

# Toward a Modern Theory of Adaptive Networks: Expectation and Prediction

Richard S. Sutton and Andrew G. Barto  
Computer and Information Science Department  
University of Massachusetts—Amherst

Many adaptive neural network theories are based on neuronlike adaptive elements that can behave as single unit analogs of associative conditioning. In this article we develop a similar adaptive element, but one which is more closely in accord with the facts of animal learning theory than elements commonly studied in adaptive network research. We suggest that an essential feature of classical conditioning that has been largely overlooked by adaptive network theorists is its predictive nature. The adaptive element we present learns to increase its response rate in anticipation of increased stimulation, producing a conditioned response before the occurrence of the unconditioned stimulus. The element also is in strong agreement with the behavioral data regarding the effects of stimulus context, since it is a temporally refined extension of the Rescorla-Wagner model. We show by computer simulation that the element becomes sensitive to the most reliable, nonredundant, and earliest predictors of reinforcement. We also point out that the model solves many of the stability and saturation problems encountered in network simulations. Finally, we discuss our model in light of recent advances in the physiology and biochemistry of synaptic mechanisms.

One way to bridge the gap between behavioral and neural views of learning is to postulate neural analogs of behavioral modification paradigms. Hebb's suggestion that when a cell A repeatedly and persistently takes part in firing another cell B, then A's efficiency in firing B is increased, is the most familiar of these postulates (Hebb, 1949). This rule for synaptic plasticity is a neural analog of associative conditioning and continues to exert a powerful influence on theoretical and experimental research in learning and memory. Neural network models designed to explore the behavioral possibilities of modifiable structures typically em-

ploy a pre- and postsynaptic correlation rule for altering connectivities as a mathematical representation of Hebb's postulate (e.g., Anderson, Silverstein, Ritz, & Jones, 1977; Brindley, 1969; Grossberg, 1974; Kohonen, 1977; Marr, 1969; von der Malsburg, 1973). However, in addition to the fact that there is no direct experimental support for the Hebbian rule as a model of neural plasticity, several different bodies of evidence have accumulated which suggest that such simple contiguity rules can account neither for the behavioral facts of learning nor for the theoretical necessities of successful adaptation.

The analysis of elemental processes of learning has a long tradition within animal learning theory. To a large extent it has been successful: Fundamental laws of wide, if not complete, applicability have been found. Animal learning theory constitutes a large body of carefully explored and tested theories about fundamental processes of learning. Given this, it is surprising how little contact and interaction there has been between animal learning theory and adaptive systems theory, particularly insofar as the latter attempts to mimic neural networks or biological adaptive systems in general.

---

This research was supported by the Air Force Office of Scientific Research and the Avionics Laboratory (Air Force Wright Aeronautical Laboratories) through Contract F33615-77-C-1191. We are most grateful to A. H. Klopff, whose theories and guidance have contributed substantially to this research. We also thank R. Lara, J. W. Moore, and the reviewers for valuable commentary and criticism and S. Parker for essential help in preparing the manuscript.

Requests for reprints should be sent to Richard S. Sutton, Computer and Information Science Department, University of Massachusetts, Amherst, Massachusetts 01003.

Numerous adaptive systems papers have made brief reference to basic animal learning processes such as classical and instrumental conditioning. However, almost exclusively, inadequate models of these conditioning processes have been used, and in some cases they are so inadequate that although a theorist derives support for his model by citing a learning process, in reality the experimental evidence and modern learning theory contradict even the simplest predictions of the model. Classical conditioning involves an interplay between expectations and stimulus patterns that is too complex to incorporate into a simple correlation rule such as Hebb's. The common modifications of a correlation rule, for example, the introduction of delay in input or output pathways, result in model behavior still not in agreement with experimental data. Moreover, as we argue below, the phenomena actually observed in classical conditioning may perhaps be crucial for sophisticated adaptive behavior.

The history of attempts to construct adaptive networks of neuronlike components also suggests that something essential is not preserved by the Hebbian model and its variants. Network approaches to adaptive system design have been notable in their failure to produce learning behavior beyond a rather low level of sophistication. The information processing success of adaptive networks is restricted almost entirely to moderate success in the recognition, processing, and associative storage and retrieval of spatial patterns. There is a conspicuous absence of nontrivial processing of temporal patterns. It may well be true that in the brain some kinds of temporal patterns are processed by being represented spatially, as, for example, suggested by Lashley (1951), and some models use this principle (e.g., Fukushima, 1973; Grossberg, 1969; Spinelli, 1970). However, little progress has been achieved in our understanding of how a system can both learn and effectively use knowledge while interacting in real time with a complex environment. Yet these temporal aspects of a system's interaction with its environment are central to much intelligent behavior.

In the time since the first computational

experiments with adaptive networks were carried out, remarkable advances in the understanding of the cellular basis of behavior have occurred. In recent years invertebrate animals have been successfully used to study aspects of the neural basis of behavioral modifications (e.g., Kandel, 1976, 1978). Although this approach has not yet elucidated the cellular basis of associative learning, simpler but possibly related forms of nonassociative learning have been successfully analyzed at the cellular level. These studies reveal that neurons employ a wide variety of biochemical modulatory processes that interact in complex ways with electrical activity and that this interaction mediates forms of behavioral modification (Kandel, 1976, 1978).

Despite this evidence that neurons are capable of very complex information processing, adaptive network theorists continue to produce idealized neural element designs that are constrained by the early view that neurons are essentially switching elements, with little internal processing power. Although one of the most important aspects of model building is simplification, the lack of significant progress in adaptive network theory, together with the high complexity of cellular and synaptic machinery, suggests that these idealizations leave out some mechanisms that are essential for producing sophisticated adaptive behavior.

In this article we introduce an adaptive element model that is more reasonably in accord with the facts of modern animal learning theory than models commonly used in adaptive network research. After discussing several forms that adaptive element analogs of classical conditioning have taken in the past, we briefly introduce our model. We then present the basic elements of a view of classical conditioning that is more realistic than that commonly used in adaptive network studies. We show how the behavior of our model is in good agreement with a variety of aspects of animal learning data. We then discuss how our model relates to a variety of other adaptive elements that form part of adaptive system theory. No attempt is made to be exhaustive. Learning theory is a complex subject with many controver-

sies, and adaptive system theory is extremely diverse. We have tried to abstract from the very large animal learning theory literature those points on which there is a reasonable amount of agreement and that we consider to be most pertinent for adaptive network modeling and simulation.

Despite recent advances, it is still premature to propose a testable molecular model of associative learning. However, even though we see our model as being of interest primarily from behavioral and theoretical perspectives, we speculate as to how the cellular mechanisms that are beginning to be elucidated could implement the required computations. Our purpose in doing this is two-fold. First, we desire to demonstrate that processing of the proposed complexity is clearly possible at a cellular or simple network level. Second, some aspects of the proposed learning rule can be implemented so naturally by known mechanisms that a discussion of these mechanisms in light of our behavioral and theoretical observations, although speculative, may contribute to experimental efforts to understand neuronal plasticity.

Although we restrict attention in this article to classical conditioning, our research was motivated by an interest in more complex forms of learning and, in particular, the novel suggestion by Klopff (1979, in press, Note 1) that neurons may be reinforcement learning devices of a kind fundamentally different from those previously proposed in neural theories. The aspects of classical conditioning that we consider here form a necessary prelude to moving beyond the restrictions of the classical conditioning paradigm.

Finally, although our theory is an attempt to explore the consequences of attributing quite complex computational power to individual adaptive elements, it is not our intention to suggest that all of the mechanisms must necessarily reside in each element. Rather, our program of endowing a single adaptive element with behavior that has detailed properties of classical conditioning represents our feeling that these properties are fundamental to adaptive behavior. In particular, what we call an adaptive element may not correspond to a single neuron.

## Adaptive Element Analogs of Classical Conditioning

In a simple classical conditioning experiment, the subject is repeatedly presented with a neutral conditioned stimulus (CS),<sup>1</sup> that is, a stimulus that does not cause a response other than orienting responses, followed by an unconditioned stimulus (UCS), which reflexively causes an unconditioned response (UCR). After a number of such pairings of the CS and the UCS-UCR, the CS comes to elicit a response of its own, the conditioned response (CR), which closely resembles the UCR or some part of it. For example, a dog is repeatedly presented with first the sound of a bell (the CS) and then its food (the UCS), which causes the dog to salivate (the UCR). Eventually, the sound of the bell alone causes salivation (the CR).

In studies of the cellular basis of learning and in purely theoretical studies of adaptive systems, it is frequently convenient to postulate neuronlike mechanisms that embody various types of "learning rules." The rules describe how the strengths of interconnectivity change between units that are intended to be crude models of neurons. In keeping with this tradition, we shall sometimes refer to synapses, synaptic weights, and so forth, but the reader should remain mindful that the relationship between models of this form and neural plasticity is often one of coarse analogy. We prefer to think of the rules as describing the behavior of *adaptive elements*.

Figure 1 shows an element with input signals  $x_1, \dots, x_n$ ; connection weights  $w_1, \dots, w_n$ ; output  $y$ ; and a specialized "teacher" input  $z$ . Since in this article we wish to focus only on rules for changing the weights,  $w_i$ , we will not pay particular attention to the input-output function of the element. For our purposes, it suffices to say that  $y$  is some function of the weighted sum of the inputs, that is, for any time  $t$ ,

<sup>1</sup> Strictly speaking, this stimulus is not a conditioned stimulus until the animal has begun to be conditioned to it. However, as is often done, we simplify notation in this article by referring to any stimulus that is meant to be considered as eventually or potentially becoming conditioned as a conditioned stimulus.

$$y(t) = f\left[\sum_{j=1}^n w_j(t)x_j(t)\right], \quad (1)$$

where  $f$  is a function that may look something like the one shown in Figure 2.<sup>2</sup> Of course, when an adaptive element is proposed as an analog of animal learning, the form of this function becomes crucial in making precise predictions about behavioral data. In these cases, the function is related to response mapping rules (e.g., Frey & Sears, 1978). If the adaptive element is proposed as a neuron model, this function relates, for example, the firing frequency of a neuron to its membrane potential.

For an adaptive element analog of conditioning, the presence of  $CS_i$ ,  $i = 1, \dots, n$ , is indicated by activity on the corresponding input pathway  $x_i$ . For example, if  $x_i(t)$  denotes the signal on pathway  $x_i$  at time  $t$ , then the presence of  $CS_i$  at time  $t$  can be indicated by letting  $x_i(t) = 1$ . If  $CS_i$  is not present,  $x_i(t) = 0$ . The associative strength of each  $CS_i$  at time  $t$  is  $w_i(t)$ , the weight associated with pathway  $x_i$ . The CR is identified with the output  $y$  so that by Equation 1 the associative strengths of the  $CS_i$ ,  $i = 1, \dots, n$ , determine the magnitude of the CR. Learning rules take the form of equations for changing the values of the weights  $w_i$ ,  $i = 1, \dots, n$ , over time as functions of various aspects of the element's input and output. Usually the element's behavior is intended only to qualitatively resemble animal learning data.

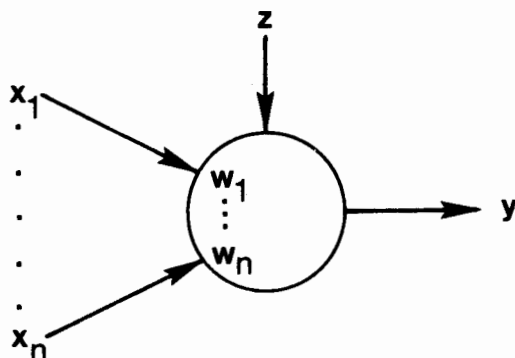


Figure 1. An adaptive element with  $n$  modifiable input pathways  $x_i$ ,  $i = 1, \dots, n$ ; connection weights  $w_i$ ,  $i = 1, \dots, n$ ; a specialized input  $z$  required by some adaptive elements to transmit the signals of a "teacher"; and an output labeled  $y$ .

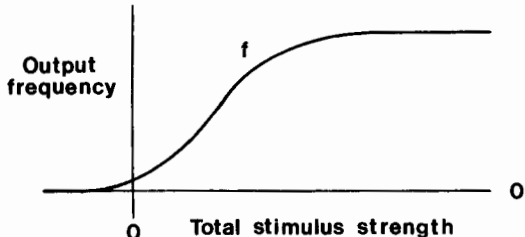


Figure 2. A common form of nonlinear input-output function used in neural and adaptive element models. (When these models are used in analogs of conditioning experiments, this function becomes a response mapping rule.)

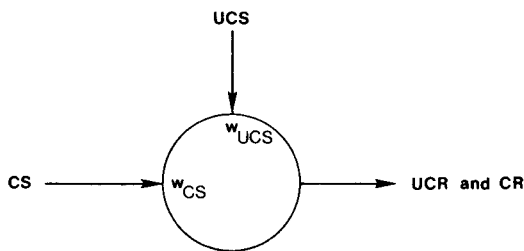
The most well-known example of an adaptive element analog of classical conditioning is based on Hebb's neural postulate that persistent pairing of pre- and postsynaptic activity increases a pathway's efficacy (Hebb, 1949). Although Hebb did not provide a mathematical formulation of this rule, the following expression has been widely used to implement his postulate:

$$w_i(t+1) = w_i(t) + cx_i(t)y(t), \quad (2)$$

where  $c$  is a positive constant determining the rate of learning. Here, and throughout this article, we use a time step of one unidentified unit that can be set equal to various values to suit particular interpretations of a model. For the case in which the input signals  $x_i$  and the output signal  $y$  are binary valued,  $w_i$  is incremented by  $c$  whenever an input pulse arrives and the cell fires and is unchanged otherwise. For the case of real valued signals,  $w_i$  becomes a rough measure of the correlation between input signal  $x_i$  and output signal  $y$ . Unlike several other rules, this rule does not require the specialized "teacher" input shown in Figure 1.

It is easy to see how a learning rule of the Hebbian kind can implement a simultaneous contiguity view of classical conditioning (Figure 3). Suppose a Hebbian adaptive element has an excitatory UCS input pathway

<sup>2</sup> According to Equation 1, the adaptive element computes its output  $y$  instantaneously from its inputs  $x_i$ . In order to remedy the problematic consequences of this when networks are considered, one can assume a small delay exists in the communication links between elements. For our present purposes we do not need to consider this detail.



*Figure 3.* A Hebbian element as an analog of classical conditioning. (The weight  $w_{UCS}$  associated with the UCS pathway is sufficiently large so that UCS occurrence causes the element to respond with the UCR. The weight  $w_{CS}$  of the CS pathway is initially too small for the CS alone to elicit a response but increases with repeated simultaneous pairing of the CS and UCS until the CS alone can elicit a response—the CR.)

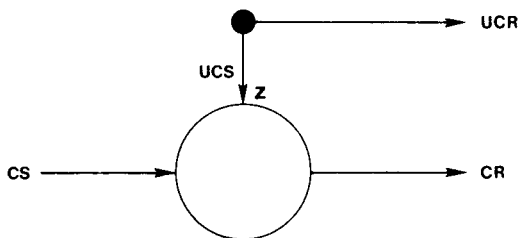
with weight  $w_{UCS}$  sufficiently large so that UCS occurrence causes the element to respond with the UCR. If the element also has an input pathway for the CS with an initially low weight  $w_{CS}$ , then after sufficient simultaneous pairing of the UCS and CS,  $w_{CS}$  will increase to a value at which the CS will elicit a response, the CR, in the absence of the UCS.

One reason the Hebbian rule has remained influential among theorists is that it provides a very simple hypothesis to account for a stimulus substitution view of classical conditioning. It is a common, though not universally accepted, theoretical position that in classical conditioning the CS comes to elicit a CR by effectively substituting for the UCS. This explains the similarity between the CR and the UCR, since it implies that the two responses occur via the same response pathway's being activated by two different stimulus pathways. This view, known as stimulus substitution theory, has proved to be a reasonable generalization from the data (see discussion and review in Mackintosh, 1974, pp. 100–109). In the Hebbian model of classical conditioning (Figure 3), the CR and UCR share the same pathway, so one would expect them to be similar.

Other adaptive element analogs of classical conditioning do not provide so natural an account of the similarity between CR and UCR because they require the UCS to be a specialized input to the adaptive element that does not excite it (Figure 4). In these

cases, separate pathways are required for the CR and UCR. To account for the similarity of the CR and UCR, it is necessary to postulate that the CR and UCR pathways converge in some manner “downstream” from the adaptive element. The “perceptron” of Rosenblatt (1962) and the “informon” of Uttley (1979) require this organization to form analogs of classical conditioning.

Aside from providing a simple explanation for the similarity of the CR and UCR, that the UCS is an unspecialized input in the case of the Hebbian element also means that the activity of any input pathway can cause changes in other pathways. In particular, pathways whose efficacies have become strengthened through previous training can further affect other pathways. A model with this property can produce behavior suggestive of higher order conditioning in animals: A previously conditioned CS can act as a UCS for a second CS. This property has also contributed to the interest in the Hebbian rule among theorists. It is not necessary to fix from the start the source of reinforcement. Any correlations among the input signals to an element will tend to be reflected in the connection weight values. The requirement for reinforcement to be provided only from a fixed source, on the other hand, raises the problem of somehow providing appropriate reinforcing signals at the appropriate times. The significance of this problem may be reflected in the lack of success in constructing powerful adaptive networks of per-



*Figure 4.* An adaptive element with a specialized UCS pathway. (Some adaptive element analogs of classical conditioning require a specialized UCS pathway that causes modifications in the CS pathway but does not have an excitatory effect on the element. This implies that the UCR and CR pathways are separate so that stimulus substitution does not occur at the element. Additional assumptions must be made to account for the similarity of the UCR and CR.)

ceptor elements (see Minsky & Papert, Note 2; Minsky & Selfridge, Note 3).

In this article we present a new adaptive element analog of classical conditioning that uses the stimulus substitution organization shown in Figure 3. We briefly introduce the model here and discuss it in detail in the following sections. In addition to the stimulus signals  $x_i$ ,  $i = 1, \dots, n$ , and the output signal  $y$ , our model requires the use of several other variables. First, for each stimulus signal  $x_i$ ,  $i = 1, \dots, n$ , we require a separate *stimulus trace*, which we denote by  $\bar{x}_i$ . By this we mean that the occurrence of  $CS_i$  at time  $t$ , indicated by  $x_i(t) = 1$ , initiates a prolonged trace given by nonzero values of a separate variable  $\bar{x}_i$  for some period of time after  $t$ . This is accomplished by letting  $\bar{x}_i(t)$  be a weighted average of the values of  $x_i$  for some time period preceding  $t$ . Similarly, we require a trace of the output  $y$ . Let  $\bar{y}(t)$  denote a weighted average of the values of the variable  $y$  over some time interval preceding  $t$ . In the computer simulations that produced the data shown below we generated these traces using the first-order linear difference equations

$$\bar{x}_i(t+1) = \alpha \bar{x}_i(t) + x_i(t) \quad (3)$$

and

$$\bar{y}(t+1) = \beta \bar{y}(t) + (1-\beta)y(t), \quad (4)$$

where  $\alpha$  and  $\beta$  are positive constants, and  $0 \leq \alpha, \beta < 1$ . Appendix A gives the values actually used in the simulations.

The behavior of the adaptive element is therefore described by the values over time of the two variables  $y$  and  $\bar{y}$  and the values of the three variables  $x_i$ ,  $\bar{x}_i$ , and  $w_i$  for each input pathway  $i = 1, \dots, n$ . In terms of these variables, the model takes the form of a set of difference equations for successively generating the values of the associative strengths: For each  $i$ ,  $i = 1, \dots, n$ ,

$$w_i(t+1) = w_i(t) + c[y(t) - \bar{y}(t)]\bar{x}_i(t), \quad (5)$$

where  $c$  is a positive constant determining the rate of learning.

We can describe the process given by Equation 5 as follows: Activity on any input pathway  $i$ ,  $i = 1, \dots, n$ , possibly causes an immediate change in the element output  $y$  but also causes the connection from that

pathway to become "tagged" by the stimulus trace  $\bar{x}_i$  as being eligible for modification for a certain period of time (the duration of the trace  $\bar{x}_i$ ). A connection is modified only if it is eligible and the current value of  $y$  differs from the value of the trace  $\bar{y}$  of  $y$ .

The effectiveness of the reinforcement for the conditioning process depends on the difference  $y(t) - \bar{y}(t)$ , which determines how the eligible connections actually change. The simplest case, and the one used in our simulations, results from letting  $\beta = 0$  in Equation 4 so that  $\bar{y}(t) = y(t-1)$ . Then  $y(t) - \bar{y}(t) = y(t) - y(t-1)$ , which is a discrete form of the rate of change of the variable  $y$ .

Our use of stimulus traces to create periods of "eligibility" was borrowed from the neural hypothesis by Klopff (in press, Note 1) that the temporal characteristics of conditioning, both classical and instrumental, can be produced if one set of conditions makes synapses eligible for modification of their transmission efficacies, but actual modifications occur due to other influences during periods of eligibility. This differs from related theories in that eligibility is seen as being indicated in some way completely separate from electrical activity. That is, instead of being marked as eligible for modification by a transient increase in efficacy or by prolonged presynaptic activation, a pathway would be marked by some mechanism that does not participate directly in the electrical signaling of the cell, such as a transient increase in the concentration of a particular chemical.

The weight change rule given by Equation 5 can be roughly understood by analogy with the Hebbian rule. Whereas the Hebbian rule detects correlations between input and output signals, this rule detects correlations between traces of input stimuli and changes in output. These differences have subtle and sometimes surprising consequences, which will be discussed in the following sections.

## Temporal Relationships

The use of the stimulus traces  $\bar{x}_i$  and the output trace  $\bar{y}$  in our model permits it to reproduce some of the intratrial temporal relationships between stimuli and responses

observed in classical conditioning experiments. Here we discuss interstimulus interval dependency and CR latency and review how earlier adaptive element models account for these aspects of classical conditioning. We then present simulation experiments which show that our model produces behavior in good agreement with experimental data.

We have said a pairing between the CS and the UCS is necessary for a classical conditioning association to form. In fact, many aspects of the temporal relationship between CS and UCS will affect the strength and rapidity of conditioning. The words *pairing* and *associative learning* commonly used in reference to classical conditioning seem to imply a symmetrical relationship between the CS and the UCS, and many theorists have created models in which associations are formed when CS and UCS (or their theoretical analogs) occur simultaneously. Experimentally, however, simultaneous presentation of CS and UCS typically results in very poor conditioning, if any (e.g., Smith, Colman, & Gormezano, 1969).

An effective pairing of CS and UCS in classical conditioning is not a symmetric one—the CS must occur first. The crucial variable with respect to the CS-UCS temporal relationship is the time interval between the onset of the CS and the onset of the UCS (the interstimulus interval, or ISI). Associative strength between the CS and the

CR is usually found to be an inverted-U-shaped function of this interval, being zero at simultaneous presentation, being maximal at intermediate values (that depend strongly on the particular response system), and then falling toward zero at longer ISIs. Conditioning for negative ISIs, or backward conditioning, is generally considered not to occur (see Mackintosh, 1974, pp. 58–60). Figure 5 shows an example of this relationship.

A second important aspect of the intratrial temporal relationships in classical conditioning is the time of occurrence of the CR relative to the CS and UCS. The time difference between CS onset and CR onset is called the *CR latency*. For a particular response there is usually a finite minimum value for the CR latency due to intrinsic delays of various kinds. For the nictitating membrane response, for example, the minimum CR latency is on the order of 70–80 msec. When the ISI is shorter than the minimum CR latency, then the CR necessarily occurs after UCS onset. In the more usual case in which the ISI is longer than the minimum CR latency, the CR begins before the UCS (Mackintosh, 1974, p. 61). Two examples are shown in Figure 6.

In Figure 6, Panel a, the CR begins nearly immediately after the CS just as the UCR begins nearly immediately after the UCS. However, in many experiments posttraining behavior much like that shown in Figure 6,

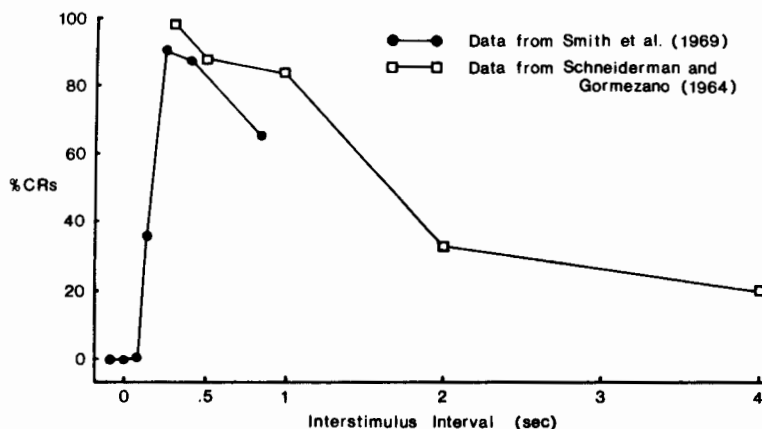


Figure 5. Asymptotic associative strength versus interstimulus interval in rabbit nictitating membrane response delay conditioning. (Optimal ISI times vary widely from small fractions of a second for some response systems to up to a minute and perhaps longer for others.)

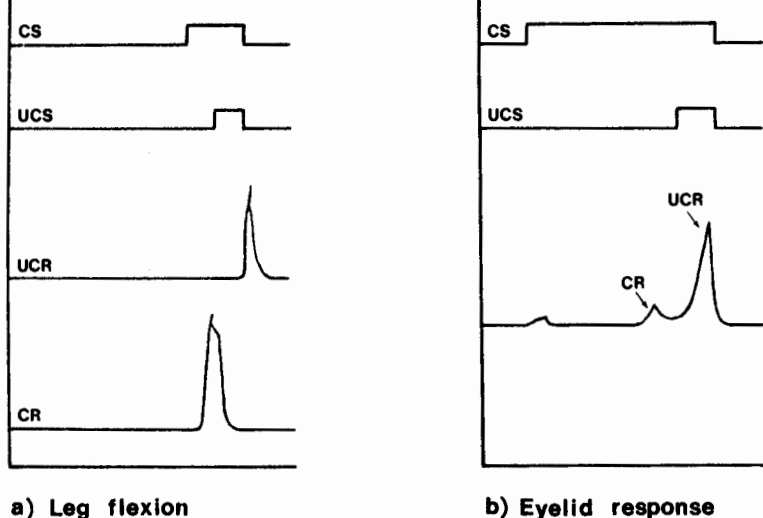


Figure 6. Tracings of CRs and UCRs in studies of leg flexion and eyelid conditioning. (In each case CR onset occurs before UCS onset. Panel a shows leg flexion CR and UCR in dogs [after Kellogg, 1938]. Panel b shows eyelid CR and UCR in a human subject [after Hilgard, 1936].)

Panel b, is observed, in which the CR begins much later than a minimum CR latency after CS onset. This phenomenon, known as inhibition of delay, appears to be the result of the animal's discriminating between earlier and later parts of the CS and treating them as different CSs. The CR initially begins soon after the overt CS onset and then gradually shifts to later with continued training (Kimmell, 1965). This shifting is made more rapid by increasing the discriminability of earlier and later parts of the CS. In these cases also the CR is experimentally found to precede the UCS. Summarizing we can state: Except in the case of an ISI less than the minimum CR latency, a classically conditioned CR will begin before its UCS (Mackintosh, 1974, p. 61).

It is on the basis of these temporal relationships that we say that the CS is a *predictor* of the UCS and the CR is a *prediction* of the UCS. Many learning theorists (e.g., Dickinson & Mackintosh, 1978; Kamin, 1969) have emphasized the importance of the CS's being an informative predictor of the UCS rather than just occurring appropriately paired with the UCS. To this we add that in order for the predictive information made available by the CS to be useful, it must be available before the event predicted. This suggests that the fact that

the CR usually begins before the UCS in classical conditioning may be an important aspect of the classical conditioning behavior.

However, not one of the adaptive element models currently in the literature is capable of producing behavior whose temporal structure is in agreement with that observed in animal learning as described above. It is usual practice to add additional mechanisms, such as a delay in the CS pathway, in order to account for some of the temporal relationships between stimuli and responses. In most cases, however, the resulting adaptive elements display only superficial aspects of this temporal structure.

### Delays

As a first step it is important to understand what can and what cannot be achieved by the addition of delays in input or output pathways of elements requiring simultaneous pairing for changing weights. Consider the two different ways of using adaptive elements to model classical conditioning that we have described with reference to Figures 3 and 4. These models differ in that the latter have a specialized UCS pathway and a UCR pathway that is different from the CR pathway.

First consider the consequences of adding a delay in the CS pathway in either type of



model (Figure 7, Panel a). When the delayed CS and the UCS temporally overlap, the associative strength of the CS increases. This means that maximal learning occurs when the UCS follows the CS by the time of the delay, thus exhibiting a rough form of the experimentally observed ISI dependency. Suppose now that conditioning continues until the CS elicits the CR. Since the CS is delayed, the CR is also delayed, so that it cannot begin earlier than the UCS, that is, the CR latency is always greater than or equal to the ISI. The delay in the CS pathway necessarily also delays the CR, thus preventing it from being a useful prediction.

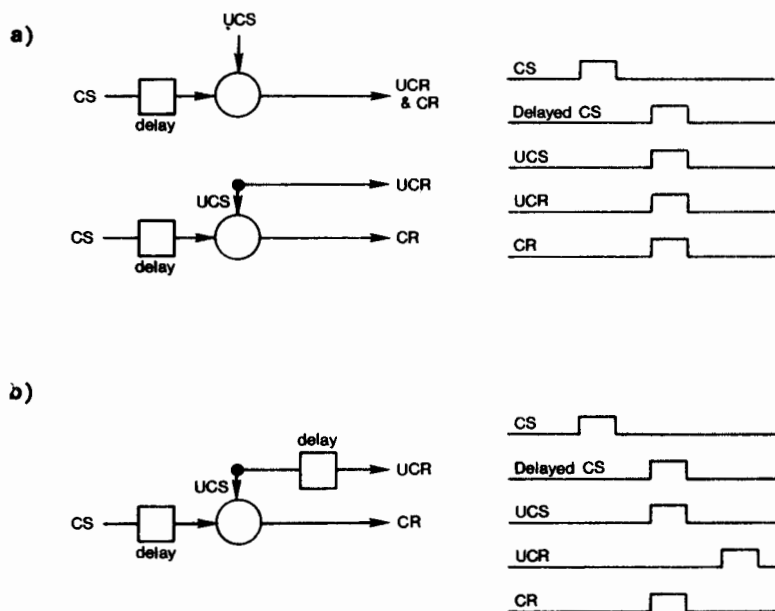
For the case in which there are separate pathways for the CR and the UCR (Figure 4), one can consider adding delays to both the CS and the UCR pathways, as is done, for example, by Uttley (1975). In this case, the CR cannot begin earlier than the UCS for the same reason discussed above, but it can begin earlier than the UCR due to the delay in the UCR pathway (Figure 7, Panel b). However, in classical conditioning it is the UCS that is anticipated by the CR. That is, an animal can predict stimuli by becoming sensitive to external signals that regu-

larly precede those stimuli. Merely producing a response earlier than it previously appeared, but not before the previously eliciting stimulus, simply results in increased speed of response. This is indeed a useful strategy, but it can be accomplished more simply by reducing the delay in the UCR pathway. In classical conditioning, on the other hand, a response can begin earlier than the occurrence of the stimulus that previously elicited it. This, of course, requires the availability of predictive information in the environment (a CS).

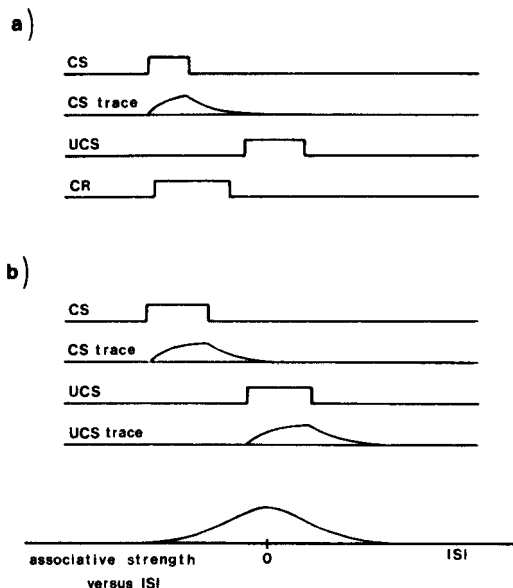
For elements requiring simultaneous pairing of stimuli for forming associations, no combination of simple delays in the CS, UCS, and UCR pathways can produce this kind of anticipatory response. The delays essentially just slow the system down. In addition, delays on the order of seconds or even longer required for this approach are very hard to justify neurophysiologically.

### Stimulus Traces

The notion that a stimulus sets up an internal neural trace that persists after the stimulus ends has a long history in theories



**Figure 7.** The use of delays in attempts to approximate the temporal relationships observed in classical conditioning. (Panel a shows that a delay in the CS pathway of both types of classical conditioning models necessarily also delays the CR. Panel b shows that delays in the CS and UCR pathways permit the CR to precede the UCR but not the UCS.)



**Figure 8.** Stimulating stimulus traces. (Panel a shows that if the CS initiates a prolonged stimulating trace and the UCS does not, then the CR can anticipate the UCS, but the CR will tend also to be prolonged unless some additional mechanism is postulated. Panel b shows that if both the CS and the UCS initiate traces which stimulate one of the adaptive elements described in the text, then there will always be backward conditioning. Shown here is an ISI dependency curve for the case in which the CS and UCS produce identically decaying exponential traces.)

of learning, notably in Hull's (1943), and has been used in neural network theories, as, for example, by Grossberg (1974). Although a simple delay is one form of stimulus trace, the kind of trace to be considered now is one that unlike a delayed signal, persists in some form throughout the temporal interval. In particular, such a trace is present in the interval's early as well as late portions. There are two general classes of possibilities for stimulus trace mechanisms: (a) traces are maintained by the firing levels of some neurons, possibly by means of reverberatory circuits, and (b) they are maintained by something other than neuronal electrical activity, perhaps by chemical concentrations. From our theoretical point of view, the most important difference between these two possibilities is that the former employs the same means for storing traces as is used for signaling stimuli and producing responses. In

the latter case, these two functions are performed by separate mechanisms. The first type of trace, which we call a stimulating trace, is more frequently hypothesized, and we discuss this possibility first.

Suppose the CS gives rise to a stimulating trace that persists long enough to span the interval between CS and UCS presentations (Figure 8, Panel a). If this trace serves as the CS input to an adaptive element requiring simultaneous pairing, and the UCS does not produce such a trace, one can obtain an ISI dependency curve whose shape resembles that of the stimulus trace function.<sup>3</sup> If the UCS leaves a similar stimulus trace that acts as input to the adaptive element, then the ISI dependency curve shows substantial learning for negative CS-UCS intervals, that is, for cases in which the UCS onset precedes the CS onset (Figure 8, Panel b). Uttley (1975) suggests the use of a long CS trace and a short UCS trace in order to minimize (but not eliminate) backward conditioning.

A stimulus trace consisting of a prolonged CS signal does permit the CR to anticipate the UCS, since the signal trace, unlike a delayed signal, is present at the beginning as well as the end of the ISI. For example, if we assume that an element produces a response whenever the weighted sum of its input signals exceeds a threshold, then after sufficient training the CS will elicit a CR whenever the CS trace, multiplied by the connection weight of the CS pathway, exceeds the threshold (Figure 8, Panel a). As training continues one would expect the duration of the CR to lengthen as longer intervals of the stimulus trace exceed threshold. Although various characteristics of the CR change as training continues, there are

<sup>3</sup> Hull (1943) apparently believed that an experimental ISI curve could be accounted for by assuming a neural trace of the same shape. As Hilgard and Bower (1975) point out, however, level of conditioning is such a complex function of the ISI along with many other factors that this form of explanation is untenable. It should be noted, though, that since we are discussing adaptive elements out of which adaptive systems can presumably be constructed, this objection holds less force, since externally observed behavior is likely to be a product of the interaction of a variable mixture of local traces.

no data indicating a tendency for the CR to persist throughout the ISI: The CR generally resembles the UCR. Some additional mechanism would have to be postulated to prevent the prolonged stimulating trace from being manifested in overt behavior as a prolonged response.

### *Nonstimulating Traces*

We now consider what one would expect if the stimulus trace were provided by a signal different from the stimulating signal. Several proposed mechanisms fall into this category. It has been suggested, for example, that a stimulus might leave a temporarily persistent trace in the form of an altered threshold of the postsynaptic element (Milner, 1957; Rosenblatt, 1962, p. 55) or that a transient increase in synaptic efficacy follows presynaptic activity and is made more permanent by subsequent firing of the postsynaptic cell (Rosenblatt, 1962, p. 57).

The use of a stimulus trace variable entirely separate from the major signaling variable has been proposed by Klopff (1979, in press, Note 1). He suggests that when activity at a synapse satisfies certain criteria, then that synapse becomes eligible for modification and remains eligible for a period of several seconds. The extent to which an eligible synapse is modified depends on the reinforcement level during the period of eligibility. Each synapse is therefore viewed as possessing its own local trace mechanism, which mediates synaptic modification but does not directly alter any other aspect of the unit's behavior. Such a trace can persist, as Klopff suggests, for the relatively very long times suggested by classical (and instrumental) conditioning data without interfering with ongoing signal transmission. Further, the large variation in ISI dependency for different response systems might be accounted for by variations in eligibility trace durations. This is the kind of stimulus trace provided by the term  $\bar{x}_i$  in our model.

Our model implies that a synapse becomes eligible for modification whenever a presynaptic signal occurs there and that eligibility forms a curve like that of the trace in Figure 8, Panel a (see also Figure 10). Since learn-

ing occurs due to an interaction between the UCS signal and a nonstimulating eligibility trace initiated by the CS, important temporal aspects of classical conditioning can be produced. In particular, the CR will begin immediately after the CS onset and, unlike the case of a stimulating trace, the CR will not extend in duration as conditioning proceeds (see below). This is possible because the trace is different from the stimulating signal.

Although either stimulating or nonstimulating traces might be postulated to account for temporal aspects of classical conditioning, a nonstimulating trace has the advantage of permitting a clear distinction to be maintained between actual stimuli and traces of stimuli. There are two countervailing requirements that need to be met. First, fast electrical signals are necessary to indicate as precisely as possible the time of occurrence of specific events. It is to an organism's advantage to perceive events as occurring as closely as possible to their actual time of occurrence, and particularly as early as possible. Second, it is necessary to retain the knowledge of these occurrences so that they can be associated with later events. In a two-variable system, these two requirements are both satisfied, whereas in a single-variable system, such as one using reverberatory activity, one of these requirements can only be satisfied at the expense of the other. If the association of events depends on their precise temporal relationship, as indeed it appears to, then we can expect there to be a high priority on precise temporal localization of events. Thus, it seems most reasonable not to confuse the need for a short, distinct signal with the need for a prolonged trace by using a single trace for both purposes.

A common argument for a reverberatory activity theory is based on certain studies of attention and distraction and their effect on learning. These studies indicate that reverberatory activity is probably important in the central nervous system. However, this we do not mean to debate. Reverberatory activity can be expected to play an important role—for example, it can determine what information is picked up or relayed to higher centers. We believe it is unwarranted, how-

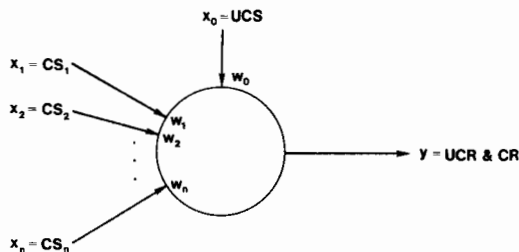


Figure 9. Our adaptive element as an analog of classical conditioning. (There are  $n$  modifiable CS input pathways and a pathway with fixed weight  $w_0$  that carries the UCS. The element output  $y$  represents both the UCR and the CR.)

ever, to proceed from this to the conclusion that reverberatory activity is the primary mechanism for spanning the time between the sequential events on which learning is contingent.

### Model Behavior in Classical Conditioning With a Single CS

One uninteresting steady state of our model occurs when all the connection weights are equal to zero. In this case  $y$  remains at zero so that no modifications to the weights can occur. A simple way to exclude this steady state is to set at least one weight to a fixed nonzero value. In an analog of classical conditioning, this fixed input pathway carries the UCS, and the resultant effect on the element is the UCR (see Figure 9).

It is useful to consider the simplest special case of a single rectangular CS signal that ends when the UCS starts. The discussion is also simplified if we assume that the UCS is sufficiently long so that all synapses have lost their eligibility by the time of its offset. Figure 10 shows this CS and the eligibility it generates as well as a UCS and the reinforcement signal generated. We have assumed that  $w$ , the associative strength of the CS, is initially equal to zero and that the term  $\bar{y}$  takes the simplest form,  $\bar{y}(t) = y(t-1)$ , resulting from letting  $\beta = 0$  in Equation 4. This makes  $y - \bar{y}$  a rough form of the derivative of  $y$ . The rectangular CS signal causes an increase in the eligibility of the CS pathway that persists for some time after the CS offset. The rectangular UCS signal, active through a fixed excitatory input of strength  $\lambda$ , causes a positive change

in  $y$  at its onset and an equal but negative change at its offset. The weight, or associative strength, of the CS experiences a net increase: At the UCS onset it increases by a certain amount and decreases by a lesser amount at the UCS offset (in this case the decrease is zero, since the eligibility has decayed to zero by the time of UCS offset).

After one trial  $w$  is positive so that on the next trial the occurrence of the CS increases the output level  $y$ . Consequently, CS onset causes a transient increase in  $y - \bar{y}$ , which has no effect on the CS pathway, since CS pathway eligibility is zero at CS onset. However, the level of  $y$  is raised by the CS so that UCS occurrence causes less of an increase in  $y$  than it did on the preceding trial. This means that the value of  $y - \bar{y}$  at the time of UCS onset causes a further increase in  $w$ , but one of smaller magnitude than in previous trials. With additional trials this process continues until the value of  $y - \bar{y}$  at the UCS onset is equal to zero, that is, until the CS produces activity equal to that produced by the UCS (Figure 11). Growth in

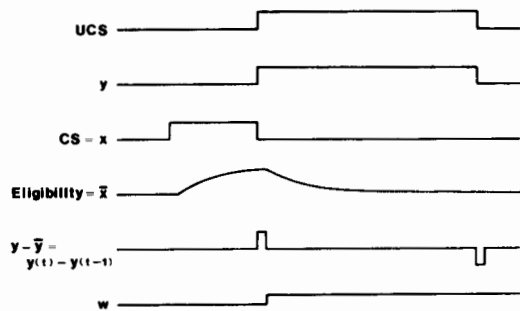


Figure 10. Time courses of element variables for a trial in which a neutral (associative strength  $w = 0$ ) CS is followed by a UCS. (For ease of explanation CS offset and UCS onset coincide and the UCS is of sufficient duration so that  $\bar{x}$  is zero at UCS offset. The trace  $\bar{x}$  of the CS signal  $x$  indicates the eligibility for modification of the CS pathway. This trace increases during CS presentation and persists after CS offset. Element output  $y$  shows no change during CS presentation, since  $w = 0$ , but since the UCS stimulates the element via a fixed positive weight, the shape of the time course of  $y$  follows that of the UCS signal. This causes  $y - \bar{y}$  to indicate UCS onset with a positive pulse and UCS offset with a negative pulse. The CS associative strength  $w$  changes according to the product of  $\bar{x}$  and  $y - \bar{y}$ . Consequently,  $w$  increases at UCS onset and decreases by a lesser amount [here, by zero] at UCS offset, thus experiencing a net increase.)

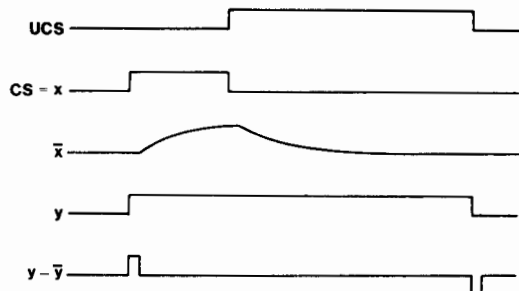


Figure 11. Time courses of element variables after the asymptotic CS associative strength has been reached due to a series of trials. (Element output  $y$  changes at CS onset, since  $w$  is now positive. UCS onset causes no additional increase in  $y$  over the level produced by the CS. The CS pathway eligibility  $\bar{x}$  is zero for the positive pulse of  $y - \bar{y}$  and, assuming a sufficiently long UCS, also zero for the negative pulse. Under these circumstances  $w$  does not change.)

associative strength therefore is negatively accelerated and stops when  $y$  remains constant during CS pathway eligibility. Figure 14, Trials 0–10, shows the form of the acquisition curve produced by computer simulation.

The equilibrium reached after a number of trials and shown in Figure 11 has the following important properties. First, the CS has an excitatory effect on the adaptive element when the effect of the UCS is also

excitatory. This permits a stimulus substitution model of classical conditioning in which the CR and UCR share the same pathway (Figure 3). Second, the CR produces an output level  $y$  of magnitude equal to that produced by the UCS, and third, the CR begins earlier than the UCS. The element increases its output level in anticipation of UCS occurrence.

Similar behavior is produced when UCS onset precedes or follows CS offset by some time interval or when the eligibility trace outlasts the UCS. In these cases, however, the CR will differ in magnitude from the UCR in a manner depending on the precise temporal arrangement of the CS and UCS. In addition, the equilibria in these cases are dynamic rather than static. The CS associative strength continues to change during each trial, but eventually there is zero net change per trial. The behavior approaches a stable limit cycle. Appendix B contains a related formal analysis.

Figure 12 shows the resultant asymptotic connection weight for a series of simulation experiments in which the time interval between CS onset and UCS onset is varied. The connection weight becomes the strongest when the CS ends just as the UCS begins (ISI = 3 time steps). At ISIs less than 3 time steps there is less time for the eligibility of

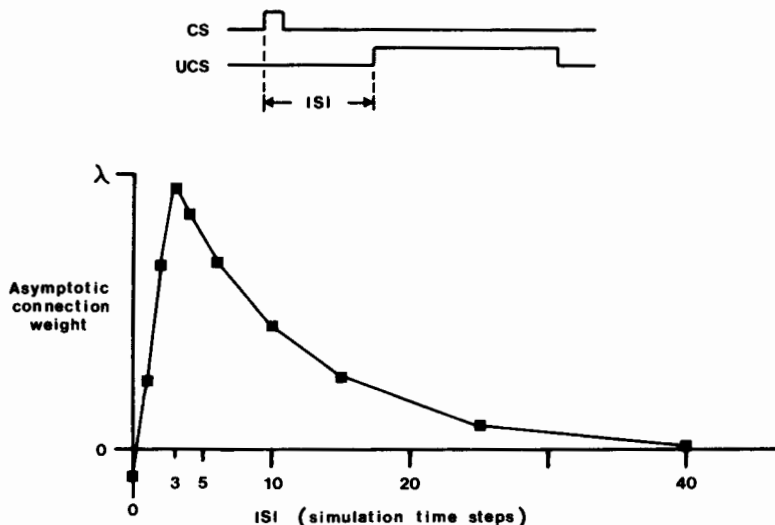


Figure 12. Asymptotic connection weight versus interstimulus interval in a simulated classical conditioning paradigm. (The interstimulus interval [ISI] was varied between 0 and 40 time steps, CS length was 3 time steps, and UCS length was 30 time steps.)

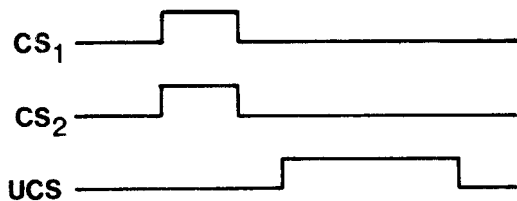


Figure 13. Temporal relationship between stimuli in the second part of a standard blocking experiment. (Learning to each of the component stimuli  $CS_1$  and  $CS_2$  will depend on the associative strength of the other component stimulus.)

the CS pathway to increase before the arrival of the UCS. At ISIs greater than 3 intervals, the eligibility decays toward zero, since the CS is not present for some interval between CS offset and UCS onset. These results have the same overall form as those observed in animals.

However, in animal experiments optimal ISIs are not so strongly tied to overt CS duration, although longer optimal ISIs have been observed for long fixed delay CSs than for short trace CSs (Schneiderman, 1966). The behavior of our adaptive element can be reconciled with the experimental observations if it is assumed that "effective," or "internal," CS duration is not identical to overt, external CS duration. A long CS is ignored shortly after it begins, whereas even an instantaneous overt CS causes an internal representation of some significant duration. This internal duration, rather than overt CS duration, then, would determine optimal ISI.

Behavior similar to that discussed above is produced by our model if  $\bar{y}$  is a more prolonged trace than that used for the preceding discussion. Letting  $\beta$  be nonzero (but still less than one) in Equation 4 results in an exponentially decaying trace  $\bar{y}$  similar to the eligibility trace  $\bar{x}$ . In this case, the term  $y - \bar{y}$  used in our model is a measure of the deviation of the current output level from an average of past values. The low-pass filtering characteristic of this measure prevents high-frequency fluctuations in  $y$  from significantly influencing the associative strengths. Equation 4 implies that for any  $\beta$ ,  $0 \leq \beta < 1$ , if  $y$  remains constant over time, then  $y - \bar{y}$  will approach zero, thus providing for deceleration of the learning process in a manner qualitatively similar to that produced when  $\beta = 0$ .

These illustrations of our model's behavior show that it is sensitive to the temporal relationships between stimuli within classical conditioning trials and is capable of producing CRs that begin before the UCS onset. It is evident from our discussion of how these properties follow from Equations 3, 4, and 5 that considerable behavior subtlety can be generated by the interaction of eligibility traces and a measure of output change. In general, the quantitative aspects of our model's behavior depend on the timing, durations, and shapes of the CS and UCS signals, the forms of the eligibility traces  $\bar{x}_i$  and the output trace  $\bar{y}$ , and the character of the output mapping function  $f$ . This complex of dependencies provides considerable latitude for making quantitative predictions about particular response systems, and we restrict our attention in this article to the qualitative aspects of our model's behavior. Appendix B contains a mathematical analysis of some of these dependencies for a simplified version of our model.

### Context and Expectation

Another aspect of classical conditioning that should be included in even a very simple theory is the effect of the context of a CS. The associative strengths of the stimuli that act as context for a CS on a trial can nullify or even reverse the effect of the occurrence of the UCS on that trial. This can be seen in numerous experimental paradigms, of which the simplest is known as blocking.

In blocking, as in all stimulus context experiments, a compound stimulus consisting of at least two stimulus components (one of which is frequently thought of as a conglomerate background stimulus component) is used as a CS. In Part 1 of a typical blocking experiment, one stimulus component  $CS_1$ , which might be a light, is paired with a UCS at an appropriate ISI until associative strength between  $CS_1$  and the CR reaches its asymptotic value. In Part 2, the experimenter continues to pair  $CS_1$  with the UCS but also pairs  $CS_2$ , say a bell, with identical temporal relationship as diagrammed in Figure 13. In effect, the compound stimulus  $CS_1 + CS_2$  is being paired with the UCS.

The result of this procedure is that  $CS_2$ , which is appropriately paired with the UCS

in Part 2, conditions very poorly, if at all, compared to a control group without Part 1 conditioning to CS<sub>1</sub>. This is not an isolated result. Effects of the associative strengths of context stimuli on conditioning occur in a great variety of experimental paradigms, in both classical and instrumental conditioning, of which blocking, overshadowing, and conditioned inhibition are only some of the more prominent examples (see Hilgard & Bower, 1975, pp. 571-573). Context stimuli can have such large effects on resultant associative strength that they cannot satisfactorily be ignored by a nontrivial theory of classical conditioning.

The simplest and most successful theory describing the effects of stimulus context is generally considered to be that of Rescorla and Wagner (1972). They state their theory in cognitive terms as follows:

Organisms only learn when events violate their expectations. Certain expectations are built up about the events following a stimulus complex; expectations initiated by the complex and its component stimuli are then only modified when consequent events disagree with the composite expectation. (p. 75)

Applying this analysis to the blocking experiment, Part 1 builds up an expectation that the UCS will follow CS<sub>1</sub>. The events of Part 2 do not violate this expectation, so there is no learning. Other stimulus context effects can be dealt with in similar fashion. However, similar ideas have been advanced by others. What distinguishes Rescorla and Wagner's theory is that it is given a precise mathematical form:

$$\Delta V_A = \alpha_A \beta [\lambda - V_{AX}], \quad (6)$$

where  $\Delta V_A$  is the change in associative strength to a CS  $A$ ,  $\lambda$  is the asymptotic value of associative strength possible with the UCS,  $V_{AX}$  is the associative strength already present to the stimulus complex  $A + X$ , where  $X$  is a conglomerate background stimulus, and  $\alpha_A$  and  $\beta$  are positive constants depending respectively on the CS being changed ( $A$ ) and the particular UCS used. Implicit here is that Equation 6 is only applied to a CS  $A$  if it is present on the trial and that the complex  $A + X$  is precisely all stimuli present on the trial. Using the simplest assumption that  $V_{AX} = V_A + V_X$ , with  $c = \alpha_A \beta$ , and letting  $S$  be the set of (indexes

of) all stimuli present on a trial, Equation 6 can be written as

$$\Delta V_{CS_i} = \begin{cases} c[\lambda - \sum_{j \in S} V_{CS_j}] & \text{for } i \in S \\ 0 & \text{for } i \notin S. \end{cases} \quad (7)$$

Part 1 of the blocking experiment causes  $V_{CS_1}$  to reach the value  $\lambda$  because CS<sub>1</sub> is the only stimulus present. In Part 2,  $V_{CS_2} = 0$  initially, and since

$$\sum_{j \in S} V_{CS_j} = V_{CS_1} + V_{CS_2} = \lambda + 0 = \lambda,$$

no changes in associative strength take place. It should be clear how this equation implements Rescorla and Wagner's cognitive theory expressed in the quotation cited above: The expectations that are built up are the associative strengths, and these are modified when events such as the UCS, represented by  $\lambda$ , differ from the composite expectation (the summation of associative strengths of present stimuli).

This theory can account for blocking and a wide range of the other stimulus context effects. The theory is not a completely satisfactory one, the two most prominent and best established shortcomings being the following: (a) There has been repeated failure to demonstrate the extinction of conditioned inhibitors predicted by the Rescorla-Wagner model (the return to zero of negative associative strengths when their stimuli occur without any correlation to the UCS; Zimmer-Hart & Rescorla, 1974). (b) The strict application of the Rescorla and Wagner equation requires the prediction of a strictly negatively accelerated acquisition curve. The consensus is that this curve is initially positively accelerating (Mackintosh, 1974, p. 11).

The Rescorla-Wagner theory also does not correctly predict the microstructure of individual response sequences (Prokasy & Gormezano, 1979). Recent extensions to the Rescorla-Wagner model have been proposed to remedy some of these problems (Frey & Sears, 1978).

Our adaptive element uses a form of expectation closely related to that of the Rescorla-Wagner model. Whereas in that model the associative strengths are changed based on the difference between received and expected UCS levels, in our model weights are changed based on the difference between

actual activity level  $y$  and expected activity level  $\bar{y}$ . In fact, our model results in all the stimulus context behavior of the Rescorla-Wagner model.

This can be seen most clearly by considering another special case. Assume that there are many CS pathways, on which rectangular pulse CSs may or may not be present, and that all CSs present on a trial begin simultaneously (and, as before, end as the UCS begins). If the UCS signaled by  $x_0$  begins at time  $T$  and has a duration longer than the eligibility traces, then the connection weight  $w_i$ , corresponding to  $CS_i$ , can only change at  $T$ . This is the only time at which  $y$  changes when an input pathway can be eligible. Then the total change in  $w_i$  on a particular trial is  $\Delta w_i(T)$ .

From Equation 5 we have:

$$\Delta w_i(T) = c[y(T) - \bar{y}(T)]\bar{x}_i(T).$$

Taking the simplest case  $\bar{y}(t) = y(t-1)$ , and  $\bar{x}_i(t) = x_i(t-1)$ :

$$\Delta w_i(T) = c[y(T) - \sum_{j=0}^n w_j x_j(T-1)] \times x_i(T-1).$$

Letting  $y(T) = w_0 x_0(T) = \lambda$ , and noting that  $x_0(T-1) = 0$ , we obtain

$$\Delta w_i(T) = c[\lambda - \sum_{j=1}^n w_j x_j(T-1)] x_i(T-1).$$

And since  $x_j(T-1) = 1$  indicates CS presence, we can write

$$\Delta w_i(T) = \begin{cases} c[\lambda - \sum_{j \in S} w_j] & \text{for } i \in S \\ 0 & \text{for } i \notin S, \end{cases}$$

where  $S$  is the set of stimuli present on the trial. Since  $\Delta w_i(T)$  is the total change in connection weight on the trial, this result is identical to the Rescorla-Wagner equation (Equation 7).

Computer simulations illustrate this result in a variety of standard stimulus context experiments. The results of a computer simulation of our model in a blocking experiment is illustrated in Trials 0–20 of Figure 14. For the first 10 trials of the simulation experiment,  $CS_1$  is presented alone and followed by the UCS as discussed earlier. The con-

nection weight  $w_1$  of  $CS_1$  quickly rises to the UCS level  $\lambda = .6$  (see Figure 14, Trials 0–10; Figure 11 shows the steady state element behavior; additional details on the simulations are in Appendix A). The acquisition curve is purely negatively accelerated as in the Rescorla-Wagner theory.

For Trials 11–20,  $CS_1$  is presented identically paired with  $CS_2$ , and both are followed by the UCS. This is the blocking paradigm. Since it provides no new information about UCS arrival,  $CS_2$  is redundant. During these trials,  $w_1$  and  $w_2$  do not change, as shown in Figure 14. This result can be understood by examining the behavior of the relevant element variables during one of these trials (Figure 15). The decrease in  $y$  occurs too long after the occurrence of the CSs for them to still be eligible, and the increase in  $y$  occurs just as the CSs begin, and thus before they are eligible.

Elements that implement the Rescorla-Wagner equation find input signals whose presence is associated with the UCS and that are not redundant. Each such signal generates an expectation equal to the additional UCS magnitude indicated by its presence. If there are many signals, the sum of their expectations is of appropriate magnitude. For example, if the compound stimulus  $CS_1 + CS_2$  is paired with a UCS of strength  $\lambda = .6$  while  $CS_1$  alone is concurrently paired with a UCS of strength  $\lambda = .4$ , then the two associative strengths (connection weights)  $w_1$  and  $w_2$  will stabilize at .4 and .2, respectively (assuming  $x_i = 1$  indicates  $CS_i$  present; in general,  $w_1 x_1$  and  $w_2 x_2$  will stabilize at .4 and .2). A simulation experiment confirmed this conclusion for the adaptive element we have introduced (Figure 16).

Elements that implement the Rescorla-Wagner equation also have a tendency to find the input pathways whose activity is most reliably associated with the UCS and to ignore all others. For example, let  $CS_1$  be paired with 100% of the UCSs while  $CS_2$  is paired with only 75% of the UCSs. Even if  $CS_2$  is initially dominant in terms of associative strength ( $w_2 = \lambda$ ,  $w_1 = 0$ ), eventually  $CS_1$  becomes completely dominant ( $w_1 = \lambda$ ,  $w_2 = 0$ ). This result contrasts strongly with the blocking experiment in which equally reliable CSs do not change their dominance



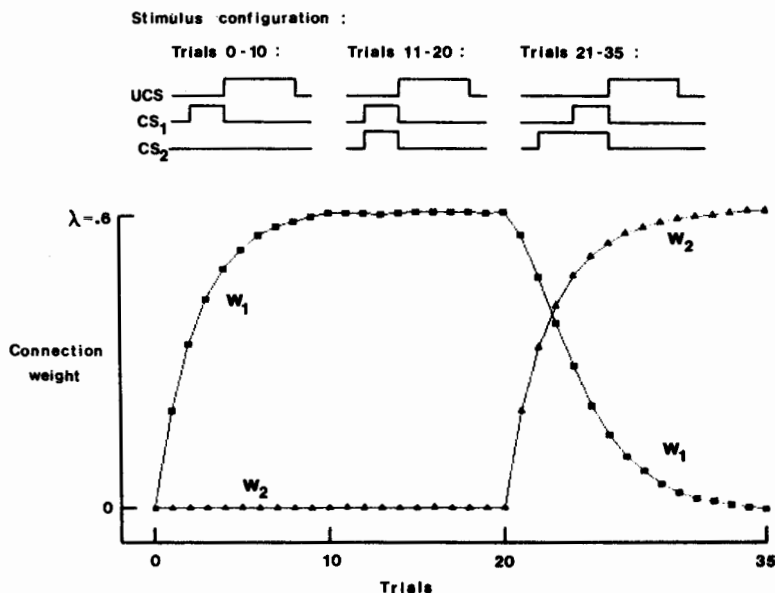


Figure 14. The connection weights at the end of each trial in a simulation experiment. (The intratrial time courses of the variables involved are not shown. In Trials 0-10, presentation of CS<sub>1</sub> alone followed by the UCS results in  $w_1$  increasing. In Trials 11-20, CS<sub>1</sub> and CS<sub>2</sub> presented together followed by the UCS produces no change, since CS<sub>2</sub> is redundant. This is the blocking paradigm. In Trials 21-35, CS<sub>2</sub> begins earlier than CS<sub>1</sub>. The element becomes sensitive to the earlier predictor and loses sensitivity to the later.)

relation (Figure 14, Trials 11-20). A simulation of our element in this situation produced results shown in Figure 17.

These simulations confirm that when viewed at the trial level and given the assumptions made above, our model behaves like the Rescorla-Wagner model and, in particular, produces the stimulus context effects of that model. When viewed at the level of trials, our model also shares the shortcomings of the Rescorla-Wagner model regarding extinction of conditioned inhibitors and the shape of the acquisition curve. Extensions of the Rescorla-Wagner model proposed to eliminate these shortcomings (Frey & Sears, 1978) are also applicable to our model. However, even with these extensions, the Rescorla-Wagner model applies only at the level of trials. It cannot supply predictions about the effects on conditioning of the intratrial temporal relationships between stimuli. As we have seen, our model does apply to this intratrial structure for the case of a single CS, having behavior consistent with data on CR latency and ISI dependence. In addition, our model provides an

extension of the Rescorla-Wagner use of expectation to a form having meaning within trials. This leads to several novel and interesting forms of model behavior.

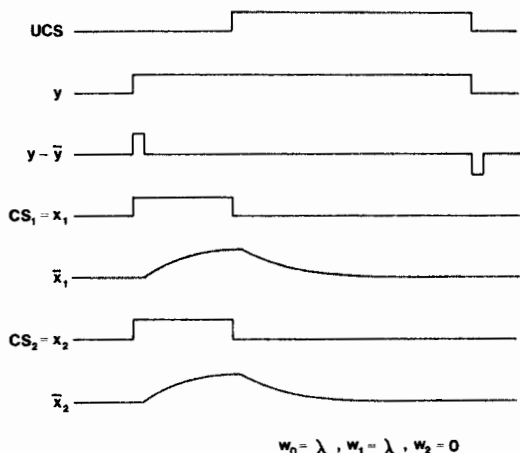


Figure 15. Intratrial time courses of element variables in Part 2 of a blocking experiment (Trials 11-20 shown in Figure 14). (Since the weight associated with CS<sub>1</sub> has already reached its asymptotic value of  $\lambda$ ,  $y - \bar{y}$  is zero whenever CS<sub>2</sub> pathway eligibility  $\bar{x}_2$  is nonzero. Consequently, no changes in weight values occur.)

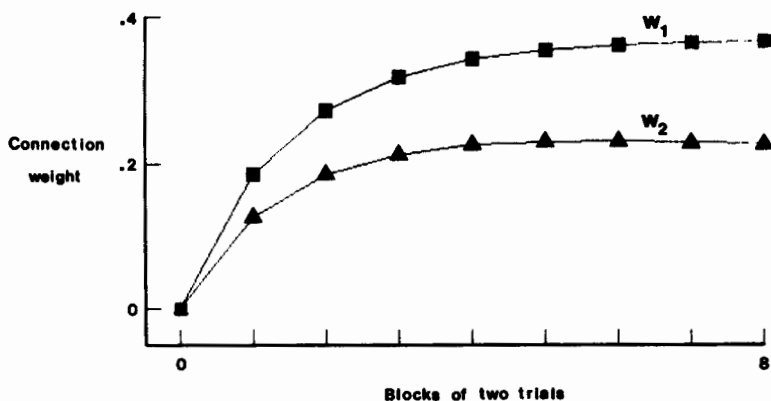


Figure 16. Simulation results of an experiment with two CSs, each of which accounts for a particular portion of the UCS's reinforcement. (Trials of  $CS_1$  paired with a UCS of strength  $\lambda = .4$  were alternated with trials in which the compound  $CS_1 + CS_2$  was paired with a UCS of strength  $\lambda = .6$ .)

The adaptive element we have presented finds the earliest predictors and ignores redundant later predictors. A CS that arrives simultaneously with or after a UCS is useless as a predictor. By the same reasoning, predictors that occur earlier than others are in some sense more predictive and potentially more useful. A later predictor can be redundant to an earlier one in the same sense that an unreliable predictor can be redundant to an identically timed but reliable predictor. For example, let  $CS_1$  and  $CS_2$  both always be followed by reinforcement, but let  $CS_2$  start earlier than  $CS_1$ . Then even if ini-

tially  $CS_1$  is dominant ( $w_1 = \lambda$ ,  $w_2 = 0$ ), eventually  $CS_2$ , the earlier predictor, will completely dominate  $CS_1$  as a predictor of the UCS (eventually  $w_1 = 0$ ,  $w_2 = \lambda$ ). The result of a simulation of this experiment is shown as Trials 21–35 of Figure 14 (recall that at Trial 20,  $w_1 = \lambda$  and  $w_2 = 0$ ). Although both stimuli are being presented in Trials 11–20 and in Trials 21–35, in the former case  $CS_2$  is blocked by  $CS_1$ , whereas in the latter the associative strength of  $CS_2$  increases quickly and  $CS_2$  comes to completely dominate  $CS_1$ . In the earlier trials  $CS_2$  is redundant to  $CS_1$ , which had already been

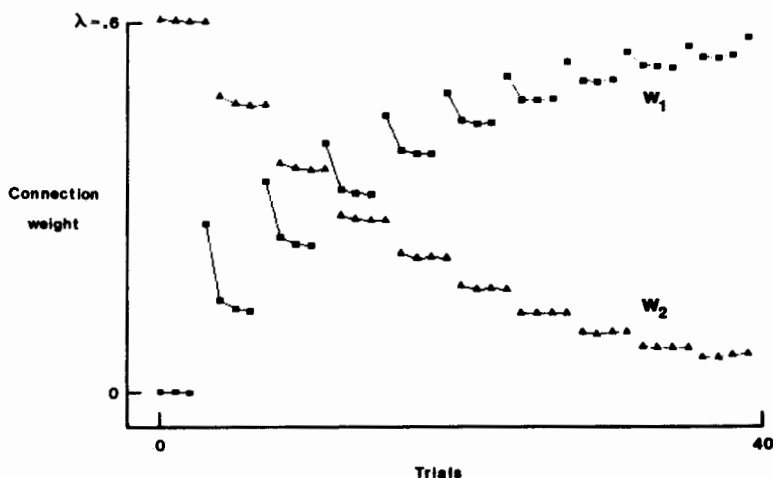


Figure 17. Simulation results of an experiment with two CSs differentially associated with the UCS. ( $CS_1$  precedes every UCS and  $CS_2$  is absent every fourth UCS. Although initially  $CS_2$  is dominant [ $w_1 = 0$ ,  $w_2 = \lambda$ ], eventually  $CS_1$ , the more reliably associated CS, dominates [ $w_1 = \lambda$ ,  $w_2 = 0$ ].)

conditioned, but in these later trials  $CS_2$  provides important new information: It is the earliest indicator that the UCS will occur. This advantage, combined with the fact that  $CS_1$  is totally redundant to  $CS_2$ , produces complete conditioning to  $CS_2$  and the elimination of conditioning to  $CS_1$ .

The approach to this steady state is quick and orderly, but the reasons for this behavior are somewhat difficult to explain. Very briefly, on each trial the associative strength  $w_2$  of  $CS_2$  increases and then decreases by a lesser amount for a net gain, whereas the associative strength  $w_1$  of  $CS_1$  only decreases:  $w_2$  increases because  $CS_2$  predicts the onset of  $CS_1$ 's excitation, and both  $w_1$  and  $w_2$  decrease at the offset of  $CS_1$  and  $CS_2$  because these two stimuli together produce too much expectation.

Although this property of the adaptive element to become sensitive only to earliest predictors of a UCS when the later ones provide no new information is reminiscent of some learning theory results (notably the work of Egger & Miller, 1962, on conditioned reinforcement), our primary interest in it stems from adaptive systems considerations. We feel that a simple mechanism that finds the earliest, most reliable, and nonredundant predictors of important events is potentially very useful for constructing powerful adaptive systems.

## Higher Order Conditioning

Although much of the discussion has been in terms of fixed pathways that (corresponding to UCSs) cause changes in plastic pathways, signals on these plastic pathways, since they also can affect  $y$ , can also cause such changes. The simplest example of this corresponds to what is known as higher order conditioning in animal learning theory. A signal on a plastic pathway ( $CS_1$ ) is paired with a fixed input (UCS) until the connection weight  $w_1$  reaches its asymptotic value. Then a signal on a second plastic pathway ( $CS_2$ ) is paired with a signal on the first plastic pathway ( $CS_1$ ). In this second pairing  $CS_1$  acts as a reinforcing UCS for  $CS_2$ . With repeated pairings, the second connection weight  $w_2$  grows to the level of  $w_1$ , but  $w_1$ , since its use is not followed by a UCS, gradually falls to zero. The result is that  $w_2$  rises to the level of  $w_1$  and then follows  $w_1$  to zero. The results of a simulation of this experiment are shown in Figure 18.

## Adaptive System Theory

In this section we discuss how the model we have presented is related to a variety of other learning rules used in adaptive system research. This discussion will serve to place the model within a theoretical framework and indicate how it differs from learning

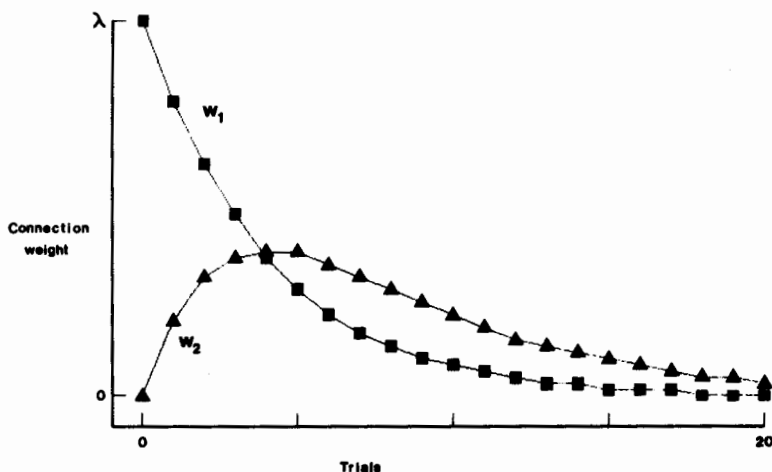


Figure 18. Connection weight values at the end of each trial in a simulation of higher order classical conditioning. ( $CS_1$  has been paired with a UCS until the weight  $w_1$  reached the asymptotic value  $\lambda$ . For the trials shown,  $CS_2$  and  $CS_1$  are sequentially presented in the absence of the UCS causing  $w_2$  to increase as  $CS_1$  acts as a reinforcing stimulus for  $CS_2$ . Since  $CS_1$  is not being followed by the UCS,  $w_1$  decreases to zero, causing a similar decrease in  $w_2$ .)

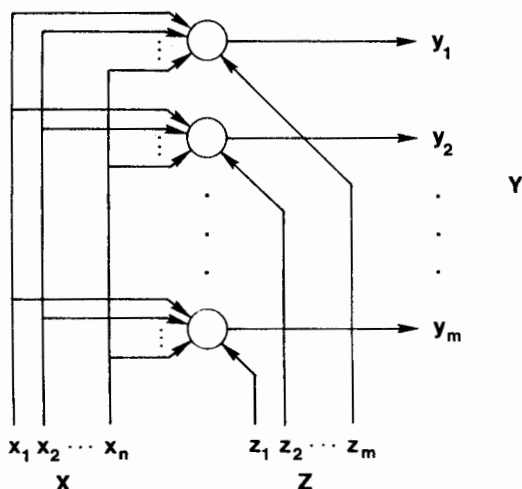


Figure 19. An associative memory network consists of a bank of  $m$  adaptive elements sharing the same  $n$  input pathways. (Any of the many types of adaptive elements proposed can be studied in this configuration. Although each type of element leads to different storage and retrieval capabilities, all such networks show the properties of generalization, noise resistance, and content addressability that have stimulated interest in these structures.)

rules proposed in the past. The history of adaptive systems research is too long and too diverse to review exhaustively here. Useful reviews are provided by Minsky (1963), Minsky and Selfridge (Note 3), Hawkins (1961), Holland (1975), and Klopff (1979). Even by restricting attention to adaptive systems based on "neural" mechanisms, we would be unable to give more than a cursory treatment. Arbib, Kilmer, and Spinelli (1976) provide a good, though also nonexhaustive, review of adaptive neural models. Here we focus only on rules that have received the most attention and are most closely related to our model.

Consider a generalized learning rule (as in Amari, 1977a): A synaptic weight increases or decreases in proportion to a reinforcement signal  $r$ :

$$w_i(t+1) = w_i(t) + cr_i(t), \quad (8)$$

where  $c$  = positive learning rate constant,  $w_i(t)$  = weight of synapse  $i$  at time  $t$ , and  $r_i(t)$  = reinforcement signal to synapse  $i$  at time  $t$ . We are using the term *reinforcement signal* simply to denote the signal that determines the changes in connection weights.

For some of the learning rules, this signal only vaguely resembles what would be called reinforcement in animal learning studies.

### Hebbian Rule

Within this framework the Hebbian postulate, in the form which we briefly discussed above, is formulated by lettering  $r_i(t) = x_i(t)y_i(t)$  in Equation 8. The most well-developed application of the Hebbian learning rule is its use in networks that implement associative information storage (e.g., Amari, 1977a, 1977b; Anderson et al., 1977; Kohonen, 1977; Nakano, 1972; Wigström, 1973). The network shown in Figure 19 transforms stimulus patterns  $X = (x_1, \dots, x_n)$  to response patterns  $Y = (y_1, \dots, y_m)$ . The inputs  $z_i$  act on the elements in exactly the same manner as the inputs  $x_i$  but are used to specify patterns  $Z = (z_1, \dots, z_m)$  to be associated with the stimulus patterns  $X$ . Repeated presentations of  $k$  different pairings of stimulus patterns  $(X_1, Z_1), \dots, (X_k, Z_k)$ , causes the network to learn, using the Hebbian learning rule, to elicit  $Z_\alpha$  when presented with  $X_\alpha$  alone,  $\alpha = 1, \dots, k$ . This occurs provided the patterns  $X_1, \dots, X_k$  form an orthogonal set. The matrix of synaptic weights  $[w_{ij}]$  converges to the correlation matrix of the patterns  $X_\alpha$  and  $Z_\alpha$ ,  $\alpha = 1, \dots, k$ .<sup>4</sup>

What accounts most strongly for the current widespread interest in associative memory networks is that they exhibit properties suggestive of the aspects of memory emphasized by Gestalt or mass action theorists (e.g., Freeman, 1975; John & Schwartz, 1978). Since information can be stored in distributed form, associative performance may not be seriously impaired by various kinds of "lesions" (e.g., Wood, 1978). Distributed storage also provides for interesting forms of generalization and content addressability (e.g., Kohonen, 1977; Nakano, 1972; Wigström, 1973).

These models provide evidence that learning rules that are essentially connectionistic

<sup>4</sup> It actually converges to the correlation matrix, also called the covariance matrix, only if the averages of the input patterns  $X_\alpha$  and  $Z_\alpha$  are zero. Amari (1977a) shows how the Hebbian rule can be modified to remove this restriction.

in character need not imply a locationalistic view of memory. The theory of associative memory networks is well understood, and as research continues on mechanisms of this type, the result emerging is that any application of simultaneous, or spatial, correlation can be cast in a form that a Hebbian rule can implement.

However, as we indicated when discussing temporal relationships, the temporal subtleties of classical conditioning are not produced by the Hebbian rule even with the use of delays and other modifications. One would therefore not expect the processing capabilities of networks of Hebbian adaptive elements to extend far beyond spatial correlation.

### Widrow-Hoff Rule

For the Widrow-Hoff rule the reinforcement signal is defined as follows:

$$r_i(t) = [z(t) - y(t)]x_i(t), \quad (9)$$

where

$$y(t) = \sum_{j=1}^n w_j(t)x_j(t), \quad (10)$$

and  $z(t)$  and  $x_i(t)$  are real numbers. This rule requires the use of a specialized signal  $z$  that acts differently from the other input signals due to its special role in Equation 9 and the fact that it does not participate at all in the computation of the output  $y$  given by Equation 10. This rule causes the weights to converge so that the response is a particular desired real number for each stimulus. If, for example, during the presentation of each stimulus pattern  $X_\alpha$ ,  $\alpha = 1, \dots, k$ , the value to be associated with it is presented as  $z$ , call it  $z_\alpha$ , then after sufficient repetitions of the pairs  $(X_\alpha, z_\alpha)$ , the element will respond with  $z_\alpha$  when presented with  $X_\alpha$  alone,  $\alpha = 1, \dots, k$ . The rule implements an iterative algorithm for computing a solution to a set of linear equations. A solution exists if the stimulus patterns  $X_1, \dots, X_k$  are linearly independent.

If the stimulus patterns are not linearly independent, convergence can still occur if the rule is modified by making the learning rate parameter  $c$  a variable whose value approaches zero as the trials continue, for ex-

ample,  $c(t) = c/t$ . In this case, and provided the pairs  $(X_\alpha, z_\alpha)$  occur with sufficient frequency in the input sequence, the weights converge so as to minimize the sum of the squared error over the stimulus patterns, that is,

$$\sum_{\alpha=1}^k (y_\alpha - z_\alpha)^2$$

is minimized where  $y_\alpha$  is the element's output for pattern  $X_\alpha$ , and  $z_\alpha$  is the desired output. In this form the Widrow-Hoff rule is an iterative algorithm for forming the Moore-Penrose pseudoinverse of a linear operator, which is the same as saying that it computes a linear regression. Duda and Hart (1973) provide a useful discussion of this and closely related stochastic approximation procedures. This rule was proposed in the form of an adaptive element by Widrow and Hoff (1960).

Amari (1977a, 1977b) discusses associative memory networks of neuronlike elements that rely on the Widrow-Hoff rule to form associations. In discussing the associative network shown in Figure 19, we said that when Hebbian synapses are employed, perfect recall of  $z_\alpha$  on presentation of  $X_\alpha$ ,  $\alpha = 1, \dots, k$ , was possible only when the stimulus patterns formed an orthogonal set. Using the Widrow-Hoff rule, perfect recall occurs even if the stimulus pattern set is only linearly independent. Amari (1977a, 1977b) calls this *orthogonal learning*, since nonorthogonal patterns are "orthogonalized" by the network. For sets of stimulus patterns that are not linearly independent, recall of the best pattern in the least-mean-square sense can be achieved.

A fact that is not generally realized is that the Widrow-Hoff rule is essentially identical to the Rescorla-Wagner equation. To see this, identify  $t$  with the trial number, each input with a CS, and the  $z$  signal with the UCS, so that  $z = \lambda$  when the UCS is present and  $z = 0$  otherwise. In the Rescorla-Wagner equation (Equation 7), the presence of a CS<sub>*i*</sub> input signal on a trial is indicated by the set notation  $i \in S$ , whereas the Widrow-Hoff form uses  $x_i$  nonzero to indicate input signal presence on a trial and  $x_i = 0$  to denote absence. The relevant equations are, for Rescorla-Wagner,

$$\Delta V_{CS_i} = \begin{cases} c[\lambda - \sum_{j \in S} V_{CS_j}] & \text{for } i \in S \\ 0 & \text{for } i \notin S; \end{cases}$$

for Widrow-Hoff,

$$\Delta w_i = c[z - \sum_{j=1}^n w_j x_j] x_i.$$

The correspondences between the two equations are (a)  $w_i = V_{CS_i}$ ; (b)  $z = \lambda$  if the UCS is present, otherwise  $z = 0$ ; and (c)  $x_i = 1$  if  $CS_i$  is present, otherwise  $x_i = 0$ .

That these two models are, in fact, identical is striking, since they were constructed for very different purposes. The Widrow-Hoff rule was formulated as an algorithm to solve sets of linear equations, and its theory addresses convergence properties. Not only are stimulus context effects not discussed in this theory, but their existence is entirely incidental. The Rescorla-Wagner theory was proposed to compactly describe a wide variety of effects observed in animal learning experiments. That it also provides an important algorithm with a strong connection to very useful areas of applied mathematics is fortuitous. We feel that the confluence of mathematical and empirical facts represented by what we shall call the Rescorla-Wagner/Widrow-Hoff rule might have considerable significance for understanding associative learning.

Due to its similarity to the Rescorla-Wagner model, the Widrow-Hoff rule provides a more adequate model of classical conditioning than does the Hebbian model. Unlike the Hebbian model, however, it does not provide a simple explanation for a stimulus substitution view of conditioning. Figure 4 shows the Widrow-Hoff rule as a model of classical conditioning. The specialized input  $z$  corresponds to the UCS. Since  $z$  does not directly influence the element's output, the UCR and CR must use separate pathways (compare to Figure 3). Also unlike the Hebb rule, the Rescorla-Wagner/Widrow-Hoff rule has the property that weight modifications can only be driven by the specialized "teacher" input  $z$ .

A learning rule closely related to the Rescorla-Wagner/Widrow-Hoff rule is the perceptron rule of Rosenblatt (1962). If  $z(t)$  in Equation 9 is restricted to taking only the values 0 and 1 and the output is similarly restricted by the use of a threshold, Equation

9 gives the fixed increment perceptron rule. This rule is an iterative procedure for solving a set of linear inequalities. A solution exists if the desired response is a linearly separable function of the stimulus patterns. Nilsson (1965) provides several proofs of convergence, and Minsky and Papert (1969) discuss its limitations as a pattern recognition system. Despite these limitations, the perceptron learning rule has resurfaced, in slightly disguised form, as a way of storing data in associative memory structures (Albus, 1979; Amari, 1977a, 1977b). These applications illustrate that in certain applications and using certain ways of representing data, the limitations of linear learning rules are not as devastating as once thought.

The perceptron seems to be most often thought of as a model of instrumental conditioning in which reinforcement is contingent on the response rather than of classical conditioning, which involves no response contingencies. This view, however, is mistaken. If the error signal  $z(t) - y(t)$  in Equation 9 is taken as being computed by the perceptron's environment, then the perceptron can be viewed as a response-contingent system: If the response is correct, the error is 0; if it is incorrect, the error is 1 or  $-1$ . However, this feedback through the environment is of such stereotyped form that it can be eliminated, for arbitrary environments, by just letting the error be computed by the perceptron itself, with the environment always simply providing the desired response rather than an error signal. Viewed in this manner, the perceptron is essentially the same as the Rescorla-Wagner model: It compares its own response (expectation) with the correct one (UCS) and modifies the weights in order to make them agree. The instrumental conditioning paradigm, on the other hand, involves essential feedback through the organism's environment, that is, feedback that cannot be eliminated in a uniform way for all environments. Nontrivial forms of response-contingent learning have received very little attention by adaptive network theorists.

### *Rescorla-Wagner/Widrow-Hoff Predictor*

The Rescorla-Wagner/Widrow-Hoff rule does not produce the predictive aspect of

classical conditioning. Here we discuss the minimal modifications to that rule which enable it to produce predictive or anticipatory responses. From the resulting rule, which we call the Rescorla-Wagner/Widrow-Hoff predictor, it is possible to see what additional properties our model provides. Although we know of no instance in which the Rescorla-Wagner/Widrow-Hoff predictor is used in an adaptive network theory, it is an example of a linear prediction procedure and is part of a larger theory of prediction or forecasting (see, for example, Box & Jenkins, 1976).

For the Rescorla-Wagner/Widrow-Hoff predictor the reinforcement signal is defined as follows:

$$r_i(t) = c[z(t) - y(t - \tau)]x_i(t - \tau), \quad (11)$$

where  $y(t)$  is as defined by Equation 10 and  $\tau$  is some positive constant. Changes in connection weights are such as to reduce the difference between  $z(t)$  and  $y(t - \tau)$  so that an equilibrium is approached at which  $z(t) = y(t - \tau)$ , or  $z(t + \tau) = y(t)$ . This means that the element will learn to produce activity that anticipates by  $\tau$  the activity of the UCS pathway  $z$ , if the input contains enough predictive information. More precisely, with reference to the discussion of the Widrow-Hoff rule above, if  $c$  is allowed to decrease as conditioning proceeds, this element will produce a best least squares prediction by  $\tau$  of the signal  $z$ . All of the stimulus context effects of the Rescorla-Wagner/Widrow-Hoff rule are also produced by the predictor.

The process defined by the predictor can be described as follows: Activity on an input pathway possibly causes a response but also causes the connection from that pathway to become eligible for modification a certain period of time ( $\tau$ ) later. An eligible connection is modified only if the UCS signal strength differs from the expected strength. Thus, each time  $z(t)$  deviates from  $y(t - \tau)$ , the input pathways that were active earlier (and thus eligible) will modify their connection weights, or associative strengths,  $w_i$ . The reinforcement signal is a measure of how strongly the current UCS confirms or contradicts the previously formed expectation or prediction.

As a model of classical conditioning, the

predictor defined by Equation 11 requires an ISI exactly equal to  $\tau$  for any conditioning to occur. This limitation can be eliminated, along with the arbitrariness of the choice of  $\tau$ , by replacing the delayed signals  $x_i(t - \tau)$  and  $y(t - \tau)$  in Equation 11 by more general forms of traces such as those used in our model. Let  $\bar{x}_i(t)$  and  $\bar{y}(t)$  be some weighted averages of their respective function values over some time interval preceding  $t$ , produced using Equations 3 and 4. Then the reinforcement signal for the Rescorla-Wagner/Widrow-Hoff predictor becomes

$$r_i(t) = c[z(t) - \bar{y}(t)]\bar{x}_i(t). \quad (12)$$

The temporal relationships implied by this rule depend on the characteristics of the CS and UCS, the form of the traces  $\bar{x}_i$  and  $\bar{y}$ , and the parameters of the experimental paradigm. Some details of these dependencies are presented in Appendix B.

### Uttley's Informon

Uttley (1970, 1975, 1976a, 1976b, 1976c, 1979) has suggested a learning rule that is closely related to the Rescorla-Wagner/Widrow-Hoff procedure except that it conforms to some of the constraints of the Hebbian rule. Starting with the Widrow-Hoff rule (Equation 9), let  $z(t) = -w_0x_0(t)$ , where  $w_0$  is a fixed positive number. That is, let the specialized "teacher" input be a signal to a fixed inhibitory pathway. If it is further assumed that this fixed signal participates in the computation of the output  $y$  just like any other input signal, then Equation 9 can be rewritten as follows:

$$\begin{aligned} r_i(t) &= [-w_0x_0(t) - \sum_{j=1}^n w_j(t)x_j(y)]x_i(t) \\ &= -[\sum_{j=0}^n w_j(t)x_j(t)]x_i(t) \\ &= -y(t)x_i(t). \end{aligned} \quad (13)$$

This is the Hebbian rule except for the minus sign. Uttley argues that this change of sign is desirable because it changes the positive feedback inherent in the Hebbian rule to negative feedback desirable for its stabilizing influence. Coincidence of pre- and postsynaptic discharges decreases rather

that increases synaptic strength, and equilibrium weight values are those that result in zero total input to the element. Uttley notes the similarity of this rule to the Rescorla-Wagner model and illustrates how it can produce much of the same behavior (Uttley, 1975).

Uttley describes his model in the manner discussed above but actually simulates a more complex model based on the concept of *mutual information*. He uses exponentially weighted time averages to estimate the negative of the mutual information between input and output signals. At each time step the weights are set to these estimates. Although the concept of mutual information led Uttley to the informon model and provides an interesting view of the stimulus contingencies that produce learning, it is an unnecessary complication to what is essentially the Widrow-Hoff rule.

If the special input labeled  $z$  in Figure 4 is regarded as a fixed inhibitory input, then that figure shows the use of Uttley's element in an analog of the classical conditioning paradigm. This is identical to the corresponding situation for the Rescorla-Wagner/Widrow-Hoff and perceptron models. Here, however, there is the additional consequence that the UCS actually inhibits the CR both before and after learning. This is due to the treatment of the UCS as an inhibitory signal that is used in the computation of the element's output. In the Rescorla-Wagner/Widrow-Hoff and perceptron models, the UCS is a special input that never influences the output of the element except indirectly through the learning process.

In order to obtain the stability and stimulus context effects of the Rescorla-Wagner/Widrow-Hoff rule while at the same time adhering to the basic constraints of the Hebbian rule, Uttley had to abandon the simple stimulus substitution view of conditioning provided by the Hebbian rule and make the behaviorally unsupportable assumption that the UCS inhibits the CR.

Since it retains the form of the Hebbian rule, however, the informon has the property that even though there are fixed, prespecified, classifying input channels, these channels are not the only source of signals that can cause weight modifications. This is an important property, but it can be obtained

in an entirely different manner (as illustrated by our model) that also has the advantage of the Rescorla-Wagner/Widrow-Hoff rule but retains a stimulus substitution view, produces appropriate ISI dependency, and permits the CR to occur before the UCS. Although we feel that Uttley's approach represents an independent discovery of the advantages of the Rescorla-Wagner/Widrow-Hoff rule, we also feel that it needlessly adheres too closely to the original Hebbian postulate.

### Our Model

Within the framework provided by Equation 8, our model uses a reinforcement signal defined as follows:

$$r_i(t) = [y(t) - \bar{y}(t)]\bar{x}_i(t), \quad (14)$$

where  $y$  is as defined by Equation 10 and  $\bar{y}$  and  $\bar{x}_i$  are traces of their respective signals as described above. This differs from the Rescorla-Wagner/Widrow-Hoff predictor (Equation 12) by the substitution of  $y(t)$  for the specialized reinforcing signal  $z(t)$ . This eliminates the requirement that reinforcement be provided only by a fixed reinforcing pathway. Since  $y(t)$  can be affected by activity on any input pathway, any input signal can bring about changes in the efficacies of other pathways. This permits the adaptive element to extract predictive relationships among its inputs in the same way that a Hebbian element or the informon extracts simultaneous associations. Unlike the informon, however, our model retains the stimulus substitution properties of the Hebbian model, since the CR and the UCR share the same pathway.

We have been able to eliminate the need for a distinct channel for reinforcing signals by, in effect, providing a distinct time (with respect to a CS) for reinforcement. This was suggested by the work of Klopff (in press, Note 1) in which a similar method was proposed for eliminating the requirement that response-contingent reinforcement be delivered over a specialized channel. We have restricted this idea to classical conditioning and must postpone a discussion of its implications for the response-contingent case to a later paper.



We note that it is possible to use our model in an associative memory system such as those described above that rely on the Hebbian rule or the Rescorla-Wagner/Widrow-Hoff rule. This would result in a network capable of exhibiting the properties of our model together with the properties of distributed, associative information storage. We have not yet systematically explored the implications of such a system, but it is unlikely that it would lack any of the properties that have stimulated interest in this kind of associative memory structure. In particular, such a system would show that our model, although connectionistic in character, need not imply a locationalistic theory of memory.

### Stability and Saturation

Some issues that were not directly addressed in the preceding section concern technical problems that occur when networks of elements based on various learning rules are simulated. For example, a literal application of the Hebb postulate implies a positive feedback loop (increases in excitatory synaptic weights cause higher correlation between pre- and postsynaptic activity and hence further weight increases). Excitatory synaptic weights tend to become large irrespective of the significance of the input signals, and some additional mechanism is required to prevent the strengths of all connections from growing without bound or from reaching and remaining at their maximum values. Early computer simulations illustrated the importance of solving these problems for preventing network "seizures" (Rochester, Holland, Haibt, & Duda, 1956).

Here we discuss several approaches to solving the stability and saturation problems associated with the Hebb rule and relate them to the solution provided by the learning rule we have developed. Our point is to show that learning rules that are based on the Rescorla-Wagner/Widrow-Hoff rule, such as ours and the one proposed by Uttley, not only provide more valid models of classical conditioning than the Hebbian rule but also solve these technical problems in a simple way. Although there is no logical or empirical necessity for the stimulus context effects accounted for by the Rescorla-Wagner/Widrow-Hoff rule to arise at a cellular level,

it is suggestive that if they did then additional mechanisms would not be required in order to solve stability and saturation problems.

Some of the current approaches for solving these problems (notably, Grossberg's, 1969, 1974, 1976a, 1976b) stress the importance of careful network design and use of inhibitory connections for controlling network stability. Other approaches attempt to achieve similar results by modifying the original Hebbian postulate so as to incorporate local stabilizing mechanisms that operate irrespective of an element's network environment. This latter approach has not been shown to be sufficient for solving network stability problems but does contribute to their solution by making the adaptive changes inherently more manageable. Although we feel that network level considerations (i.e., a priori structure) are very important, they are strongly influenced by the choice of local learning rules, and here we focus on element level issues.

There are two fundamentally different ways of preventing unbounded weight growth in theoretical models of plastic synapses. The first technique is to impose an upper bound at some fixed, predetermined value. Whether this is done by setting the weight back to the preset maximum whenever an increment makes it larger or by modifying the learning rule in order to make the value asymptotically approach a preset finite limit, the same problem arises: Unless weights decrease in some circumstances, all excitatory weights will tend to reach and remain at their maximum values. Saturation of some weights may be desirable, but if all the weights always eventually reach their maximum value, then all learning eventually ceases and all stored information is eventually forgotten. Either learning must occur slowly enough to postpone this ultimate state of forgetfulness for as long as necessary, the plasticity of some connections must be temporary, or a means for decreasing weights must be introduced.

The second technique for preventing unbound weight growth relies on the boundedness of the reinforcement signal that drives the weight modification process. Instead of a fixed, predetermined limit's being enforced by the learning rule, the limit is a

function of the external reinforcing input to the element. Larger reinforcement can always cause weights to increase, but—and this is the crucial point—arbitrarily prolonged periods of nonzero reinforcement must not produce arbitrarily large weights. Several of the methods discussed below solve stability and weight saturation problems in this manner.

### Normalization

One of the most common techniques used in simulations invokes a “conservation of total synaptic strength,” or normalization, principle. This technique is a particular way of presetting weight bounds. The total saturation problem is avoided by requiring some weights to decrease in order to maintain the sum of all the weights at a constant value. New weight values  $w'_i$  are computed according to the Hebbian rule, and then each  $w'_i$  is divided by the sum of all of the  $w'_i$  to obtain the actual next weight values, that is (cf. von der Malsburg, 1973),

$$w_i(t+1) = w'_i / \sum_{j=1}^n w'_j, \quad (15)$$

where

$$w'_i = w_i(t) + cx_i(t)y(t). \quad (16)$$

This normalization procedure is successful in permitting those pathways to dominate whose activity is most strongly correlated with postsynaptic activity. One can view synaptic strength modification computed in this way as a competition among pathways for proportions of the sum. In many models, this procedure is absolutely essential, not only for stable operation of the model but also for the generation of behavior that resembles experimental data (e.g., von der Malsburg, 1973).

Although it adequately solves some of the technical problems associated with the Hebbian rule, this normalization procedure has several deficiencies. First, it was pointed out by Uttley (1976a) that although perfect normalization often produces desired results, small departures from this ideal can cause rather drastic changes in behavior. If, for example, one synapse is consistently favored by even a very small amount in the normalization process, then it can gain much more than its share of the total synaptic

strength. The weight values can reflect normalization asymmetries rather than the desired correlation measures.

A second criticism of the normalization procedure holds to the extent that a faithful representation of classical conditioning phenomena is desired. Although stimulus context effects are produced by normalization, they are different from those observed experimentally. For example, suppose each input  $x_i$ ,  $i = 1, \dots, n$ , to an adaptive element using a normalized Hebbian scheme has a constant value  $x_i(t) = x_i$  for all  $t$ . Using Equations 15 and 16, it is not hard to show that the equilibrium weights are

$$w_i = x_i / \sum_{j=1}^n x_j, \quad i = 1, \dots, n. \quad (17)$$

These equilibrium values are independent of the initial weight values. The stimulus context effects observed experimentally, however, require that the associative strengths at the beginning of a series of trials crucially determine their values at its end.

Consider blocking, for example. Suppose an element has two binary valued inputs  $x_1$  and  $x_2$  corresponding to conditioned stimuli  $CS_1$  and  $CS_2$  and an input of arbitrary fixed strength representing an unconditioned stimulus. Assume that the associative strengths  $w_1$  and  $w_2$  of  $CS_1$  and  $CS_2$  initially equal zero and are thereafter required always to sum to one. Pairing  $CS_1$  with the UCS until equilibrium is reached results by Equation 17 in  $w_1 = 1$  and  $w_2 = 0$ . Now, starting with these values and pairing both  $CS_1$  and  $CS_2$  with the UCS results by Equation 17 in equilibrium values  $w_1 = w_2 = .5$ . This is the same result that would have been produced if the weights were both zero at the commencement of the paired trials. Blocking, on the other hand, would occur if the series of paired trials did not change the weights from the values they had when it began, that is,  $w_1 = 1$ ,  $w_2 = 0$ .

Another criticism of the normalization technique can be made if a model using this method is intended to reflect what might occur in actual neurons. Although it has been suggested that synaptic modifications and normalization may be the result of the redistribution of a constant amount of receptor protein (Stent, 1973), this hypothesis

positive feedback inherent in the Hebbian rule by changing it into negative feedback by reversing the sign (Equation 13). Coincidence of pre- and postsynaptic discharges decreases rather than increases the synaptic strength. This produces the stimulus context effects observed in classical conditioning experiments while at the same time solving stability and saturation problems (and producing the undesirable consequences discussed earlier). However, it is not the precise form of Uttley's model that produces these solutions, but rather its resemblance to the Rescorla-Wagner/Widrow-Hoff rule.

### Autonomous Decay

If it is assumed that synaptic strength slowly decays in the absence of a reinforcement signal, then a bound on weight size is imposed that is a function of reinforcement level and the decay rate. A weight can always increase when its reinforcement signal increases, but if the reinforcement signal remains bounded, then no matter how long the signal persists the weight also remains bounded. Thus, learning can occur whatever the system's "age," but experiences are always "forgotten" within a certain period of time. In system theoretic terms, the adaptive element has *definite memory*: It cannot remember anything that occurred arbitrarily far in the past. Moreover, the weight bound is inversely proportional to the length of time that memory traces can be retained. That is, if weights are to be kept below rather low levels, then the decay of the weights must be rather fast. The normalization method described above, in contrast, has *indefinite memory*, meaning that information is not lost unless actively replaced by new information.

Despite the lack of indefinite memory, learning rules incorporating autonomous decay lead to behavior that can be understood in mathematical terms. For example, if the decay is sufficiently slow, then the long-term statistical properties of the reinforcement signal can be reflected in the weight values (e.g., Amari, 1977a, 1977b; Uttley, 1976a, 1976b, 1976c).

### Negative Feedback

Although normalization and autonomous decay schemes employ negative feedback, rules resembling the Rescorla-Wagner/Widrow-Hoff procedure use a more explicit form. Uttley (1976a) directly eliminates the

positive feedback inherent in the Hebbian rule by changing it into negative feedback by reversing the sign (Equation 13). Coincidence of pre- and postsynaptic discharges decreases rather than increases the synaptic strength. This produces the stimulus context effects observed in classical conditioning experiments while at the same time solving stability and saturation problems (and producing the undesirable consequences discussed earlier). However, it is not the precise form of Uttley's model that produces these solutions, but rather its resemblance to the Rescorla-Wagner/Widrow-Hoff rule.

Recall that for the Rescorla-Wagner/Widrow-Hoff rule, weights change according to

$$\Delta w_i(t) = c[z(t) - \sum_{j=1}^n w_j(t)x_j(t)]x_i(t).$$

A weight therefore cannot change if either the input signal on that pathway is zero ( $x_i(t) = 0$ ) or the total stimulus strength,

$$\sum_{j=1}^n w_j(t)x_j(t),$$

equals the training signal  $z(t)$ . Thus, the weights are always bounded, yet never become saturated so as to be insensitive to further changes in the environment. Negative feedback is provided in the form of the expectation term

$$-\sum_{j=1}^n w_j(t)x_j(t).$$

Learning can always occur when the reinforcement differs from the expected level, and the asymptotic weight values depend on the magnitude of the reinforcement signal. Moreover, the rule permits memory of events that occurred arbitrarily far in the past, that is, it has indefinite memory.<sup>5</sup> A weight will

<sup>5</sup> It is curious that the model Uttley actually simulates does not possess indefinite memory. Exponentially weighted time averages are used to estimate the negative of the mutual information between input and output signals. This makes the informon's stability an obvious property, but memory traces always decay to zero due to the exponential decay used to estimate mutual information. He discusses an additional hysteresis mechanism that produces indefinite memory (Uttley, 1970). If no estimates of mutual information were used, then indefinite memory would be present without an additional mechanism, and none of the rule's other advantages would be lost.

not decrease unless a stimulus occurs that is not reinforced to the expected level.

The model we have presented uses negative feedback in the form of an expectation term that is a weighted average of past values of the element's output. The weight associated with a pathway cannot change unless that pathway is eligible and the current output value differs from the weighted average of past output values. As a weight grows, the signals arriving on that pathway cause larger output values and hence larger expectations. As the expectations grow they exert negative effects on weight growth. The stability of this method is evident from the simulation results shown above.

One consequence of this form of negative feedback in our element is that if a signal arrives via a modifiable pathway, for example, as a positive rectangular pulse, but is not followed by other activity within the eligibility period, then the weight of that pathway will, if positive, decrease toward zero. This will occur because this signal's offset will coincide with positive eligibility to cause a negative change in weight. With repeated presentations of stimuli not directly followed by other activity, weights will converge to zero. This is why in the simulations of classical conditioning presented above, we required the UCS to arrive over a pathway of fixed weight  $\lambda$ . If this weight were not fixed, then UCS presentation would eventually have no effect on the output of the adaptive element. This does not imply, however, that our model has definite memory. In the absence of incoming signals, a pathway will exhibit no change (since it will never be eligible for modification) no matter how long the period of inactivity lasts.

It is useful to compare the form of negative feedback employed by models resembling the Rescorla-Wagner/Widrow-Hoff rule with that of normalization schemes. In the former case, the feedback signal is the total stimulus strength,

$$\sum_{j=1}^n w_j(t)x_j(t),$$

whereas in the latter it is simply the sum of the weights,

$$\sum_{j=1}^n w_j(t).$$

Although one form of feedback is additive and the other is multiplicative, the major difference is that the Rescorla-Wagner/Widrow-Hoff rule uses information from the current stimulus pattern, whereas the normalization scheme does not. Without this information, the stimulus context effects observed experimentally cannot be produced. Further, if one is arguing for the cellular plausibility of a learning rule, then negative feedback in the form of total stimulus strength is easier to account for than feedback in the form of total synaptic efficacy. Since total stimulus strength is reflected in neurons by the membrane potential, it is plausible to hypothesize that this signal is available, at least approximately, at each synaptic site.

We have seen, then, that not only does the Rescorla-Wagner/Widrow-Hoff rule provide a strong model of classical conditioning and a powerful iterative method for solving sets of linear equations, it also solves some of the technical problems that always accompany the use of the Hebbian rule. Our model retains these advantages while accounting for some of the intratrial temporal structure of classical conditioning.

### Cellular Mechanisms

There is always a risk in speculating about cellular mechanisms for learning processes. On the one hand, not enough is known about the cellular changes that occur during associative learning to permit the construction of detailed models, and on the other hand, experimental progress in this area is occurring so rapidly that any postulated mechanism is likely to be soon invalidated by concrete fact. Despite these hazards, we feel that a discussion of our model in light of current electrophysiological and biochemical knowledge of cellular mechanisms can be of value, since the model is empirically supported at a behavioral level and is of interest from a theoretical perspective. In addition, the concepts of "eligibility" and "expectation" in our model are not only of critical importance in accounting for animal learning behavior and, we believe, essential for adaptive behavior of artificial systems but can be associated quite naturally, albeit speculatively, with certain processing capabilities of neurons.

There are four aspects of our rule to consider. First, the notion of eligibility would be realized if a synapse were "tagged" by a nonstimulating trace for some period of time after each discharge of the presynaptic cell. This indication of previous stimulation would be required to endure for a period on the order of at least a few seconds rather than a few milliseconds. This trace should remain local to the synapse. Second, some way of registering changes in the postsynaptic cell's firing rate from its previous level is required. This determines the reinforcement that facilitates or inhibits eligible synapses. The length of time over which the reference firing rate is determined is not critical but should be relatively long, perhaps with a time scale similar to that of eligibility. Third, it is necessary for the measure of eligibility, which is local to the synapse, to interact with the reinforcement signal, which is a global feature of the postsynaptic cell. This interaction should occur at each synapse. Finally, the result of the interaction between the eligibility of a synapse and the reinforcement level must regulate modifications of the transmission efficacy of that synapse.

The notion of a synaptic marker indicating previous presynaptic discharge could be realized either postsynaptically or presynaptically. We discuss a postsynaptic site for eligibility first. There is good evidence that in some cells the binding of a neurotransmitter to its receptor site regulates postsynaptic concentrations of an adenosine 3', 5'-monophosphate (cyclic AMP) or guanosine 3', 5'-monophosphate (cyclic GMP). It has been hypothesized that these cyclic nucleotides may mediate, as second messengers, the action of several neurotransmitters in generating slow postsynaptic potentials. This hypothesis is supported by several lines of electrophysiological and pharmacological evidence. For reviews see Greengard (1976), Nathanson (1977), or Rasmussen, Jensen, Lake, Friedmann, & Goodman (1975).

However, some studies have suggested that postsynaptic increases in cyclic nucleotide concentrations may have roles other than the generation of postsynaptic potentials. For example, it has been shown that the administration of cyclic AMP and cyclic GMP to cells in a sympathetic ganglion of

the bullfrog does not cause appreciable changes in membrane potential even though synaptic stimulation increases both cyclic AMP and cyclic GMP in these cells (Busis, Weight, & Smith, 1978). It has been suggested that in addition to the role cyclic nucleotides may play in simple neurotransmission, they may also carry more indirect messages that might, for example, mediate a stimulus trace that temporally links events in associative learning at a cellular molecular level (Woody, 1976).

Although the role of cyclic nucleotides in synaptic transmission and its regulation is not yet clear, it is evident that in some cells, and for some neurotransmitters, postsynaptic concentrations of cyclic nucleotides do reflect the amount of presynaptic stimulation received and can register previous stimulation for a time that is very long compared to the millisecond times of electrical activity. A difficulty, however, with the hypothesis that postsynaptic chemical concentrations provide stimulus traces as required by our model is that these traces would probably not remain local to their initiating synapses.

The locality of the trace suggests that a presynaptic site might be more plausible. Studies of the presynaptic mechanisms that are responsible for the nonassociative phenomena of post-tetanic facilitation and habituation suggest that the notion of eligibility could be represented presynaptically. Since these phenomena involve time scales much longer than that of electrical activity, we might postulate that some of the same mechanisms realize the notion of eligibility used in our model of classical conditioning. For example, intracellular concentration of free  $\text{Ca}^{2+}$  or  $\text{Ca}^{2+}$  conductance characteristics (e.g., voltage dependence) could provide relatively prolonged records of presynaptic activity. The mechanisms that result in post-tetanic facilitation or habituation for some temporal stimulus patterns might provide important record keeping facilities that operate whatever the stimulus characteristics are.

If eligibility were recorded presynaptically, then we would need to postulate some way in which the activity of the postsynaptic cell could influence the presynaptic terminal. Although it has been shown that postsynaptic activity can influence a presynaptic ter-

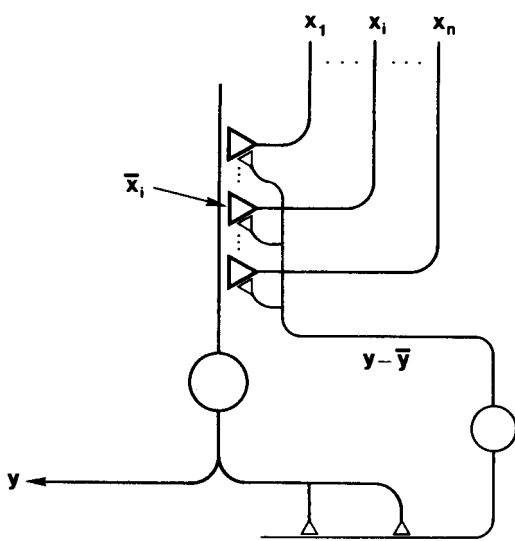


Figure 20. The adaptive element implemented via a feedback interneuron. (Eligibility  $\bar{x}_i$  is computed presynaptically, and the difference between actual and expected firing rates, computed by the feedback interneuron, modulates synaptic strengths through synaptosynaptic connections.)

minimal by altering the ionic content of the surrounding medium (Weight & Erulkar, 1976), we discuss instead mechanisms whose roles in plasticity are much better understood. These involve presynaptic facilitation via synapto-synaptic connections. Figure 20 shows a simple circuit in which presynaptic modulation is provided by extracellular feedback from the postsynaptic cell. The figure shows a single feedback interneuron, but a multisynaptic pathway is clearly also possible. In fact, the feedback pathway could pass through a brain region that integrates the signal with other information in a manner not accounted for by our model. For example, the signal may be integrated with other stimulus context information by the septo-hippocampal complex in a way similar to that suggested by Moore (1979). Figure 20 also shows episynaptic connections from the interneuron to all of the incoming fibers. As formulated here our model requires this feature, but it should be regarded as a convenient simplification. Fibers not contacted by the interneuron would not exhibit plasticity (or, at least, not plasticity of the same form), and episynaptic connections carrying signals from other than the postsynaptic cell

would permit processing of a form more complex than that considered here.

It remains to suggest how the feedback signal from the interneuron could represent the reinforcement signal, that is, the deviation of the postsynaptic cell's firing rate from previous levels. Perhaps the simplest possibility is that the interneuron, or the network of interneurons, responds only to changes in its input, as has been commonly observed for some cells responding to sensory stimuli. A more complex hypothesis would be that the presynaptic input from the feedback interneuron produces two superimposed effects on each synapse it modulates. One effect would be a fast change in the voltage dependence of the  $\text{Ca}^{2+}$  conductance so that depolarization would cause increased  $\text{Ca}^{2+}$  influx. The amplitude of this effect would depend on the eligibility of the modulated terminal determined by its previous history of depolarization. This increased influx of  $\text{Ca}^{2+}$  would facilitate the transmission effectiveness of the synapse by increasing transmitter release. The second effect would be a slower and less dramatic decrease in the peak  $\text{Ca}^{2+}$  conductance during depolarization that would cause decreased transmitter release. The magnitude of this second effect would also depend on the terminal's eligibility. If one assumed that these two effects linearly superimpose and that the effects of different discharges of the presynaptic terminal superimpose, then the resultant change in synaptic efficacy would depend on eligibility and the amount of change in activity of the feedback pathway as required by our model. Summing the fast positive and slow negative effects would produce a form of differentiation. The "expectation" would be represented by the negative component of the presynaptic effect.

We summarize this discussion of cellular mechanisms by making several observations. The model of classical conditioning we have presented was formulated on the basis of empirical evidence from behavioral experiments and from a sensitivity to the technical difficulties that have beset theoretical adaptive network studies. Although the evidence that this model or some variant of it might be implemented at the level of single neural units or simple neural circuits is not strong,

there is evidence that the relatively long lasting and nonstimulating memory of previous activity required by the notion of eligibility is indeed present at a cellular level. The monosynaptic phenomena of post-tetanic facilitation and habituation show that synapses themselves do possess nontrivial forms of short term memory that do not require one to hypothesize that reverberatory electrical activity provides the means for storing reflections of previous activity. Other evidence exists suggesting that within a single cell there can exist mechanisms for short-term stimulus traces as well as longer term memory (e.g., Alkon, 1979; Libet, Kobayashi, & Tanaka, 1975; von Baumgarten, 1970; Weight, Schulman, Smith, & Busis, 1979; Woody et al., 1974).

### Summary and Conclusions

Although the spirit of Hebb's theory still seems to be relevant, there is little support for the use of a literal interpretation of the Hebbian rule in adaptive network studies. As a model of classical conditioning, it is not up to the standard of sophistication now readily available in the learning theory literature. As a model of neural plasticity, it lacks experimental support and is based on a view of the processing capabilities of neurons and synapses that does not take into account the wealth of data now available. Although networks employing Hebbian rules have been successful in producing some interesting effects, their behavior is far from the level of sophistication required for complex tasks. Finally, models relying on Hebbian rules require rather ad hoc additional mechanisms to ensure stable and flexible behavior.

The Rescorla-Wagner/Widrow-Hoff rule, to which the perceptron and Uttley's informon are closely related, provides a more valid model of classical conditioning by incorporating stimulus context effects while at the same time cleanly solving a number of stability and saturation problems. That the Rescorla-Wagner equation was developed to account for animal learning behavior, whereas the nearly identical Widrow-Hoff rule was formulated to approximate the solutions of sets of linear equations, suggests

that these rules describe some ingredient essential for adaptive behavior. One important aspect of the Rescorla-Wagner/Widrow-Hoff rule's behavior is the extraction of reliable and nonredundant information that correlates with reinforcement. The experimental results regarding stimulus context effects in classical conditioning indicate that animals similarly form reliable and nonredundant associations.

We have presented a rule that preserves the properties of the Rescorla-Wagner/Widrow-Hoff rule but also incorporates the predictive nature of classical conditioning. The problems of making useful and accurate predictions seem to be solved by the ability to generate expectations. The actual events are then compared with those predicted, and appropriate incremental changes are made if the two differ. The Rescorla-Wagner equation does this while lumping together, as far as time of occurrence, all stimuli present on a trial. One contribution of the adaptive element developed here is to provide a mechanistic implementation of the descriptive Rescorla-Wagner theory of classical conditioning. In taking this lumped-trial theory to a mechanistic form in which system behavior is specified at all times within the trial, it becomes possible to make distinctions between inputs based on their relative time of occurrence. Rather than extracting reliable and nonredundant information that correlates with reinforcement, this rule extracts reliable, nonredundant, and early predictors of reinforcement. Moreover, an adaptive element employing this rule is able to use its sensitivity to predictive information to make predictions that occur earlier than the events predicted. A prediction made at the same time as, or later than, the event predicted is no more useful in guiding behavior than no prediction at all.

In addition, the adaptive element presented here preserves the simple account of stimulus substitution provided by the Hebbian rule. This is true because the UCR and CR share the same pathway—probably the simplest hypothesis accounting for the similarity of the UCR and the CR. Also as in the case of the Hebb model (and Uttley's informon), activity on any input pathway can cause changes in other pathways. This



produces some higher order conditioning effects and permits the element to extract regularities whose constituents have not been predetermined by a priori network structure. Unlike the Hebb and informon models, however, it extracts spatio-temporal rather than just spatial regularities.

Although we feel that our model includes some of the aspects of classical conditioning that have adaptive significance, the model is not a completely valid model of classical conditioning and obviously does not go beyond this restricted learning paradigm. It does not, for example, include the effects of experience on stimulus salience. In addition, like the original Rescorla-Wagner model, our model makes only ordinal predictions about behavioral data. Recent extensions of the Rescorla-Wagner model to deal with these shortcomings (Frey & Sears, 1978) can perhaps also be applied to our adaptive element model. Our theory also does not address stimulus representation problems. We have assumed that input signals arrive at an adaptive element on discrete pathways of fixed meanings. In a more sophisticated model, these meanings would be changed by circuits "upstream," as, for example, might occur in configural learning in which a compound stimulus is treated as a nonlinear combination of its parts. Our theory does not indicate how the adaptive mechanisms we have suggested can be extended to extract arbitrary nonlinear regularities.

The model presented here also does not address the issues arising from response-contingent reinforcement paradigms. Although the exact nature of the relationship between classical and instrumental conditioning remains elusive (e.g., Rescorla & Solomon, 1967), the attention given to temporal processing in our model makes possible extensions that incorporate response contingencies. Klopff's (in press, Note 1) theory, which forms the basis of several aspects of our model, incorporates response contingencies, and Sutton (1978, Note 4) has extended this theory to a single process view of expectation in classical and instrumental conditioning using an adaptive element closely related to that presented here. However, a thorough formal treatment of these issues is beyond the scope of the present article.

Also beyond this article's scope is a discussion of the kinds of behavior that can be expected from networks of adaptive elements like those proposed here. Can such a network perform sophisticated learning tasks? This question is central from a theoretical perspective and notoriously difficult to answer for any type of primitive component. We merely suggest that the predictive capabilities of the adaptive element presented here may permit adaptive networks to exhibit forms of behavior not yet obtained from network models. Our reason for believing this is that predictive capabilities permit response alternatives to be evaluated before overt action is taken.

The model we have developed need not be thought of as a neural model. It is supported at a behavioral level and has potentially significant theoretical implications. However, the search for neural analogs of behavioral conditioning continues to guide learning and memory research in the neurosciences. From our discussion of cellular mechanisms it is clear that although there is no shortage of machinery for implementing almost any learning model one might construct, there is evidence indicating that some of the essential aspects of our model could be implemented in a natural manner. The stimulus trace required by the notion of eligibility could involve either presynaptic or postsynaptic biochemistry. Our definition of reinforcement as the difference between actual and expected output levels could be realized via fast excitatory and slow inhibitory effects.

At the very least, our discussion of cellular mechanisms makes it clear that the concept of a neuron as a biological logic gate that still pervades much neural network theory is much too simple. Neurons and their synapses possess processing capabilities that can utilize relatively long-term histories of pre- and postsynaptic activity. In the terms of system theory, they possess a rich internal state space that can support behavior requiring nontrivial forms of memory. Neural network theorists have focused largely on synaptic weights as a form of memory and have postulated only relatively simple rules for controlling these memory variables. Other forms of memory have generally been as-



used to be metabolic and genetic, and, to a first approximation, not significantly implicated in computational behavior. Notably absent from the theoretical literature is a consideration of potentially powerful forms of synaptically local short-term memory and their possible role in synaptic modulation. The growing understanding of the role of biochemical mechanisms in synaptic action indicates that there is considerable internal memory linking events that occur at intervals of seconds, minutes, hours, and days. Moreover, this processing interacts with physiological events that occur in milliseconds. It seems certain that these mechanisms are crucially involved in neural plasticity. The model we have proposed takes a step toward recognizing the theoretical importance of the first few links in this chain.

## Reference Notes

1. Klopff, A. H. *Brain function and adaptive systems—A heterostatic theory*. Air Force Cambridge Research Laboratories Research Report AFCRL-72-0164, Bedford, Massachusetts, 1972 (AD742259). (There is a summary in *Proceedings of the International Conference on Systems, Man and Cybernetics*, IEEE Systems, Man and Cybernetics Society, Dallas, Texas 1974.)
2. Minsky, M., & Papert, S. *Artificial intelligence progress report*. MIT Artificial Intelligence Laboratory, Memo No. 252, January 1, 1972.
3. Minsky, M., & Selfridge, O. G. *Learning in random nets*. MIT Lincoln Laboratory, Lexington, Massachusetts, Report No. 54G-0024, June 1960.
4. Sutton, R. S. *A unified theory of expectation in classical and instrumental conditioning*. Unpublished undergraduate thesis, Stanford University, 1978.

## References

Albus, J. S. Mechanisms of planning and problem solving in the brain. *Mathematical Biosciences*, 1979, 45, 247–293.

Alkon, D. L. Voltage-dependent calcium and potassium ion conductances: A contingency mechanism for an associative learning model. *Science*, 1979, 205, 810–816.

Amari, S. A mathematical approach to neural systems. In J. Metzler (Ed.), *Systems Neuroscience*. New York: Academic Press, 1977. (a)

Amari, S. Neural theory of association and concept-formation. *Biological Cybernetics*, 1977, 26, 175–185. (b)

Anderson, J. A., Silverman, J. W., Ritz, S. A., & Jones, R. S. Distinctive features, categorical perception, and probability learning: Some applications of a neural model. *Psychological Review*, 1977, 85, 413–451.

Arbib, M. A., Kilmer, W. L., & Spinelli, D. N. Neural

models and memory. In M. R. Rosenzweig & E. L. Bennet (Eds.), *Neural mechanisms and memory*. Cambridge, Mass.: MIT Press, 1976.

Box, G., & Jenkins, G. *Time series analysis: Forecasting and control*. San Francisco: Holden Day, 1976.

Brindley, G. S. Nerve net models of plausible size that perform many simple learning tasks. *Proceedings of the Royal Society (London)*, 1969, B 174, 173–191.

Busis, N. A., Weight, F. F., & Smith, P. A. Synaptic potentials in sympathetic ganglia: Are they mediated by cyclic nucleotides? *Science*, 1978, 200, 1079–1081.

Dickinson, A., & Mackintosh, N. J. Classical conditioning in animals. *Annual Review of Psychology*, 1978, 29, 587–612.

Duda, R. O., & Hart, P. E. *Pattern classification and scene analysis*. New York: Wiley, 1973.

Egger, M. D., & Miller, N. E. Secondary reinforcement in rats as a function of informative value and reliability of the stimulus. *Journal of Experimental Psychology*, 1962, 64, 97–104.

Ellias, A. A., & Grossberg, S. Pattern formation, contrast control, and oscillations in the short term memory of shunting on-center off-surround networks. *Biological Cybernetics*, 1975, 20, 69–98.

Freeman, W. J. *Mass action in the nervous system*. New York: Academic Press, 1975.

Frey, P. W., & Sears, R. J. Model of conditioning incorporating the Rescorla-Wagner associative axiom, a dynamic attention process, and a catastrophe rule. *Psychological Review*, 1978, 85, 321–340.

Fukushima, K. A model of associative memory in the brain. *Kybernetik*, 1973, 12, 58–63.

Greengard, P. Possible role for cyclic nucleotides and phosphorylated membrane proteins in postsynaptic actions of neurotransmitters. *Nature*, 1976, 260, 101–108.

Grossberg, S. Some networks that can learn, remember, and reproduce any number of complicated space-time patterns. *Journal of Mathematics and Mechanics*, 1969, 19, 53–91.

Grossberg, S. Classical and instrumental learning by neural networks. In R. Rosen & F. Snell (Eds.), *Progress in theoretical biology*. New York: Academic Press, 1974.

Grossberg, S. Adaptive pattern classification and universal recoding: I. Parallel development and coding of neural feature detectors. *Biological Cybernetics*, 1976, 23, 121–134. (a)

Grossberg, S. Adaptive pattern classification and universal recoding: II. Feedback, expectation, olfaction, illusions. *Biological Cybernetics*, 1976, 23, 187–202. (b)

Hawkins, J. K. Self-organizing systems—a review and commentary. *Proceedings of the Institute of Radio Engineers*, 1961, 49, 31–48.

Hebb, D. O. *The organization of behavior*. New York: Wiley, 1949.

Hilgard, E. R. The nature of the conditioned response. I. The case for and against stimulus substitution. *Psychological Review*, 1936, 43, 366–385.

Hilgard, E. R., & Bower, G. H. *Theories of learning* (4th ed.). Englewood Cliffs, N.J.: Prentice-Hall, 1975.

Holland, J. H. *Adaptation in natural and artificial systems*. Ann Arbor: University of Michigan Press, 1975.

- Hull, C. L. *Principles of behavior*. New York: Appleton-Century-Crofts, 1943.
- John, E. R., & Schwartz, E. L. The neurophysiology of information processing and cognition. *Annual Review of Psychology*, 1978, 29, 1-29.
- Kamin, L. J. Predictability, surprise, attention and conditioning. In B. A. Campbell & R. M. Church (Eds.), *Punishment and Aversive Behavior*. New York: Appleton-Century-Crofts, 1969.
- Kandel, E. R. *Cellular basis of behavior*. San Francisco: W. H. Freeman, 1976.
- Kandel, E. R. *A cell-biological approach to learning* (Grass Lecture Monograph 1). Bethesda, Md.: Society for Neuroscience, 1978.
- Kellogg, W. N. Evidence for both stimulus-substitution and original anticipatory responses in the conditioning of dogs. *Journal of Experimental Psychology*, 1938, 22, 186-192.
- Kimmell, H. D. Instrumental inhibitory factors in classical conditioning. In W. F. Prokasy (Ed.), *Classical conditioning*. New York: Appleton-Century-Crofts, 1965.
- Klopf, A. H. Goal-seeking systems from goal-seeking components: Implications for AI. *The Cognition and Brain Theory Newsletter*, 1979, 3(2).
- Klopf, A. H. *The hedonistic neuron: A theory of memory, learning and intelligence*. Washington, D.C.: Hemisphere, in press.
- Kohonen, T. *Associative memory: A system theoretic approach*. Berlin: Springer, 1977.
- Lashley, K. S. The problem of serial order in behavior. In L. P. Jeffress (Ed.), *Cerebral mechanisms of behavior: The Hixon Symposium*. New York: Wiley, 1951.
- Libet, B., Kobayashi, H., & Tanaka, T. Synaptic coupling into the production and storage of a neuronal memory trace. *Nature*, 1975, 258, 155-157.
- Mackintosh, N. J. *The psychology of animal learning*. New York: Academic Press, 1974.
- Marr, D. A theory of cerebellar cortex. *Journal of Physiology*, 1969, 202, 437-470.
- Milner, P. M. The cell-assembly: Mark II. *Psychological Review*, 1957, 64, 242-252.
- Minsky, M. Steps toward artificial intelligence. In E. A. Feigenbaum & J. Feldman (Eds.), *Computers and thought*. New York: McGraw-Hill, 1963.
- Minsky, M. L., & Papert, S. *Perceptrons: An introduction to computational geometry*. Cambridge, Mass.: MIT Press, 1969.
- Moore, J. W. Brain processes and conditioning. In A. Dickinson & R. A. Boakes (Eds.), *Mechanisms of learning and motivation: A memorial volume to Jerzy Konorski*. Hillsdale, N.J.: Erlbaum, 1979.
- Nakano, K. Associatron—A model of associative memory. *IEEE transactions on systems, man, and cybernetics*, 1972, SMC-2, 380-388.
- Nathanson, J. A. Cyclic nucleotides and nervous system function. *Physiological Reviews*, 1977, 57, 157-256.
- Nilsson, N. J. *Learning machines*. New York: McGraw Hill, 1965.
- Prokasy, W. F., & Gormezano, I. The effect of US omission in classical aversive and appetitive conditioning of rabbits. *Animal Learning and Behavior*, 1979, 7, 80-88.
- Rasmussen, H., Jensen, P., Lake, W., Friedmann, N., & Goodman, D. B. P. In G. I. Drummond, P. Green-gard, & G. A. Robinson (Eds.), *Advances in cyclic nucleotide research*, Vol. 5. New York: Raven Press, 1975.
- Rescorla, R. A., & Solomon, R. L. Two-process learning theory: Relationships between Pavlovian conditioning and instrumental learning. *Psychological Review*, 1967, 74, 151-182.
- Rescorla, R. A., & Wagner, A. R. A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and non-reinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory*. New York: Appleton-Century-Crofts, 1972.
- Rochester, N., Holland, J. H., Haibt, L. H., & Duda, W. Tests on a cell assembly theory of the action of the brain, using a large digital computer. *IRE Transactions on Information Theory*, 1956, 2, 80-93.
- Rosenblatt, F. *Principles of neurodynamics*. New York: Spartan Books, 1962.
- Schneiderman, N. Interstimulus interval function of the nictitating membrane response of the rabbit under delay versus trace conditioning. *Journal of Comparative and Physiological Psychology*, 1966, 62, 397-402.
- Schneiderman, N., & Gormezano, I. Conditioning of the nictitating membrane of the rabbit as a function of the CS-US interval. *Journal of Comparative and Physiological Psychology*, 1964, 57, 188-195.
- Smith, M. C., Coleman, S. R., & Gormezano, I. Classical conditioning of the rabbit's nictitating membrane response at backward, simultaneous and forward CS-US intervals. *Journal of Comparative and Physiological Psychology*, 1969, 69, 226-231.
- Spinelli, D. N. OCCAM: A computer model for a content addressable memory in the central nervous system. In K. Pribram & D. Broadbent (Eds.), *The biology of memory*. New York: Academic Press, 1970.
- Stent, G. S. A physiological mechanism for Hebb's postulate of learning. *Proceedings of the National Academy of Science, USA*, 1973, 70, 997-1001.
- Sutton, R. S. Single channel theory: A neuronal theory of learning. *Brain Theory Newsletter*, 1978, 3, Nos. 3/4, 72-75.
- Uttley, A. M. The informon: A network for adaptive pattern recognition. *Journal of Theoretical Biology*, 1970, 27, 31-67.
- Uttley, A. M. The informon in classical conditioning. *Journal of Theoretical Biology*, 1975, 49, 355-376.
- Uttley, A. M. A two-pathway informon theory of conditioning and adaptive pattern recognition. *Brain Research*, 1976, 102, 23-35. (a)
- Uttley, A. M. Neurophysiological predictions of a two-pathway informon theory of neural conditioning. *Brain Research*, 1976, 102, 55-70. (b)
- Uttley, A. M. Simulation studies of learning in an informon network. *Brain Research*, 1976, 102, 37-53. (c)
- Uttley, A. M. *Information transmission in the nervous system*. London: Academic Press, 1979.
- von Baumgarten, F. J. Plasticity in the nervous system at the unitary level. In F. O. Schmitt (Ed.), *The neurosciences second study program*. New York: Rockefeller University Press, 1970.
- von der Malsburg, C. Self-organization of orientation sensitive cells in the striate cortex. *Kybernetik*, 1973, 14, 85-100.

- Weight, F. F., & Erulkar, S. D. Modulation of synaptic transmitter release by repetitive postsynaptic action potentials. *Science*, 1976, 193, 1023-1025.
- Weight, F. F., Schulman, J. A., Smith, P. A., & Busis, N. A. Long-lasting synaptic potentials and the modulation of synaptic transmission. *Federation Proceedings*, 1979, 38, 2084-2094.
- Widrow, G., & Hoff, M. E. *Adaptive switching circuits. Institute of Radio Engineers, Western Electronic Show and Convention, Convention Record*, 1960, Part 4, 96-104.
- Wigström, H. A neuron model with learning capability and its relation to mechanisms of association. *Kybernetik*, 1973, 12, 204-215.
- Wood, C. C. Variations on a theme by Lashley: Lesion experiments on the neural model of Anderson, Sil-verstein, Ritz, and Jones. *Psychological Review*, 1978, 85, 582-591.
- Woody, C. D. If cyclic GMP is a neuronal second messenger what is the message? In D. J. Jenden (Ed.), *Cholinergic mechanisms and psychopharmacology*. New York: Plenum, 1976.
- Woody, C. D. et al. Prolonged increases in resistance of neurons in cat motor cortex following extracellular iontophoretic application of acetylcholine (ACH) and intracellular current injection. *Federation of American Societies for Experimental Biology*, 1974, 33, p. 399.
- Zimmer-Hart, C. L., & Rescorla, R. A. Extinction of Pavlovian conditioned inhibition. *Journal of Comparative and Physiological Psychology*, 1974, 86, 837-845.

## Appendix A

### A Formal Description of the Simulated Model

In the following,  $R$  denotes the real numbers,  $R^+$  the non-negative reals, and  $[0, 1]$  denotes the closed real interval.

#### Components

- $n$  modifiable pathways labeled  $1, \dots, n$
- 1 fixed input pathway labeled 0

#### Descriptive Variables

##### Input variables:

For each  $i$ ,  $0 \leq i \leq n$ ,  $x_i(t) \in R$  denotes input level on input pathway  $i$ .

##### Output variables:

$y(t) \in [0, 1]$  denotes the output level of the adaptive element.

##### State variables:

$\bar{y}(t) \in [0, 1]$  is called the element's expectation, or expected output level.

For each  $i$ ,  $0 \leq i \leq n$ ,  $w_i(t) \in R$  denotes the transmission efficacy or connection weight of input pathway  $i$ .

For each  $i$ ,  $1 \leq i \leq n$ ,  $\bar{x}_i(t) \in R^+$  denotes the eligibility of input pathway  $i$ .

#### Equations of Interaction

$$w_i(t+1) = w_i(t) + c[y(t) - \bar{y}(t)]\bar{x}_i(t) \quad (A1)$$

$$\bar{x}_i(t+1) = \alpha\bar{x}_i(t) + x_i(t) \quad (A2)$$

$$y(t) = \sum_{j=0}^n w_j(t)x_j(t)$$

$$(\text{bounded to remain in } [0, 1]) \quad (A3)$$

$$\bar{y}(t+1) = \beta\bar{y}(t) + (1 - \beta)y(t) \quad (A4)$$

$$w_0(t) = w_0. \quad (A5)$$

#### Parameters

In all simulation experiments  $n = 4$ ,  $\beta = 0$ . The other parameters change from experiment to experiment (see below).

In all simulation experiments, rectangular pulse CSs and UCSs were represented as amplitude 1 rectangular pulses in  $x_i$  and  $x_0$ , respectively. A low level of normally distributed pseudo-randomly generated noise ( $M = .005$ ,  $SD = .03$ ) was then added into the rectangular pulses. Pulse lengths varied from experiment to experiment (see below). The intertrial interval was usually 50 time steps. CS durations were 5 time steps and UCS durations 10 time steps, except where otherwise noted.

#### Simulation Experiment Particulars, by Figure Number in Which Results Appeared

##### Figure 12:

$$c = .2; \alpha = .9; w_0 = .6$$

CS duration was 3 time steps. UCS duration was 30 time steps.

##### Figure 14:

$$c = .5; \alpha = .6; w_0 = .6$$

CS<sub>2</sub> duration was 10 time steps in Trials 21-35.

##### Figure 16:

$$c = .1; \alpha = .6; w_0 = .6 \text{ alternating with } .4$$

##### Figure 17:

$$c = .2; \alpha = .6; w_0 = .6; w_2(0) = .6$$

##### Figure 18:

$$c = .1; \alpha = .6; w_0 = .6; w_1(0) = .6$$

# Analysis of Steady State Behavior of the Rescorla-Wagner/Widrow-Hoff Predictor for a Simple Case

For simplicity we treat the continuous time case in which a trial consists of a single impulsive CS of amplitude  $\alpha$  at time  $t = 0$  and a single impulsive UCS of amplitude  $\lambda$  at time  $t = T$ . Letting the time functions  $x$  and  $z$  respectively denote the CS and UCS signals, then for  $t \geq 0$ :

$$x(t) = \alpha \delta_0(t) = \begin{cases} \alpha & \text{for } t = 0 \\ 0 & \text{otherwise,} \end{cases}$$

and

$$z(t) = \lambda \delta_T(t) = \begin{cases} \lambda & \text{for } t = T \\ 0 & \text{otherwise.} \end{cases}$$

Let the element's output be the linear result of one CS input pathway:

$$y(t) = \begin{cases} \alpha w(0) & \text{for } t = 0 \\ 0 & \text{otherwise.} \end{cases}$$

For continuous time the Rescorla-Wagner/Widrow-Hoff Predictor rule (Equation 12) becomes:

$$\frac{dw}{dt} = c[z - \bar{y}]\bar{x} \quad (\text{B1})$$

where  $\bar{x}$  and  $\bar{y}$  are the eligibility generated by  $x$  and the expectation generated by  $y$ , respectively. We assume that  $\bar{x}$  and  $\bar{y}$  are exponential traces of their respective variables. That is, let

$$\bar{x}(t) = \alpha e^{-\gamma t},$$

and

$$\bar{y}(t) = \alpha w(0) e^{-\xi t},$$

where  $\gamma$  and  $\xi$  are positive decay rates. Then Equation B1 becomes

$$\frac{dw(t)}{dt} = c[\lambda \delta_T(t) - \alpha w(0) e^{-\xi t}] \alpha e^{-\gamma t}. \quad (\text{B2})$$

Here we investigate the conditions under which a trial leaves the associative strength of the CS unchanged, that is, we ask what initial weight  $w(0)$  is such that  $w(t) = w(0)$  for some time  $t$  occurring after the trial. But when is a trial over? Weight changes can occur as long as  $\bar{x}$  and  $\bar{y}$  are not both equal to zero and thus can occur during

the ISI and after the UCS offset ( $t = T$ ). Since exponential traces never return to zero, we consider the case of an infinite intertrial interval and ask what  $w(0)$  should be so that

$$\lim_{t \rightarrow \infty} w(t) = w(0).$$

Integrating Equation B2 we obtain:

$$\begin{aligned} (\lim_{t \rightarrow \infty} w(t)) - w(0) &= \int_0^\infty \frac{dw(t)}{dt} dt \\ &= c \int_0^\infty [\lambda \delta_T(t) - \alpha w(0) e^{-\xi t}] \alpha e^{-\gamma t} dt \\ &= c \lambda \alpha e^{-\gamma T} - c \alpha^2 w(0) \left( \int_0^\infty e^{-(\gamma + \xi)t} dt \right) \\ &= c \lambda \alpha e^{-\gamma T} - \frac{c \alpha^2 w(0)}{\gamma + \xi}. \end{aligned}$$

Then  $w(0)$  must be such that

$$c \lambda \alpha e^{-\gamma T} = \frac{c \alpha^2 w(0)}{\gamma + \xi},$$

or,

$$w(0) = \frac{\lambda}{\alpha} e^{-\gamma T} (\gamma + \xi). \quad (\text{B3})$$

A trial of the form we have assumed is such that if the weight is the value given by Equation B3 at its commencement, then the weight will return (asymptotically) to this value after the trial. The weight can change during the trial, however. When viewed at the trial level, Equation B3 gives the asymptotic associative strength of the CS. It depends on the CS strength  $\alpha$ , the UCS strength  $\lambda$ , the ISI length  $T$ , and the characteristics  $\gamma$  and  $\xi$  of the traces. For more general types of trials, the asymptotic associative strength will also depend on the durations and shapes of the CS and UCS.

Received November 14, 1979

Revision received April 20, 1980 ■