a key signature of classical conditioning.

As stated above, it is probably dangerous to infer mechanisms from observations of the whole-animal model to the cellular substrates (and neural network models). It would be desirable to correspond performance observations in simple invertebrate systems such as *Aplysia* (Small, *et al.*, 1989) or *Hermissenda* (Alkon, 1987) with those in birds and mammals, in order to make such inferences.

In the Adaptrode model the nature of reduction of an effective weight (the response of the Adaptrode) is through passive decay, but one which is proportional to the distance the weight is above equilibrium. However, this is accomplished in a piecewise fashion across multiple time scales. Thus over short time frames, say a series of trials with short inter-episode periods, followed by presentation of the CS only, there would appear to be a rapid decay of the efficacy weight giving the impression of active decay. Over longer time frame protocols, the extinction curve starts to look passive. This is, in fact, what has been seen in simulations. As will be discussed below, in the case of learning a new, contrary association, we even see a curve which appears to be due to active decay. These findings

suggest that the appearance of passive versus active decay may be influenced strongly by the temporal nature of the protocol used. This is certainly a testable idea.

An interesting aspect of passive decay of long-term memory in the adaptrode is that the effective weight can decay below the threshold necessary for the single adaptrode to initiate a neuron output. The memory appears, from the outside to have decayed away. However, because there is a long-term trace remaining, a new series of CS-US co-presentations causes reacquisition in a fewer number of trials which is not unlike what occurs in animal models.

The graph in Figure 20 shows memory decrement in terms of the CR amplitude from a test probe at time intervals after the CR had reached asymptote (from Figure 18). Two delay conditioning curves are shown to demonstrate the difference in protocols. One uses a 5 minute inter-episode interval (IEI - see below) while the other uses a 10 minute IEI. These are, likewise compared to a single trace conditioning extinction curve in which the IEI is five minutes. Note that the trace curve actually extends out past the 10 minute IEI delay curve even though it starts from a lower CR amplitude. This was found to be due to the fact that in trace conditioning, the total period of time during which the CS/US combination is correlated is extended, compared with the delay conditioning protocol. The extension provides additional time for the w_2 level to potentiate since the latter is not, in these simulations, gated. It is not known if this relates to any specific animal model per se. But it does provide some support to the possibility that an adaptrode-like mechanism could account for the appearance of

active forgetting in one regimen/protocol while being based on a passive decay process. This remains an open issue.

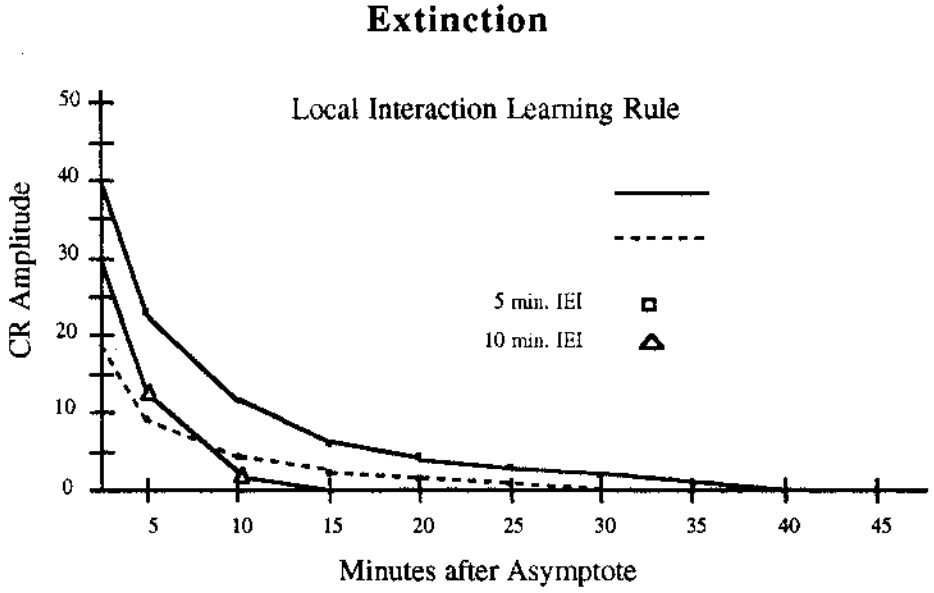

Fig. 20. Extinction of adaptrode storage terms after the reinforcement protocol is stopped.

Re-acquisition (Savings)

If some time after the extinction of a CR, the animal is exposed to a series of trials again, it re-acquires the performance of the CR in a shorter number of episodes than it took to originally train the animal. Figure 21 shows a graph of acquisition following a period over which the CR was allowed to extinguish (Figure 20). As can be seen in comparison to the acquisition curve of Figure 18, the CR is re-acquired in a significantly fewer number of trials. This is accounted for by the fact that the w_2 level remains sufficiently high that the adaptrode is not starting from zero. The w_2 level is, however, low enough relative to the neuron's threshold that the

adaptrode cannot generate a sufficient response to illicit a CR output during a test probe after extinction. The neuron appears to have completely forgotten the association, yet on being exposed to the CS-US combination again, shows some lingering memory trace by re-learning the performance in a shorter time.

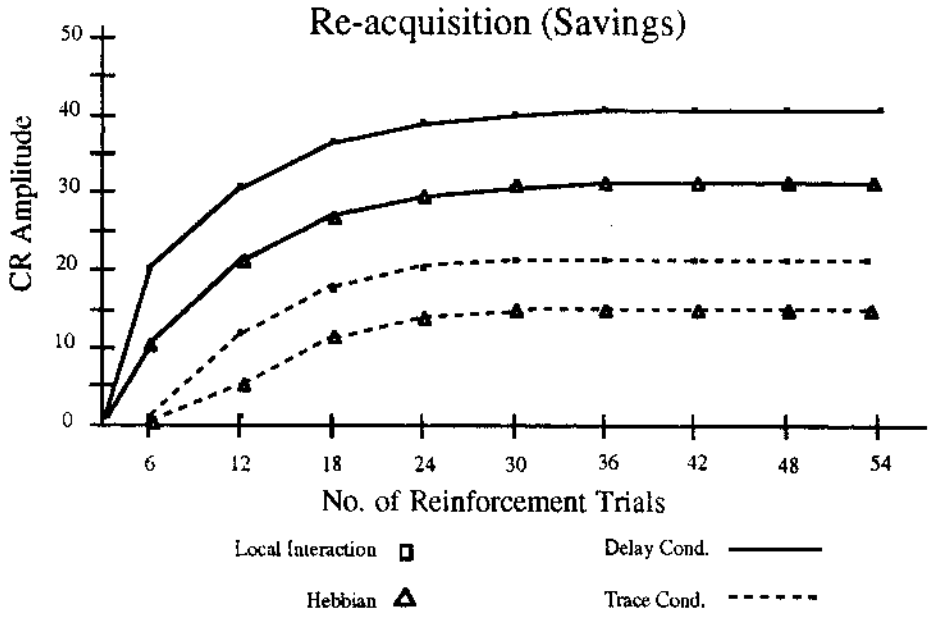

Fig. 21. A memory trace recovers toward its original strength with fewer reinforcement trials due to "savings". A longer-term trace, sufficiently below the threshold of the neuron could act as a higher starting point for re-acquisition of the response.

Inter-episode Interval

In animal models of classical conditioning it has been found that the period of time between trials has a acquisition rate effect. That is, if longer intervals are given between episodes, then it takes longer to acquire the CR and the level of performance is somewhat diminished. Since an

adaptrode forgets during the period of time between episodes, it should be the case that an acquisition rate effect should be seen for this model as well.

Indeed, as the graph in Figure 22 shows, CR acquisition declines with increasing IEI as expected. Three curves for each of the two protocols are plotted as in Figure 18 (solid lines are delay, dashed lines are trace conditioning). Each of the three curves shows the effects of a longer IEI.

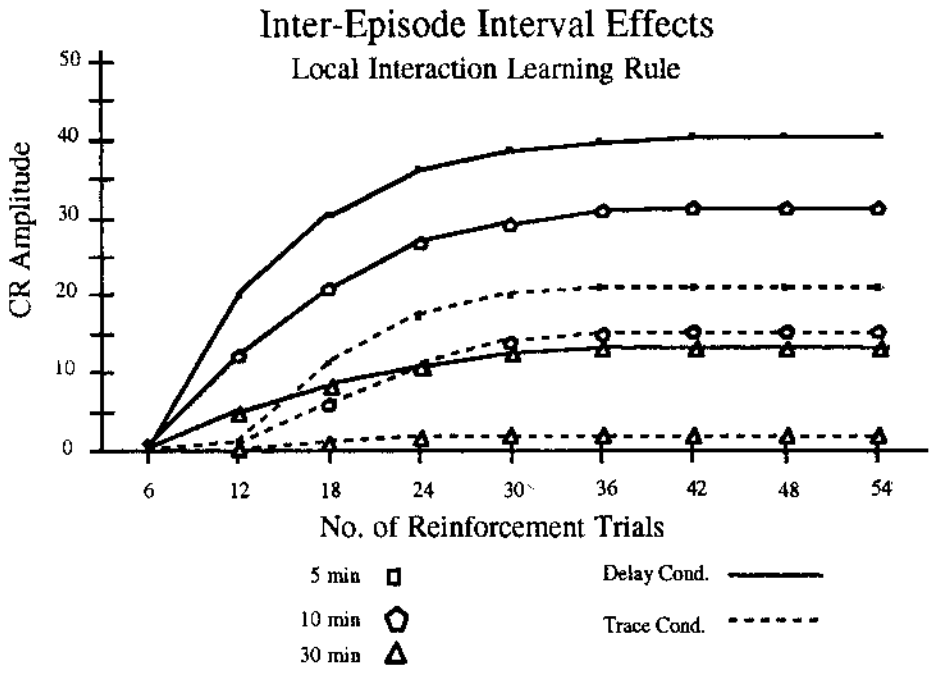

Fig. 22. It takes longer to learn something if the period between trials is extended. As the interepisode inteval (IEI) increases, the rate of CR acquisition declines.

CS and US Duration

It is known that for the delay conditioning protocol, the duration of the CS has little, if any impact on the rate or strength of acquisition, while the duration of the US has a significant effect. The general interpretation of this is that a longer US is more significant to the animal than a short one, so should cause a stronger encoding of the association.

The adaptrode-based neuron shows these same effects. Figures 23 and 24 show graphs of CR amplitude as a function of the CS and US duration, respectively. As can be seen readily, the duration of the CS has little overall impact on acquisition while the US duration has a very marked impact. This can be explained by the potentiation of w_1 which will continue over a longer time if the US is present. However, such would not be the case unless the ultimate decay of w_0 was delayed, as it is in the adaptrode equations by uncoupling the decay rate from the encoding rate. If standard exponential smoothing were used, the w_0 value would decay too rapidly to allow the longer term pull-up of w_1 to occur.

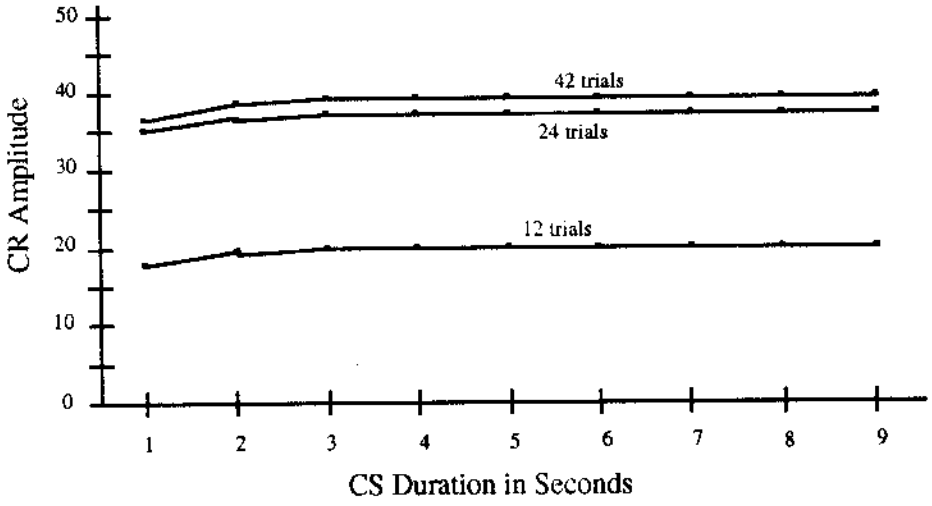

Fig. 23. CR amplitude is relatively insensitive to changes in CS duration as predicted from animal models.

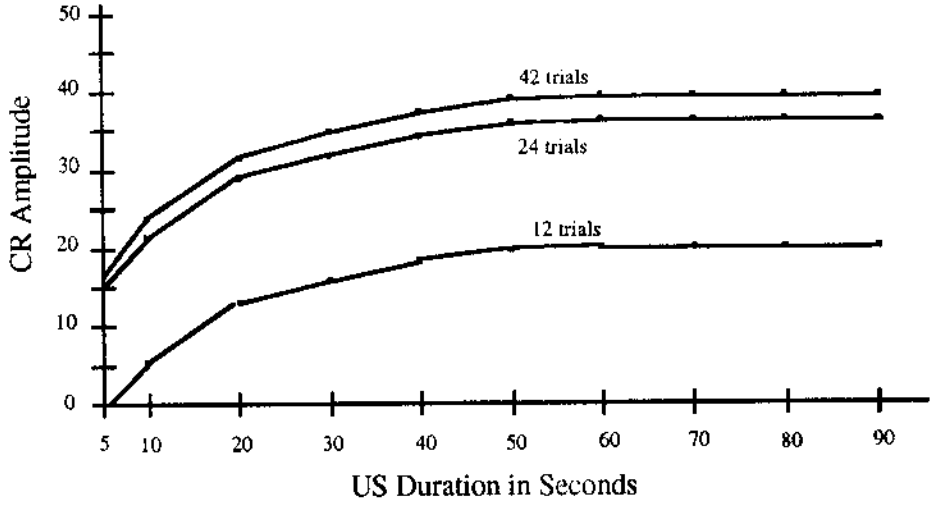

Fig. 24. CR amplitude is sensitive to changes in the US duration due to the amount of potentiation of w_1 which increases with longer US duration.

Second-order Conditioning

Following the protocols given above the CS becomes able to initiate the CR by itself. This is generally referred to as primary conditioning. Second-order conditioning comes from the pairing of a second CS with a following CS that has already been conditioned to predict the US (i.e., produce the UR). Using the Hebbian rule this phenomenon was simulated. One neuron, in this model, is comprised of one US input and two or more CS inputs. After training the system on a particular CS-US pairing using the nominal values of Table1, which produced a deep, sustained encoding, the trained CS, call it CS_1, was paired with another CS input, call it CS_2 using the same protocol as if CS_1 were now the US. The latter successfully became a predictor of the former, and that, in turn remained a predictor of the US.

If presentations of the CS signals alone were not occasionally followed by reinforcement from the US, then both memory traces extinguished in the reverse order. That is, CS_2 extinguished first followed by CS_1.

Blocking

An interesting phenomenon occurs when a second CS is correlated with the US in the same temporal frame as the first CS but for which the first CS has already encoded the association. In this case, the existence of a CS_1-CR coding blocks the encoding of the second CS, CS_2, from becoming associated with the US-UR. Since a predictor of the US already exists,

there is presumably no need for a second predictor having the same temporal relation so it is effectively ignored.

It turns out that the same mechanism that gives rise to secondary conditioning is involved in blocking, namely the Hebbian rule. The actual blocking, however, is due to the gating of input signal to the adaptrode via the hurdle signal calculation. The onset of the CS_1 signal is sufficient to generate the CR which, in turn, generates an output from the A_0 adaptrode. This response is used in the hurdle signal calculation but due to the causal temporal ordering imposed by the ρ factor in Equation (3.10), the hurdle will arrive at adaptrode A_3 before its response has reached a level sufficient to overcome the ρ and allow potentiation of A_3. The arrival of a hurdle signal prior to the buildup of a response signal prevents the input signal at x^0 from entering the adaptrode, thus preventing any growth of w^0 and any run-away effect on potentiation.

Importance of Simulation Results

What has been demonstrated is that the adaptrode-based neuron encodes memory traces in an emulation of biological systems. The above reported results mirror the invertebrate models of classical conditioning in several important features or signatures of conditioned learning. Furthermore, these simulations were conducted in such a way as to better replicate the conditions found in actual animal trials, including long inter-trial periods.

However, these results only show that the adaptrode model is at least as efficacious as some other, more abstract models such as the drive-reinforcement (Klopf, 1988) and temporal difference models (Sutton and Bartow, 1987). If this were all that was shown, the best that could be said for the adaptrode model is that, in principle, it could lead to a physical implementation more readily than these other models. Fortunately, there is more. Attention is now turned to the real benefit of the multi-time scale learning mechanism. As claimed in Chapter I, the nature of the real world is that event relationships that may be true in one time scale can change in another. Conventional neural networks nor any machine learning system that assumes a closed world boundary can accommodate changes without washing out the prior encoded associations. In the next section a small network of mutually competing neurons that will be shown to be able to learn two contrary conditions that separate in time (as opposed to space) without interference.

5.7 Basic Associative Network (BAN), New Results

A network of mutually competing neurons can be constructed as shown in Figure 25. In the simulations reported here, the local interaction model was used for all of the neurons. All neurons had the same parameter set as given in Table 2. There are several important features of this network. First, each US input to the network has an excitatory, non-learning synapse (w_0 level only) on exactly one neuron. That is, there is a one-to-one relationship between US inputs and neurons such that each neuron is said to code for a separate US. Some US inputs may have

inhibitory synapses onto other neurons to implement a priority scheme, but that will not be considered here.

Any number of CS inputs may have excitatory, learning synapses on any of the neurons. In fact in the general case, all CS inputs would have learning synapses on all neurons. This makes it *a priori* feasible that correlations between any CS and US can be encoded. The biological justification for this comes right from classical conditioning. One might wonder what possible connection could exist between the hearing a bell ring and salivating in preparation for eating. But that is exactly what Pavlov (1927) discovered and it suggests that many seemingly non-related signals have convergence points somewhere in the nervous system in order for such odd-seeming associations to be made. From the biological perspective it makes sense to be ready to exploit any causal association no matter what, since the environment is nonstationary and likely chaotic.

In these simulations, the simple, two-neuron model with a single CS input was used. The objective of the simulation was to see what happens when, after training the system to associate the CS with one of the two USs, US_1, such that a long-term memory trace is encoded (w_3 approaches asymptote), we change the situation and the system is exposed to episodes where the CS is followed by the alternate, and mutually exclusive US. The first association was learned from long experience, whereas the second association was transient by comparison.

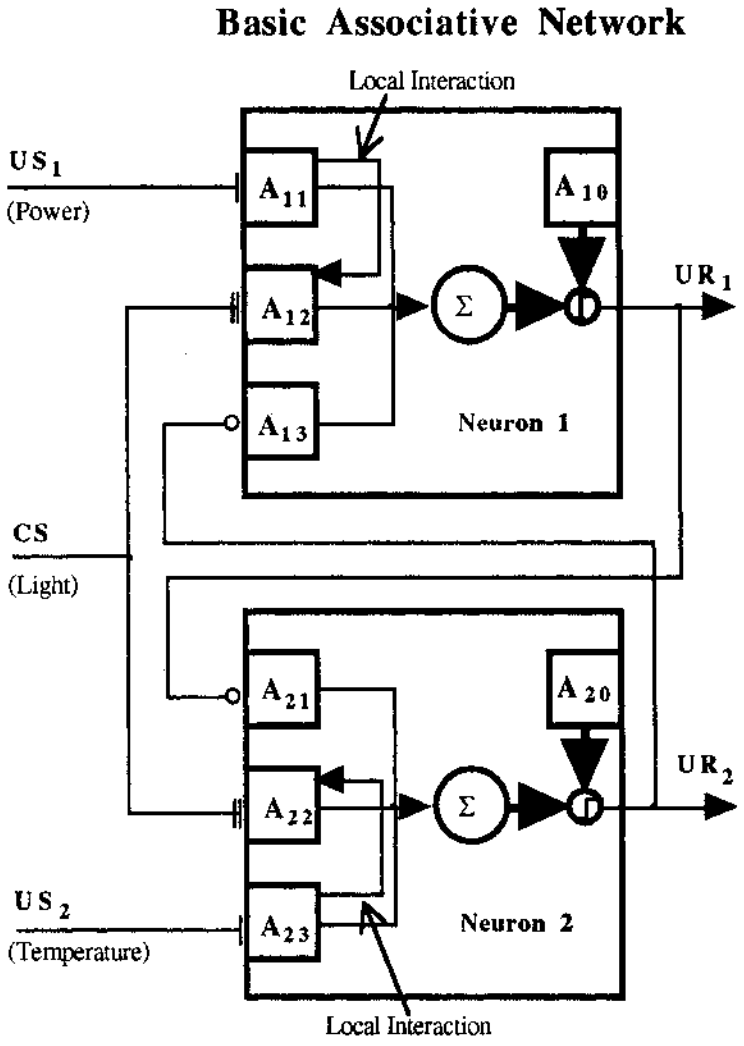

Fig. 25. Schematic representation of the BAN network for two neurons. The neurons compete with one another through mutual inhibition (circular terminus). External inputs include two unconditionable stimuli (US_1 and US_2) which are excitatory on neurons 1 and 2, respectively, through non-learning adaptrodes A_{11} and A_{23} (Note: The first subscript indicates the neuron number and the second indicates the adaptrode index within the neuron). A third input, CS, has excitatory connections to both neurons through learning adaptrodes A_{12} and A_{22}. Learning connections are indicated with a double, flat terminus. A local interaction hurdle signal (e.g., from A_{11} to A_{12}) is used to generate associative learning; see text for details.

The two neuron network shown in Figure 25 might be viewed as a simple robot decision controller. The robot, in this case, can sense three environmental conditions, the presence of a power source for recharging its batteries (i.e., food) at US_2, the presence of light at CS, and the presence of high temperature (i.e., pain) at US_1. The output of each neuron in the BAN represents an unconditioned response to the corresponding unconditioned input. Thus, a signal at US_1 will result in an output at UR_1; values for θ, the threshold, and κ, the preweighting constant, have been selected to assure that an unconditioned input results in an unconditioned output. An input to the conditioned stimulus alone will not, however, produce a response of significance in either of the two neurons even though it is wired to both of them. This is due, in part, to the slower, and weaker response output from the learning adaptrodes (A_{12} and A_{22}). It is also due to the cross inhibition between neurons. Even if a very strong signal at CS is sustained for a period long enough to drive the w^0 signals in A_{12} and A_{22} high enough to cross their respective thresholds, the cross inhibition between neurons ensures that output from each is inhibited. A signal at CS alone can activate a neuron only if the learning Adaptrode (either A_{12} or A_{22}) has potentiated to a point that gives w^0 the ability to climb to a value greater than θ and it does so before the competing neuron becomes active.

A priori, there is no reason to associate light with either pleasure or pain. It is a neutral signal. If, however, there is a causal relationship between the presence of light and the presence of one or the other of the unconditioned stimuli, following the constraints of causality discussed above, then the occurrence of light could be used as a predictor of the occurrence of the US. Such a prediction could give the robot a "head start" in reacting to the US, which, after all means something important. How then can an association be encoded in the BAN, if a causal relationship exists?

The BAN was trained with a CS-US_1 combination, delay conditioning protocol as described previously. Potentiation of w_1 and w_2 occurs in A_{12} as the result of local interaction between A_{11} and A_{12}. This potentiation is sufficient so that Neuron 1 will win the competition as a result of input at CS only. Figure 26 shows the situation at the first episode of training. The graphs in Figure 27 shows the relative levels of w^0, w^1, and w^2 in A_{22} and A_{12} at the third episode of stimulation by CS and US_1. Since A_{21} is not active, no potentiation occurs in A_{22}. The intersecting lines at the leading edge of A_{12} show the point at which the corresponding time tick and potentiation level in A_{22} occurs. These cross hairs will be used throughout to show how one of the two adaptrodes will win the competition by responding incrementally faster than the other.

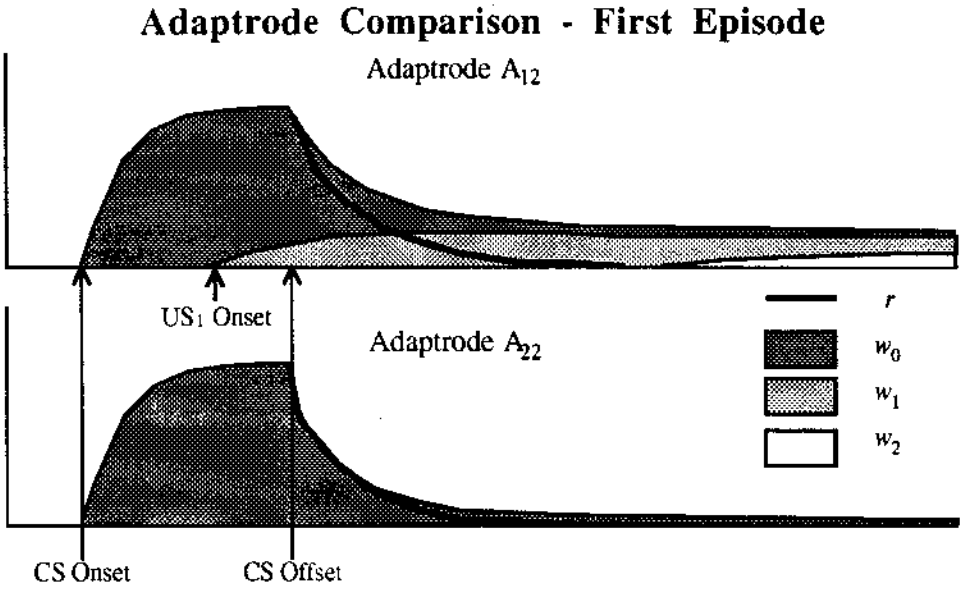

Fig. 26. Comparisons of the activities of A_{12} and A_{22}. This set of graphs shows the situation at the end of the first episode of training.

By the end of the fourth episode adaptrode A_{12} produces a substantial response to the CS input before A_{22} (Figure 28). Figure 29 gives an idea of the relationship between the activity in adaptrode A_{12} and the neuron output. In the initial trial, the neuron produces an output only when the US_1 signal comes on. By the last trial in the third episode (as in Figure 5.15) the neuron response starts earlier due to the effect of A_{12}. By the end of the fourth episode the effect is pronounced. Subsequently, the CS signal alone will cause neuron 1 to fire which in turn will strongly inhibit neuron 2. At this point the network is trained to respond to the CS by producing the CR which is some lesser value of the UR.

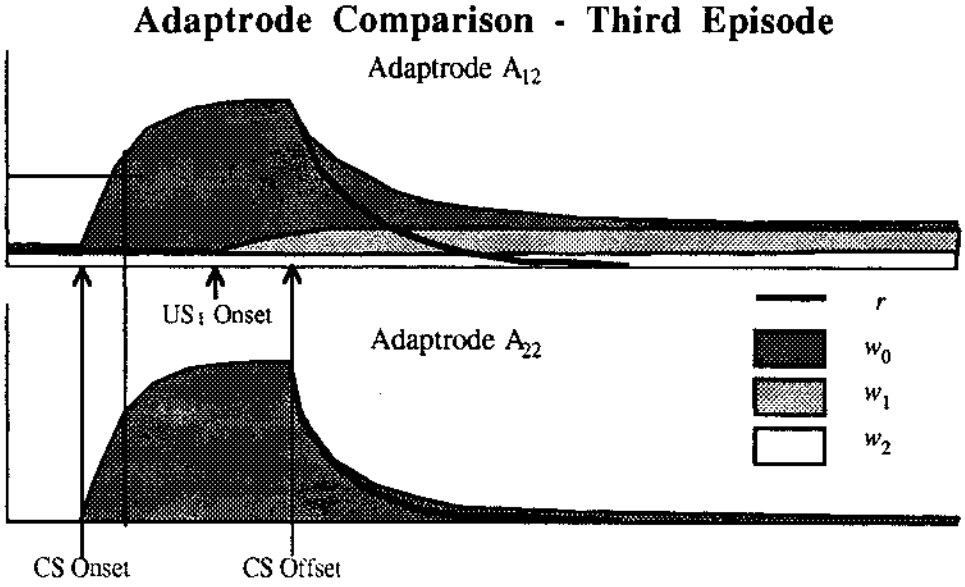

Fig. 27. Potentiation comparison between A_{22} and A_{12} at third episode. w_0 of A_{12} starts the episode with an advantage over w_0 of A_{22}. Thus the point where the lines cross in A_{12} shows that its response is above that of A_{22}. Before the end of the episode US_1 turns on and is sufficiently strong so that the response of A_{11} exceeds the gate threshold of w_1 in A_{12}, thus allowing further weight potentiation of the latter. As a result of the rise in w_1 of A_{12}, and its very slow decay compared with that of w_0, the w_0 trace remains elevated long past the end of the episode. At the same time the w_0 trace of A_{22} decays back toward zero. The level of the w_0s for these two adaptrodes represents their 'expectation' of a future occurrence of an episode involving the association of the CS and their respective US signals.

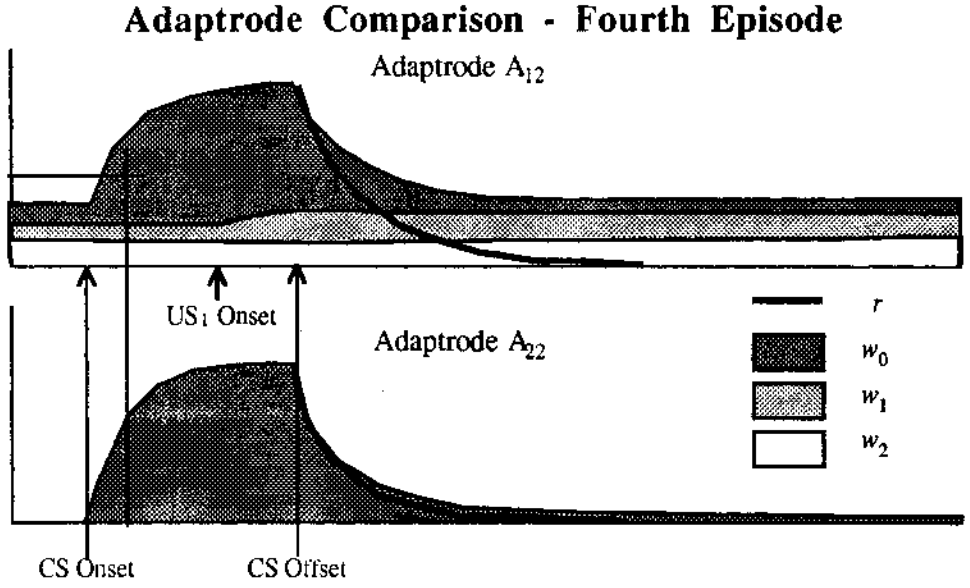

Fig. 28. The state of the two competing adaptrodes at the end of the fourth episode. A_{12} is now sufficiently potentiated that it wins the competition by producing a stronger response, quicker than A_{22}.

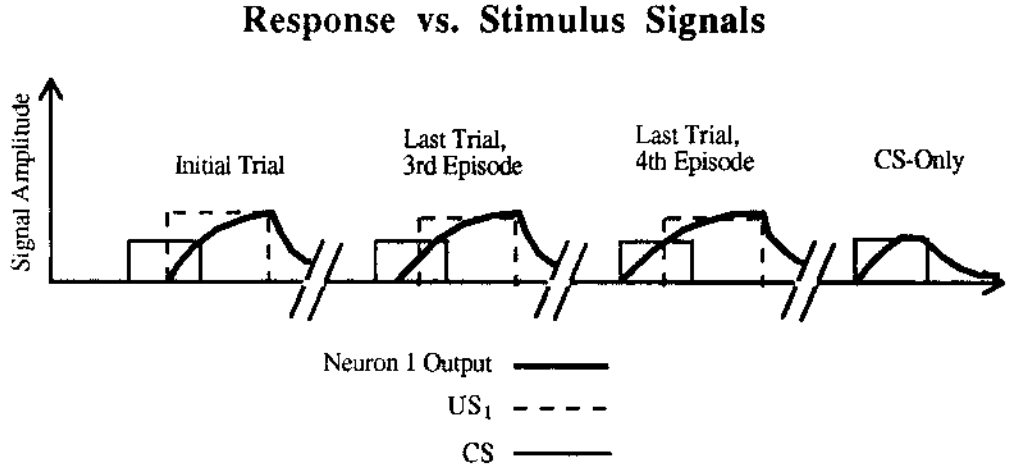

Fig. 29. The evolution of the efficacy of adaptrode A_{12} in generating neuron output. By the end of the fourth episode the neuron can be activated by the CS signal alone. This corresponds to the generation of the CR by the CS input.

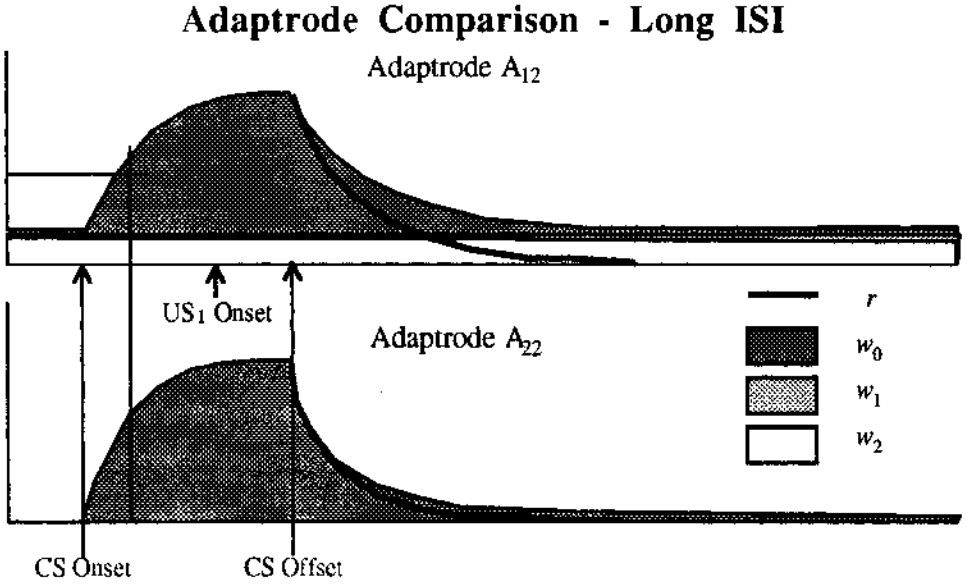

Fig. 30. After a long interval of time during which no signal is received, the long-term memory encoded in adaptrode A_{12}'s w_3 level is sufficient to produce a response in the number 1 neuron.

After the network has aged for some period of time without any inputs, Adaptrode A_{22} maintains a residual memory of the association to which it has been exposed (Figure 30 above). As a consequence of this residual, the network has established a relationship between the associated signals which can be recalled when a signal is injected at CS only. This corresponds to the robot seeing light alone.

In this way, the network has learned to exploit CS, a non-meaningful signal in its own right, as a predictor of some meaningful signal (in this case power availability). The frequency and intensity with which these two signals have occurred together in the past will determine the reliability of

in the past then the robot will have a degree of confidence that the association will be valid in the future. What happens, however, if conditions change for some reason? Should the robot forget this learned association to learn the new situation? Should it ignore the new situation, treating it as transient or noise? If the new situation is as meaningful to the robot as the old, then surely it cannot ignore it. On the other hand, if the old association is washed out by learning the new association and the new one is transient or short-term, then the robot will be penalized by having to relearn the old association again. This would take a whole new set of training trials. These are the questions explored next.

Suppose that, after primary conditioning as described above, for some short interval of time a contrary situation develops whereby a causal relationship exists between US_2 and CS as, for example, might happen if a fire were to break out in the robot's environment. First, the robot should emit the appropriate response - that is, it should avoid the fire. Second, the robot needs some means of determining the significance of this contrary condition. It has no explicit knowledge of fire or lights. It does not 'know' that fire is an unusual circumstance. Is this new situation just noise or does it represent a more permanent change in the nature of the environment? Third, in the event that this new relationship is temporary (as compared to the duration of the prior conditioning), the robot should not forget the prior association since there is some possibility that the old relationship will again be the norm in the future.

Conventional artificial neural networks cannot easily deal with this last condition. Systems that do not learn in real-time cannot deal with the first condition. The problem resembles the classic XOR problem that Minsky and Papert lamented could not be solved by the perceptron (Minsky & Papert, 1969; Rumelhart, Hinton & McClelland, 1986). In this case the problem separates in time as opposed to space and must be learned on-line, in real-time.

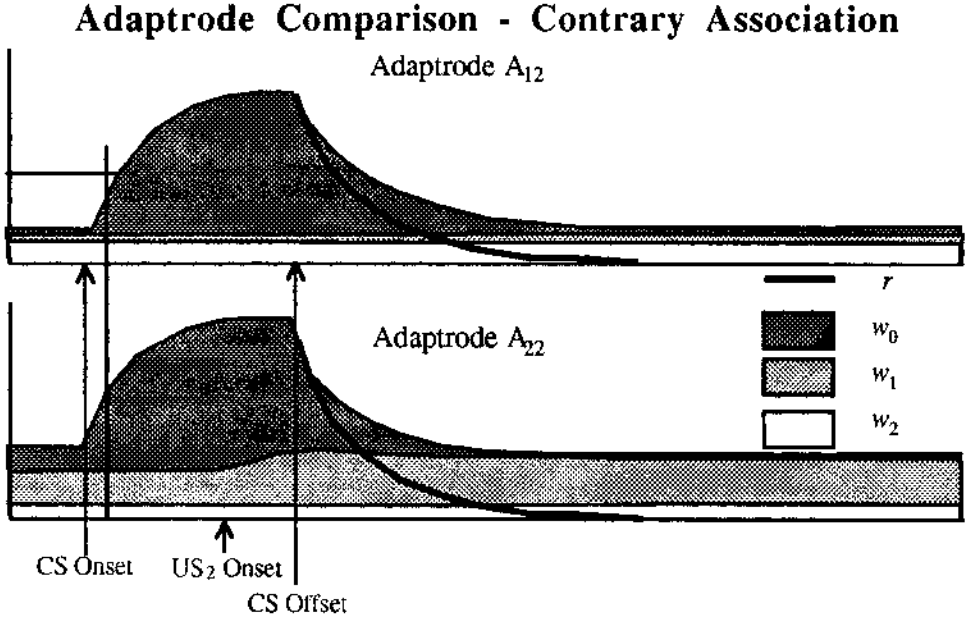

Fig. 31. After two episodes where the CS is followed by US_2, adaptrode A_{22} shows signs of potentiating to a level that is sufficient to beat adaptrode A_{12} through inhibitory feedback. Note, however that A_{22}, while showing a higher potentiated value for its w_1 level, has a lower value for its w_2 level than does A_{12}.

In this experiment the network is exposed to US_2 and CS for two episodes (Figure 31, prior page). This situation is clearly contrary to the system's prior conditioning. The system will initially start to fire the neuron in which the long-term association is encoded as if responding to the availability of power - the original semantic binding of the CS (Neuron 1). It takes a small amount of time for the w_0 value of A_{22} to build to a level sufficient to override, through the inhibitory link at A_{13}, the output of Neuron 1. However, due to the more rigorous firing of the US_2/UR_2 neuron, the latter wins the competition leading to the appropriate response (avoidance) at UR_2. The robot avoids the fire.

If the network is presented with the CS input only within a short time after this last episode (IEI = 10), it will, for a short while, respond, if only weakly, with output at UR_2. The reason is that w_1 of A_{22} has risen to a level just slightly greater than that of w_1 of A_{12}. This will persist until w_1 of A_{22} has decayed to the same level as w_1 of A_{12} at which time the response will be ambiguous. The ambiguity will not last for long. Since the exposure to the contrary condition was short, compared with the "normal" association, and w_2 of A_{22} did not rise significantly, w_1 of A_{22} will continue to decay, falling below the level of w_1 of A_{12}, which is being supported by its w_2 level. At that time, after two more episodes after the contrary conditioning, the network will respond with the original CR_1 response as the older association reemerges as dominant (Figure 32).

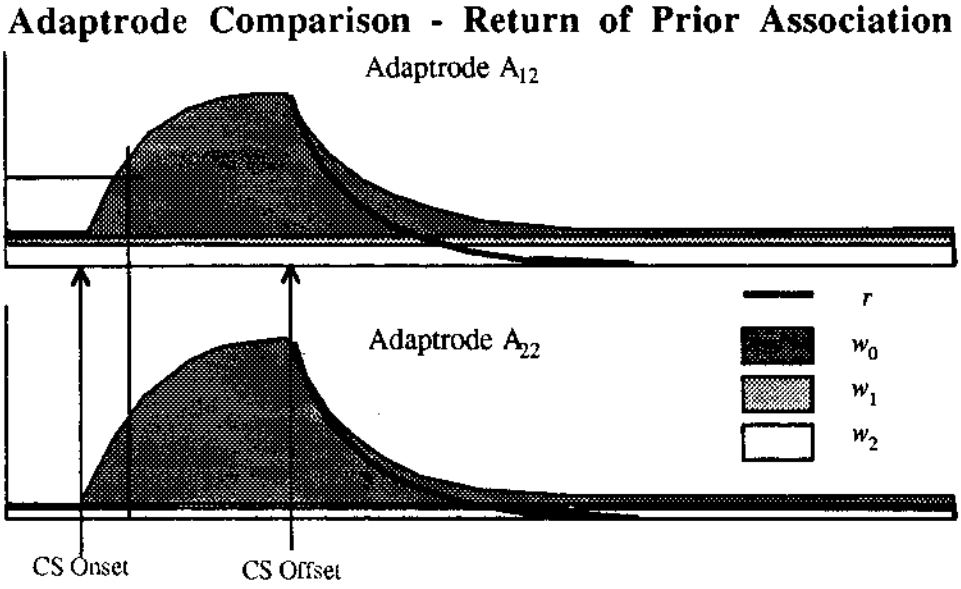

Fig. 32. The older association between CS and US_1/UR_1 emerges after potentiation of w_1 at A_{22} decays below the w_2 level of A_{12}.

This network, then, has the capacity to encode contrary associations which separate in time as opposed to space. The memory traces can be maintained without interference. If the short-term, newly encoded association is actually caused by a new causal alignment in the environment, then it will be reinforced and eventually completely override the former memory. If, on the other hand, the environment returns to the alignment which gave rise to the former encoding, the older memory will reemerge. It will be strengthened (in essentially the manner of re-acquisition) by the old reinforcement while the newer trace will decay.

5.8 Concluding Remarks

These simulations provide evidence that the adaptrode model meets many of the behavioral and performance criteria for classical conditioning demonstrated in the empirical results from animal learning experiments. It should be noted that, except for the lack of an S-shaped (logistic) acquisition curve, the results reported here are in accord with many of the existing models based on biologically-plausible considerations (Alkon, 1987; Baxter, et al., 1991). The results from the last simulation reported above extends the capacity of a neural network model to emulate some important biological memory behaviors. Namely, it provides a means for encoding memory traces that are time-domain specific. It provides a mechanism for learning associations that vary by time domain such that nonstationary processes are accommodated. Learning new, possibly transient, associations need not wash out existing memory traces.

CHAPTER VI

CONCLUSIONS AND FUTURE WORK

The adaptrode model has been advanced in order to address critical questions in machine learning with respect to encoding memory traces that are time-domain dependent, serve in nonstationary environments and capture causal relations between non-semantic and semantic cues. The argument has been put forth that these factors characterize the quality of animal learning models, specifically the paradigm of classical or Pavlovian conditioning. Further, these research findings establishe that the adaptrode, to a first approximation, meets many of the fundamental criteria of that paradigm. In Chapter I, the animal learning problem for contrary associations that separate in time was posed. The research results support the adaptrode model's claim to solving that problem.

The adaptrode was derived from considerations of the dynamics of synaptic modification and adaptive response learning. It was shown that the derived model shares some important structural features with what is now understood about the biophysical processes that underlie synaptic plasticity.

Causality relations, where a unidirectional temporal contiguity constraint applies to the occurrence of the non-semantic and semantic cues, is a natural outcome of the way the adaptrode emulates membrane reversal potential, locking out associations that violate the proper phase relationship between signals. This capacity was shown to give rise to the same inverted-U dynamics seen in conditioning acquisition experiments.

Learning and representing causal relations is an important problem in AI research. The adaptrode provides a computationally tractable means of providing on-line, real-time learning of such relations. It remains open as to the nature of a network architecture that would ideally make use of this mechanism. However, the architecture of the BAN suggests some ways in which representational issues, particularly the schism between neural networks and symbolic systems, may be resolved. In the BAN each neuron represents a semantic stimulus/action pair. Each non-semantic or conditionable input represents a different feature of the environment. Thus, each neuron is a kind of associative rule which relates temporally ordered feature occurrences to an action. Unlike the fully distributed representations of neural networks, where the 'rules' are embodied in the weight matrix as a whole, this architecture provides a means for identifying each 'rule' and by examination of the w_0 levels of each of its adaptrodes determining why the rule was activated.

Temporal reasoning, another important area in AI research, depends on representing temporal patterns. Adaptrodes explicitly represent temporal information in the dynamics of the weights. More importantly, the adaptrode encodes information in multiple time scales, providing a way

to deal with signals whose statistical properties may vary in one time domain but remain stable over a longer time scale. Due to the decay of even the longest-term weight level in the model, even long-term nonstationarity may be addressed by gradually forgetting those relations which are no longer true.

There are several near-term areas where the adaptrode model might find applicability. Currently multivariate correlation analysis for data sets involving phase relationships is computationally intensive. Though this is not yet proved, I conjecture that under certain technical conditions the adaptrode is formally *approximately* equivalent to a moving average method with a biased window on a stationary process. If so then it is also *approximately* equivalent to an ensemble average under those same conditions. It would follow that a set of adaptrodes could be used to capture the correlations between time series data in real-time and on-line. The expected values are encoded in the weight vectors. If full potentiation lockout is employed (setting ρ so that the CS must arrive before the US) then an analysis of the adaptrodes' weights would suggest causal relations. Furthermore, adaptrodes with weight values close to or at zero would show which data are candidates for elimination. While such a method would not replace traditional statistical methods it could provide a useful, and efficient adjunct to suggest more efficient handling of the data sets in the traditional methods.

Another obvious application is in the field of autonomous agent learning, which was the motivation for this work. Neural networks based on adaptrode learning may be able to emulate capabilities of animal nervous systems. Two areas for using adaptrode learning are already under consideration. The first is their use in mobile, autonomous robots. The computational efficiency of this mechanism along with its naturally parallel architecture lends itself to reactive systems operating in *real* real-time (human perception time). The BAN of Chapter V has already been suggested as a core for basic seek/avoidance control. The other area of agent research is the logical equivalent of a robot - a knowbot. A knowbot is envisioned as a personal assistant in the computer - human interface that will aid users in accomplishing a number of tasks. One such task under serious consideration is an agent that "forages" for information in a wide area network, such as the Internet. The knowbot would emulate foraging animals such as ants or bees. The success of foraging depends on the agent's ability to learn its environment.

Adaptrode learning may also play a useful role in the classic arena of neural networks - pattern recognition. By setting the ρ value of Equation (3.9) to zero, one eliminates the strict causal ordering that is required for classical conditioning. In this instance, the adaptrode simply records correlations that are aphasic, but constrained to a well defined temporal window of opportunity. Under these circumstances adaptrode-based neural networks could capture ordinary static correlations as in single image patterns. The advantage of using adaptrodes in this area stems from their real-time, on-line capabilities which allows for continuous learning.

The work reported here has laid a foundation for a more advanced machine learning capacity with respect to temporal encoding. The adaptrode algorithm is designed for real-time, on-line learning that does not destroy memory traces that separate in time. Additionally, the algorithm is computationally efficient and suitable for direct digital implementation in VLSI due to its use of simple integer operations. In order to address some of the above application possibilities additional work is required. This includes comparative as well as formal analysis of adaptrodes as correlators with respect to traditional statistical techniques. Simulations of larger networks with richer stimulus environments need to be built in order to better understand the potentials of using multiple synapse, and neuron types in a single network.

From the biological-correspondence side, it would be worthwhile to explore the effects of including nonlinear elements in the potentiation coupling terms of Equation (3.6). Some very preliminary, informal work with a nonlinear adaptrode indicates that it can generate an S-shaped acquisition curve as mentioned in Chapter V, section 5.7. The addition of nonlinearity to the model also holds promise for interesting chaotic dynamics which are thought to play a role in vertebrate nervous systems.

Both artificial intelligence and neural network investigations have helped to shed light on learning an memory in animals and man. The adaptrode model may contribute to this tradition by providing a new dimension, multiple time scales, for investigating adaptive phenomena. It might prove useful in the field of learning about learning.

REFERENCES

Alkon, D. L. (1987). *Memory Traces in the Brain*. New York: Cambridge University Press.

Alkon, D. L. (1989). Memory storage and neural systems. *Scientific American,* July, Vol. 261, No. 1, 42-50.

Alkon, D. L., Blackwell, K. T., Barbour, G. S., Rigler, A. K., & Vogl, T. P., (1990). Pattern-recognition by an artificial network derived from biologic neuronal systems. *Biological Cybernetics*. **62**, 363 - 376.

Alkon, D. L., Vogl, T. P., Blackwell, K. T. & Tam, D. (1991). Memory function in neural and artificial networks. In M. L. Commons, S. Grossberg & J. E. R. Staddon, (Eds.), *Neural Network Models of Conditioning and Action* (pp. 1-11). Hillsdale, NJ: Lawrence Erlbaum Associates.

Anderson, S., Merrill, J. W. L. & Port, R. (1989). Dynamic speech categorization with recurrent networks. In D. Touretzky, G. Hinton & T. Sejnowski (Eds.), *Proceedings of the 1988 Connectionist Models Summer School* (Pittsburg, 1988), 398-406. San Mateo, CA: Morgan Kaufmann.

Aparicio, M. (1988). *Neural computations for true Pavlovian conditioning: control of horizontal propagation by conditioned and unconditioned reflexes.* Unpublished doctoral dissertation. University of South Florida.

Aparicio, M. & Levine, D. S. (1994). Why are neural networks relevant to higher cognitive function? In D. S. Levine and M. Aparicio (Eds.), *Neural Networks for Knowledge Representation and Inference*. pp 1-26. Hillsdale, NJ: Lawrence Erlbaum Associates

Bak, P., Tang, C. & Wiesenfeld (1988). Self-organized criticallity. *Physical Review A.* **38**, No. 1, 364-374.

Baxter, D. A., Buonomano, D. V., Raymond, J. L., Cook, D. G., Kuenzi, F. M., Carew, T. J. & Byrne, J. H. (1991). Empirically derived adaptive elements and networks simulate associative learning. In Commons, M. L., Grossberg, S. & Staddon, J. E. R. (Eds.), *Neural Network Models of Conditioning and Action.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Brooks, R. A. (1991). *Intelligence without reason.* A.I. Memo No. 1293. MIT Artificial Intelligence Laboroatory.

Buonomano, D. V., Baxter, D. A., & Byrne, J. H. (1990). Small networks of empirically derived adaptive elements simulate some higher-order features of classical conditioning. *Neural Networks*: **3**. 507-523.

Byrne, J. H. & Gingrich, K. J. (1989). Mathematical model of cellular and molecular processes contributing to associative and nonassociative learning in *Aplysia.* In J. H. Byrne & W. O. Berry (Eds.), *Neural models of plasticity* (pp. 58-72). Orlando, FL: Academic Press.

Byrne, J. H., Gingrich, K. J. & Baxter, D. A. (1990). Computational capabilities of single neurons: relationship to simple forms of associative and nonassociative learning in *Aplysia.* In R. D. Hawkins & G. H. Bower (Eds.), *Computational models of learning* (Vol. 23 Psychology of learning and motivation) (pp. 31-63). New York: Academic Press.

Carew, T. J., Walters, E. T., & Kandel, E. R. (1981). Classical conditioning in a simple withdrawal reflex in *Aplysia californica. Journal of Neuroscience.***1**, 1426-1437.

Carpenter, G. A., & Grossberg, S. (1987a). A massively parallel architecture for a self-organizing neural pattern recognition machine. *Computer Vision, Graphics and Image Processing*: **37**, 54-115.

Carpenter, G. A., & Grossberg, S. (1987b). ART2: Self-organization of stable category recognition codes for analog input patterns. *Applied Optics*: **26**, 4919-4930.

Crick, F. (1989). The recent excitement about neural networks. *Nature*. **337**, 129-32.

Dretske, F. I.., (1988). *Explaining Behavior: Reasons in a World of Causes*. Cambridge, MA: MIT Press.

Elman, J. L. (1990). Finding structure in time. *Cognitive Science*: **14**, 170-211.

Elman, J. L. & Zipser, D. (1988). Discovering the hidden structure of speech. *Journal of the Accoustical Society of America*: **83**, 1615-1626.

Fodor, J. A. & Pylyshyn, Z. W. (1988) Connectionism and cognitive architecture: A critical analysis. *Cognition* **28**, 3-71.

Gelperin, A., Hopfield, J. J. & Tank, D. W. (1985). The logic of *Limax* learning. In A. I. Selverston (Ed.), *Model neural networks and behavior*. (pp. 237-261). New York: Plenum.

Getting, P. A. (1989). Emerging principles governing the operation of neural networks. *Ann. Rev. Neuroscience*. **12**, 185-204.

Gingrich, K. J. & Byrne, J. H. (1987). Single-cell neuronal model for associative learning. *Journal of Neurophysiology*. **57**, No. 6, 1705-1715

Grossberg, S. (1982). *Studies of mind and brain: Neural principles of learning, perception, development, cognition, and motor control.* Boston: Reidel Press.

Grossberg, S. (1987). Competitive learning: from interactive activation to adaptive resonance. *Coginitive Science*: **11**, 23-63.

Grossberg, S. (1991a). Nonlinear neural networks: principles, mechanisms, and architectures. In Carpenter, G. & Grossberg, S. (Eds.) *Pattern Rcognition by Self-organinzing Neural Networks* (36-109). Cambridge: The MIT Press.

Grossberg, S. (1991b). A neural network architecture for Pavlovian conditioning: reinforcement, attention, forgetting, timing. In M.L. Commons, S. Grossberg & J. E. R. Staddon, (Eds.), *Neural Network Models of Conditioning and Action* (pp. 1-11). Hillsdale, NJ: Lawrence Erlbaum Assoc.

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Acadamey of Sciences USA*, **79**, 2554-2558.

Hebb, D. O., (1949). *The Organization of Behavior.* New York: Wiley.

Hecht-Nielsen, R. (1986). Performance limits of optical, electro-optical, and electronic artificial neural system processors. In H. Szu (Ed.), *Hybrid and Optical Systems*, SPIE.

Hertz, J., Krogh, A. & Palmer, R. G. (1991). *Introduction to the Theory of Neural Computation.* Santa Fe Institute: Studies in the Sciences of Complexity, New York: Addison-Wesley Publishing Company .

Hull, C. L., (1952). *A Behavior System: An Introduction to Behavior Theory Concerning the Individual Organism.* New Haven: Yale University Press.

Jaffe, D. & Johnston, D. (1990). Induction of long-term potentiation at hippocampal mossy-fiber synapses follows a Hebbian rule. *Journal of Neurophysiology.* **64**, 948-960.

Kaelbling, L. P. (1993). Foundations of learning in autonomous agents. In *Toward Learning Robots.* W. Van de Velde (Ed.). Cambridge, MA: The MIT Press.

Kandel, E. R. & Schwartz, J. H. (1981). *Principles of Neural Science.* New York: Elsevier/North-Holland.

Kandel, E. R. & Schwartz, J. H. (1982). Molecular biology of learning: modulation of transmitter release. *Science*: 218, 433-443.

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

Klopf, A. H. & Morgan, J. S. (1990). The role of time in natural intelligence: implications of classical and instrumental conditioning for neuronal and neural network modeling. In M. Gabriel & J. Moore (Eds), *Learning and Computational Neuroscience*, (pp. 463 - 495). Cambridge, MA: The MIT Press.

Koshland, D.E. (1980). *Bacterial Chemotaxis As A Model Behavioral System*: New York: Raven Press.

Koch, C. & Segev, I. (1989). Introduction. In *Methods in neuronal modeling: from synapses to networks*. C. Koch & I. Segev (Eds.). Cambridge, MA: The MIT Press.

Kosko, B. (1986). Differential Hebbian learning. In J. S. Denker (Ed.), *AIP Conference Proceedings 151: Neural networks for computing* (pp. 277-282). New York: American Institute of Physics.

Kosko, B. (1991). *Neural Networks and Fuzzy Systems: A Dynamical Systems Approach to Machine Intelligence*. Englewood Cliffs, NJ: Prentice Hall.

Levy, W. B. & Desmond, N. L. (1985). The rules of elemental synaptic plasticity. In W. B. Levy, J. A. Anderson & S. Lemkuhle (Eds.), *Synaptic Modifications, Neuron Selectivity, and Nervous System Organization.* (pp. 105-121). Hillsdale, NJ: Lawrence Erlbaum Associates.

Lynch, G. & Baudry, M. (1985). The Biochemistry of Memory: A New and Specific Hypothesis. Chapter 26 in P. H. Abelson, E. Butz, & S. H. Snyder (Eds), *Neuroscience*. Washington, D.C.: The American Association for the Advancement of Science.

Lynch, G. Granger, R. & Larson, J. (1989). Some possible functions of simple cortical networks suggested by Computer modeling. In J. H. Byrne and W. O. Berry (Eds.), *Neural Models of Plasticity.* (pp. 329-362). San Diego, CA: Academic Press.

Mackintosh, N. J. (1983). *Conditioning and Associative Learning.* London: Oxford University Press.

McCulloch, W. S. & Pitts, W. (1943). A logical calculus of ideas immanent in nervous activity. *Bulletin of Mathematical Biophysiscs* **5**, 115-133.

Minsky, M. & Papert, S. (1969). *Perceptrons.* Cambridge, MA: MIT Press.

Moore, J. W. & Blazis, D. E. J. (1989). Simulation of a classically conditioned response: a cerebellar neural network implementation of the Sutton-Barto-Desmond model. In J. H. Byrne and W. O. Berry (Eds.), *Neural Models of Plasticity.* (pp. 329-362). San Diego: Academic Press.

Morgan, J. S., Patterson, E. C. & Klopf, A. H. (1990). Drive-reinforcement learning: a self-supervised model for adaptive control. *Network*: **1**, 439-448.

Pazzani, M. J. (1990). *Creating a Memory of Causal Relationships.* Hillsdale, NJ: Lawrence Elrbaum Associates.

Pavlov, I. P. (1927). *Conditioned reflexes.* London: Oxford University Press.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1985). *Learning internal representations by error propagation.* ICS Report 8506. Institute of Cognitive Science, University of California, San Diego.

Rumelhart, D. E., Hinton, G. E., & McClelland. (1986). A general framework for parallel distribute processing. In D. E. Rumelhart and J. L. McClelland (Eds.). *Parallel distributed processing: Explorations in the microstructure of cognition. Vol.* 1, Chapter 2. Cambridge, MA: The MIT Press.

Rumelhart, D. E.& McClelland, J. L. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition.* Vols. 1 & 2. Cambridge, MA: The MIT Press.

Sejnowski, T., Chattarji, S. & Stanton, P. (1989). Induction of synaptic plasticity by Hebbian covariance in the hippocampus. In R. Durbin, C. Miall, & G. Mitchison (Eds.) *The Computing Neuron.* New York: Addison-Wesley.

Sejnowski, T. J. & Qian, N. (1992). Synaptic Integration by Electro-Diffusion in Dendritic Spines. In T. McKenna, J. Davis & S. F. Zornetzer (Eds.) *Single Neuron Computation.* New York: Academic Press.

Shastri, L. & Ajjanagadde, V. (1993). From simple associations to systematic reasoning: A connectionist representation of rules, variables and dynamic bindings using temporal synchrony. *Behavioral and Brain Science.* **16,** 417-494.

Shepherd, G. M. (1992). Canonical Neurons and Their Computational Organization. In T. McKenna, J. Davis & S. F. Zornetzer (Eds.) *Single Neuron Computation.* New York: Academic Press.

Small, S. A., Kandel, E. R. & Hawkins, R. D. (1989). Activity-dependent enhancement of presynaptic inhibition in *Aplysia* sensory neurons. *Science*: **243**, 1603-1605.

Staddon, J. E. R. (1993). On Rate-Sensitive Habituation. *Adaptive Behavior* **1**, 4, 421-436.

Starmer, C.F. (1987). Characterizing synaptic plasticity with an activity dependent model. In M. Caudill & C. Butler (Eds.), *Proceedings of the IEEE First International Conference on Neural Networks* (Vol. IV, pp. 3-10). San Diego, CA: IEEE, Inc.

Sutton, R. S. (1988). Learning to predict by the methods of temporal differences. *Machine Learning*: **3**, 9-44.

Sutton, R. S. & Barto, A. G. (1981). Toward a modern theory of adaptive networks: expectation and prediction. *Psychological Review*. **88**, No. 2, 135-170.

Sutton, R. S. & Barto, A. G. (1987). *A temporal-difference model of classical conditioning*. (Tech Rep. No. 87-509.2) Waltham, MA: GTE Laboratories.

Sutton, R. S. & Barto, A. G. (1991). Time derivative models of Pavlovian reinforcement. In M. Gabriel & J. W. Moore (Eds.), *Learning and computational neuroscience*. Cambridge, MA: The MIT Press.

Tesauro, G. (1986). Simple neural models of classical conditioning. *Biological Cybernetics*: **55**: 187-200.

Turing, A.M. (1950). Computing machinery and intelligence. *Mind*. **59**, 433-460.

von Neumann, J. (1958). *The computer and the brain*. New Haven, CN: Yale University Press.

Widrow, B. & Hoff, M. E. (1960). Adaptive switching circuits. In *1960 IRE WESCON Convention Record*, part 4, 96-104. New York: IRE.

Winograd, T. & Flores, F. (1986). *Understanding computers and cognition*. Reading, MA: Addison-Wesley.