Behavioural Brain Research, 4 (1982) 221–235
Elsevier Biomedical Press

Research Papers

SIMULATION OF ANTICIPATORY RESPONSES IN CLASSICAL CONDITIONING BY A NEURON-LIKE ADAPTIVE ELEMENT

ANDREW G. BARTO and RICHARD S. SUTTON

Department of Computer and Information Science, University of Massachusetts, Amherst, MA 01003 (U.S.A.)

(Received August 15th, 1980)
(Revised version received May 19th, 1981)
(Accepted June 2nd, 1981)

Key words: classical conditioning – prediction – learning – neural model – adaptation

SUMMARY

A neuron-like adaptive element is described that produces an important feature of the anticipatory nature of classical conditioning. The response that occurs after training (conditioned response) usually begins earlier than the reinforcing stimulus (unconditioned stimulus). The conditioned response therefore usually anticipates the unconditioned stimulus. This aspect of classical conditioning has been largely neglected by hypotheses that neurons provide single unit analogs of conditioning. This paper briefly presents the model and extends earlier results by computer simulation of conditioned inhibition and chaining of associations.

INTRODUCTION

The concept of single unit analogs of conditioning paradigms continues to be influential in the study of the neural basis of learning and memory. Hebb's suggestion that repeated and persistent pairing of presynaptic and postsynaptic discharges facilitates synaptic transmission is an example of attributing to single units the associative learning properties observed behaviorally in classical conditioning experiments [10]. Many theoretical models of plasticity in the nervous system are based on this postulate (see ref. 2 for a review). While a literal form of Hebb's hypothesis has not been vindicated by physiological evidence, there is ample support that forms of non-associative learning can take place via synaptic

modification [14, 15], and current research is directed toward a cellular explication of associative learning [1, 36, 37].

This paper focuses on a property of classical conditioning that might also be present at the cellular level: this is the anticipatory aspect of the temporal relationship between the conditioned response (CR) and the unconditioned stimulus (UCS). Since the conditioned stimulus (CS) must usually precede the UCS for an association to be formed, the CS is a predictive cue for the occurrence of the UCS. It is consistent with behavioral data to view the learning which takes place during classical conditioning as the learning of this predictive relationship (see, for example ref. 5). Equally well established is that the CR is usually initiated before the UCS [9, 23]. Being a response to the predictive CS, the CR anticipates the UCS and the unconditioned response (UCR). Although when very short interstimulus intervals (ISIs) are used, the CR often will not begin until after UCS onset, an anticipatory CR is the more usual case. Since a prediction must be available earlier than the event predicted in order to provide useful information, the anticipatory nature of the CR may be crucial to the adaptive significance of the behavior elicited in classical conditioning experiments.

The fact that anticipatory CRs are possible at all is problematic for many neural models. For example, in a Hebbian model of conditioning, the UCS is a strongly excitatory input to a neuron that also has the CS as an initially ineffective input. Before conditioning, the UCR is the neuron discharge produced by the UCS pathway. Pairing of the CS and UCS eventually increases the efficacy of the CS pathway until the CS also causes a discharge (the CR). Many mathematical interpretations of this view require simultaneous pairing of the UCS and CS signals at the neuron, thus implying an optimal ISI of zero. A suitable delay in the CS pathway is often suggested to bring this model closer to the behavioral data on ISI dependency [4, 30]. However, this delay also delays CR onset at least until the time of UCS onset, thereby preventing the CR latency from ever being shorter than the ISI required for conditioning.

Although the model we describe is still relatively coarse behaviorally compared to the wealth of experimental data available on classical conditioning, it makes several improvements over earlier neuron-like adaptive element models. The model was introduced by the authors in ref. 29 and originated in the work of Klopf [17–19] and Sutton [28] on modelling both classical and instrumental conditioning. It is also closely related to the model of Rescorla and Wagner [25]. We discuss its relationship to these models below.

THE MODEL

Fig. 1a shows a neuron-like adaptive element with an input pathway for the UCS and an input pathway for each stimulus capable of being associated with the UCS. These latter stimuli are the conditioned stimuli which we denote by CS_i,

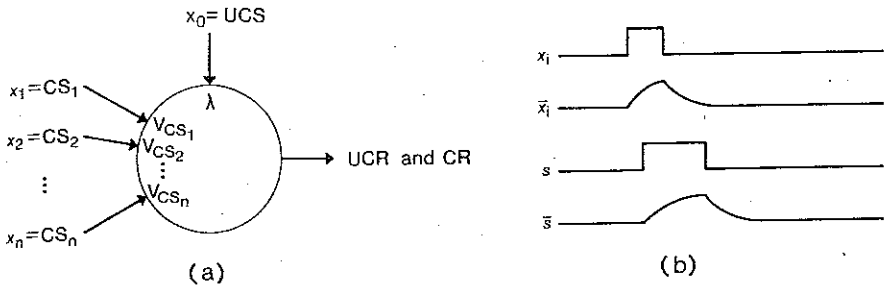

Fig. 1. a: a neuron-like adaptive element as an analog of classical conditioning. Each input pathway x_i has transmission efficacy V_{CS_i} corresponding to the associative strength of CS_i. The UCS is signalled via a pathway of fixed efficacy λ. Before conditioning, the element output contributes to the UCR, and after conditioning it contributes to both the UCR and the CR. See text for a discussion of rules for computing the element output and for updating the associative strengths. b: traces of the signals x_i and s (the temporal relationship of the signals shown is not intended to have a particular significance). The value of $\bar{x}_i$ at time t is a weighted average of the values of x_i over some preceding time interval, and similarly for $\bar{s}$. Illustrated here are exponential weightings. This causes the values for $\bar{x}_i$ and $\bar{s}$ to remain elevated for a time after the offsets of the corresponding signals.

i = 1,..., n. The associative strength of each CS_i with respect to the UCS is denoted by V_{CS_i} and characterizes the efficacy of the corresponding pathway. These associative strengths vary over time as functions of simulated classical conditioning paradigms. The efficacy of the UCS pathway is fixed at some constant value which we denote by λ. The element output level is assumed to contribute to both the UCR and the CR.

For each CS_i, i = 1,..., n, let x_i be a time function representing the presence or absence of that stimulus. That is, for each time t, $x_i(t) = 1$ if CS_i is present at time t, and $x_i(t) = 0$ otherwise. Similarly, let $x_0(t)$ indicate the presence or absence of the UCS at time t and let the associative strength at time t of CS_i be denoted by $V_{CS_i}(t)$. In addition, let s(t) denote the weighted sum of all the inputs to the element in which each input signal is weighted by the efficacy of its pathway; that is, let

$$s(t) = \lambda\, x_0(t) + \sum_{i=1}^{n} V_{CS_i}(t)x_i(t). \qquad (1)$$

The manner in which the output of the element is computed is not critical for the present discussion and, for simplicity, we assume that the output at time t is just s(t).

Several other variables are required in order to define the model. First, for each stimulus signal x_i, i = 1,..., 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 non-zero values of a separate variable $\bar{x}_i(t)$ 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 s. Let $\bar{s}(t)$ denote a weighted average of the values of the variable s over some time interval preceding t. Fig. 1b shows examples of these traces produced by exponentially weighted averages*. Although exponentially decaying traces are used in this article, traces of this form are not essential for the model behavior described. Other forms of traces produce similar but quantitatively different behavior. We use exponential traces merely because they are the simplest to generate.

The behavior of the adaptive element is therefore described by the values over time of the two variables s and $\bar{s}$, and the values of the 3 variables x_i, $\bar{x}_i$, and V_{CS_i} for each input pathway i = 1,..., 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,..., n,

$$V_{CS_i}(t + 1) = V_{CS_i}(t) + c[s(t) - \bar{s}(t)]\bar{x}_i(t) \tag{2}$$

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

We can describe the process given by Eqns. 1 and 2 as follows: activity on any input pathway i, i = 1,..., n, possibly causes an immediate change in the element output s 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 s differs from the value of the trace $\bar{s}$ of s. Thus, the traces $\bar{x}_i$ mark their corresponding pathways as being eligible for modification. In order to account for the temporal relationships observed in classical conditioning experiments, the eligibility traces must last for a time period on the order of several seconds in length.

The effectiveness of the reinforcement for the conditioning process depends on the difference $s(t) - \bar{s}(t)$ which determines how the eligible connections actually change. The simplest case, and the one used in our simulations, results from letting $\bar{s}(t) = s(t - 1)$. Then $s(t) - \bar{s}(t) = s(t) - s(t - 1)$ which is a discrete form of the rate-of-change of the variable s. The most important property of this difference is that it is zero while s is constant irrespective of the magnitude of s. This contributes to the stability of our model. Prolonged traces $\bar{s}$ like that shown in Fig. 1b can produce the same kind of effect and also filter out fast transient fluctuations of s so that associative strengths are not strongly influenced by the noise components of signals.

---

* In the computer simulations that produced the data shown below, we generated these traces using the first-order linear difference equations

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

$\bar{s}(t + 1) = \beta\bar{s}(t) + (1 - \beta)s(t)$

where α and β are constants such that $0 \leqslant \alpha, \beta < 1$.

An adaptive element operating according to the learning rule given by Eqs. 1 and 2 is able to increase its response in anticipation of increased stimulation because it uses stimulus trace variables $\bar{x}_i$ that are different from the stimulus-signalling variables x_i. The CR is produced by the CS signal, but learning is governed by the interaction of the trace of that signal with later element activity.

CLASSICAL CONDITIONING WITH A SINGLE CS

In order to understand the behavior of this model it is useful to consider the simplest special case of a single rectangular signal representing CS occurrence to the adaptive element. Fig. 2a shows an adaptive element analog of this situation. We assume that V_{CS} is initially zero and that the trace $\bar{s}$ takes the simplest form $\bar{s}(t) = s(t-1)$. The eligibility trace $\bar{x}$ is taken to have an exponential form as shown in Fig. 1b.

The rectangular CS signal causes an increase in the eligibility $\bar{x}$ of the CS pathway that persists for some time after the CS offset. The rectangular UCS signal causes a positive change in s at its onset and an equal but negative change at its offset. Since eligibility is greater at the time of the UCS onset than at its offset, the associative strength of the CS is caused to have a net increase: it increases a certain amount at the UCS onset and decreases by a lesser amount at the UCS offset. Fig. 2b shows the time courses of the signals involved.

After the start of conditioning, the CS, because its associative strength is no longer zero, causes an increase in the output level s. Hence, CS onset causes a transient increase in the reinforcement signal $s - \bar{s}$, and its offset causes a transient decrease. With additional trials the associative strength of the CS increases until the positive reinforcement of the UCS onset is counterbalanced by the negative reinforcement of the CS offset (Fig. 2c). We are assuming here, for simplicity, that the intertrial interval is long enough for the eligibility of the CS pathway to decay to zero between trials so that CS onset has no effect. Similarly, we are assuming that the UCS is long enough so that UCS offset has no effect on the associative strength.

The equilibrium associative strength V_{CS} attained by this process is a dynamic equilibrium. Except in the special case in which the CS offset and UCS onset occur at exactly the same time, V_{CS} continues to change during each trial, but eventually undergoes no *net* change per trial. By the asymptotic associative strength of a CS we therefore mean that value which eventually holds both before and after a trial. This value in general depends on the durations and amplitudes of the CS and UCS, the ISI, and the character of the traces $\bar{x}$ and $\bar{s}$. A mathematical analysis of a special case is given in ref. 29. Figure 3, trials 0–10, shows a typical acquisition curve plotting the associative strength value after each trial.

Notice from Fig. 2c that the value of s, the output of the element, shows a response to the CS and the UCS. The latter response corresponds to a component

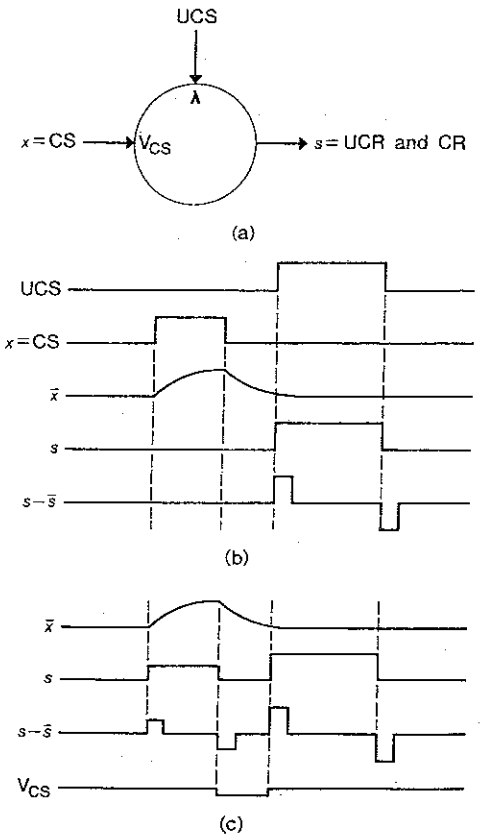

Fig. 2. Classical conditioning for a single CS. a: the adaptive element analog. The single CS x has variable associative strength V_{CS}, and the UCS has fixed strength λ. For simplicity, the output is simply the weighted sum of the input signals. b: time courses of the signals during the first trial. The associative strength V_{CS} increases due to the non-zero trace $\bar{x}$ coinciding with the positive difference $s - \bar{s}$ caused by the UCS onset. c: time courses of the signals after complete training. Since V_{CS} is now positive, CS occurrence causes changes in s. Then CS offset, coinciding with positive eligibility $\bar{x}_i$, causes a decrease in V_{CS}. At equilibrium, this decrease is exactly counterbalanced by the increase caused by UCS onset. Thus, after complete training, V_{CS} continues to change within each trial but undergoes no net change. The increase in s due to the CS represents the development of the CR.

of the UCR while the former represents a component of the CR. Note that this CR component begins earlier than the UCS indicating that the model provides a basis for the anticipatory nature of the CR. The time interval between the onset of this CR component and the arrival of the CS signal to the adaptive element is zero since we have assumed zero latency for the adaptive element response.

Another fact about the behavior of this model is that it produces an ISI-dependency similar to that found experimentally in animals. The asymptotic associative strength versus ISI curve is an inverted U with a maximum at the ISI

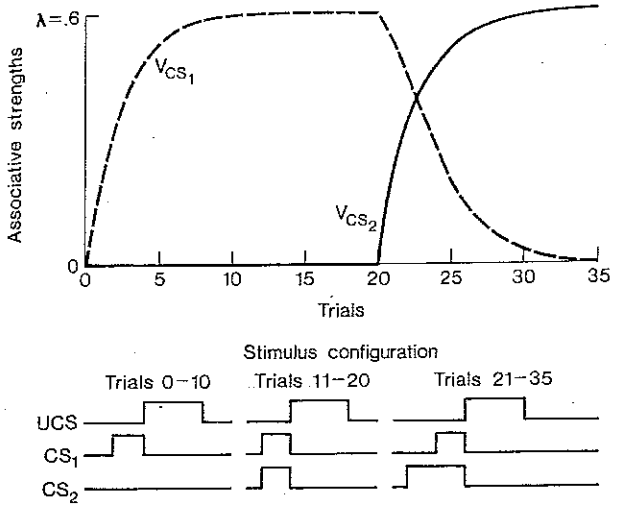

Fig. 3. Results of a computer simulation experiment. The associative strengths at the end of each trial are plotted. Changes in the associative strengths which occur within trials are not shown. Trials 0–10: presentation of CS_1 alone followed by the UCS resulted in the rise of the associative strength of CS_1 to the asymptotic level. Trials 11–20: CS_1 and CS_2 presented together followed by the UCS produced no change since CS_2 was redundant. This is the blocking paradigm. Trials 21–35: CS_2 began earlier than CS_1. The element became sensitive to the earlier predictor and lost sensitivity to the later.

equal to the duration of the CS [29]. Although in animal experiments the optimal ISI is roughly independent of overt CS duration (but see ref. 27), the model is consistent with experimental data if it is assumed that 'effective' or 'internal' CS duration is not the same as overt, external CS duration.

CLASSICAL CONDITIONING WITH SEVERAL STIMULI

In behavioral experiments 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 blocking. In part I of a typical blocking experiment, one stimulus, CS_1, is paired with a UCS at an appropriate ISI until the associative strength between CS_1 and the CR reaches its asymptotic value. In part II, CS_1 continues to be paired with the UCS, but another stimulus, CS_2, co-occurs with CS_1. Although CS_2 is appropriately paired with the UCS in part II, it conditions very poorly, if at all, compared to a control group without prior part I conditioning to CS_1. Effects of the associative strengths of context stimuli on conditioning occur in a variety of experimental paradigms, of which blocking, overshadowing, and conditioned inhibition are some of the prominent examples [11].

The results of a simulation of blocking are illustrated in trials 0–20 of Fig. 3. For the first 10 trials, CS_1 was presented alone and followed by the UCS. The associative strength V_{CS_1} quickly rose to the UCS level $\lambda = 0.6$. For trials 11–20, CS_1 was presented identically paired with CS_2, and both were followed by the UCS. During these trials V_{CS_1} and V_{CS_2} did not change (Fig. 3). Changes in V_{CS_2} were blocked since the signal s did not change while the CS pathway was eligible.

The adaptive element we have described is also able to extract from its input signals those which most reliably predict the UCS. For example, if CS_1 is paired with 100% of the UCSs while CS_2 is paired with a lesser percentage, then eventually CS_1 becomes completely dominant ($V_{CS_1} = \lambda$, $V_{CS_2} = 0$) even if CS_2 had been dominant initially (see ref. 29 for other details).

In a related experimental arrangement the model produces conditioned inhibition. If the occurrence of CS^+ alone is always followed by the UCS, but the co-occurrence of CS^+ and CS^- is never followed by the UCS, then CS^- becomes an inhibitor of the CR. The associative strength V_{CS^+} increases so that CS^+ produces a CR, but V_{CS^-} becomes negative so that a CR does not follow $CS^+ + CS^-$ (Fig. 4).

HIGHER ORDER CONDITIONING

One reason the Hebbian postulate has remained influential among theorists is that adaptive elements based on this postulate do not require specialized input pathways whose sole function is to carry signals telling when and how the connection weights on the other pathways should change (i.e. so-called 'teacher' inputs). Since the Hebbian postulate implies that the activity of any input pathway can cause changes in other pathways, pathways whose efficacies have become strengthened through previous training can further affect other pathways. A model with this property can produce behavior similar to higher order conditioning in animal learning: a previously conditioned CS can act as a UCS for a second CS. Since the reinforcement signal $s - \bar{s}$ of the element described here can be influenced by activity on any input pathway, the element also exhibits this property. When coupled to the anticipatory capabilities of our model, several novel consequences appear.

First, the adaptive element tends to respond to the *earliest* predictors of the UCS arrival. For example, assume CS_1 and CS_2 both end at the same time and are both always followed by reinforcement, but let CS_2 start earlier than CS_1. Then, even if CS_1 is dominant initially ($V_{CS_1} = \lambda$, $V_{CS_2} = 0$), eventually the earlier predictor CS_2 will completely dominate CS_1 (eventually $V_{CS_1} = 0$, $V_{CS_2} = \lambda$). See trials 21–35 of Fig. 3. Although both stimuli were presented in trials 11–20 and in trials 21–35, in the former case CS_2 was blocked by CS_1, while in the latter the associative strength of CS_2 increased quickly as the associative strength of CS_1 decreased. In the earlier trials CS_2 was redundant to CS_1, which had already

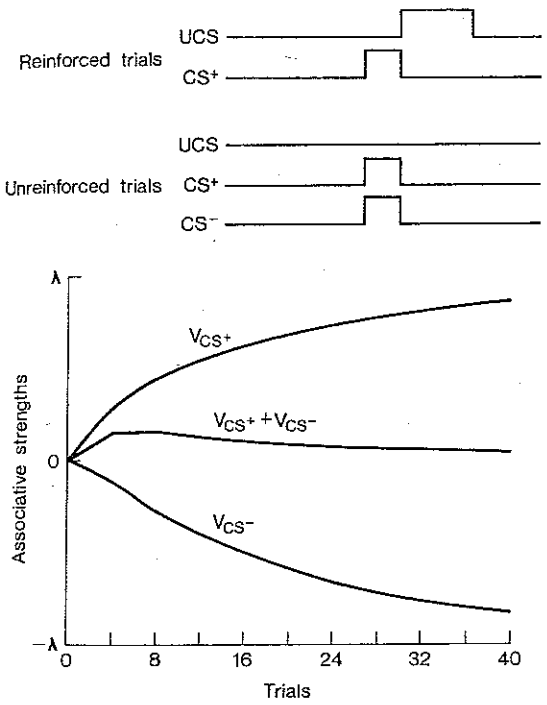

Fig. 4. Conditioned inhibition. CS^+ occurring alone predicted UCS occurrence, but CS^+ and CS^- occurring together predicted its absence. Alternating reinforced and unreinforced trials produced the learning curves shown (plotted points are in groups of 4 trials). Intratrial changes in associative strengths are not shown. Associative strengths were attained which sum to zero so that a response was produced to CS^+, but no response followed the occurence of CS^+ and CS^- together. CS^- became a conditioned inhibitor.

been conditioned, but in the later trials CS_2 provided important new information: it was the earliest indicator that the UCS would occur. This advantage, combined with the fact that CS_1 was totally redundant to CS_2, produced complete conditioning to CS_2 and the elimination of conditioning to CS_1.

This steady state is approached quickly and in an orderly manner. Very briefly, on each trial the associative strength V_{CS_2} increases and then decreases by a lesser amount for a net gain, while V_{CS_1} always decreases: V_{CS_2} increases because CS_2 predicts the onset of CS_1's excitation, and both V_{CS_1} and V_{CS_2} decrease at the offsets of CS_1 and CS_2. It is thus the facilitating effect of the onset of CS_1 which causes conditioning to CS_2.

A behavior of the adaptive element that is closely related to that just described is the ability to chain associations. Figure 5 shows an experimental arrangement in which 4 CSs with a particular temporal ordering were paired with a UCS. Also shown in Fig. 5 are the acquisition curves produced by computer simulation. The associative strength of CS_1, the CS immediately preceding the

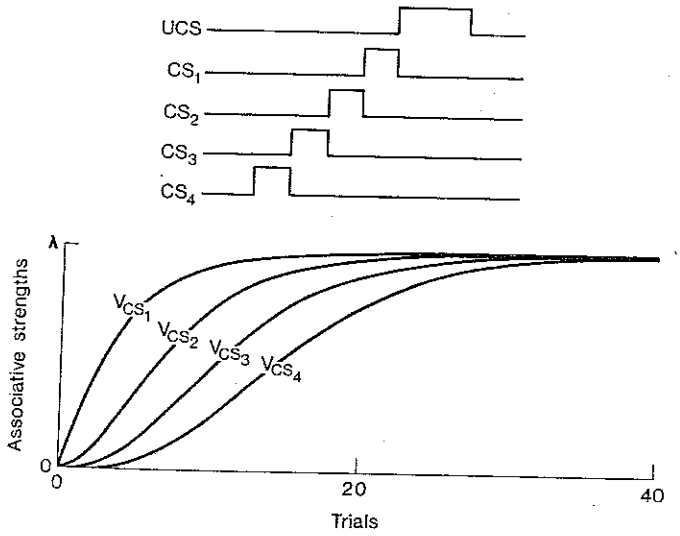

Fig. 5. Chaining of associations. Four CSs were presented to the adaptive element in trials having the temporal arrangement shown at the top of the figure. Conditioning occurred for each CS, with later CSs attaining asymptotic associative strength before the earlier CSs. The onset of a later CS comes to provide reinforcement for an earlier CS. Associations can therefore be formed for time intervals longer than the stimulus traces if intervening events reliably occur.

UCS, increased first. Then CS_1 onset acted as reinforcement for CS_2 which, in turn, came to reinforce CS_3 which then was able to reinforce CS_4. This caused the onset of the CR to move earlier in time as conditioning proceeded. For the temporal arrangement of the CSs shown in Fig. 5, the steady state was achieved in which all 4 associative strengths had the same value. For other temporal arrangements, variants of this basic behavior are produced. Notice that unlike the experiment described immediately above, the later predictors did not lose their associative strengths. This was due to the fact that here the CS offsets did not coincide. Chaining of associations in this manner permits conditioning to occur for ISIs much longer than those which can be spanned by a single stimulus trace as long as there are regularly occurring intervening events. Kehoe et al. [16] observed a strong effect of this nature for rabbit nictitating membrane response.

RELATIONSHIP TO OTHER MODELS

The adaptive element described here is closely related to the adaptive element proposed by Klopf [17–19]. Klopf suggested 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 many other adaptive element proposals in that eligibility is seen as being indicated in some way completely separate from the usual process of converting presynaptic activity to postsynaptic polarization. In the present model we have departed from Klopf's proposal in two ways. First, in the model described here an input pathway becomes eligible for modification whenever a signal arrives via that pathway. In Klopf's model, on the other hand, eligibility is triggered only if an input signal actually causes a suprathreshold response by the element. Second, in Klopf's model it is the value of s, which would correspond to neuronal membrane potential, that provides the reinforcement signal. The model presented here uses what amounts to the *change* in s to provide reinforcement. We have found that this modification of Klopf's proposal not only provides for stability but also produces the stimulus context effects.

The model described here also has strong connections to the Rescorla–Wagner model which was proposed to describe stimulus context effects [25]. The Rescorla–Wagner model is based on the often proposed view that learning occurs only when expectations are violated. According to this view, for example, blocking occurs since part I training creates an expectation of the UCS that is not disrupted in part II. We can similarly interpret the activity trace $\bar{s}$ as providing the expected value of the actual activity s as pointed out by Sutton [28]. Then the model described here can be interpreted as causing eligible pathways to be modified whenever the actual value of s differs from the expected value $\bar{s}$. The reinforcement signal $s - \bar{s}$ is a measure of how strongly the current activity (s) confirms or contradicts the previously formed expectation ($\bar{s}$). Thus, our model can be viewed as producing stimulus context effects for essentially the same reason they are produced by the Rescorla–Wagner model. Anticipatory aspects of classical conditioning and ISI dependency, however, are not addressed by the Rescorla–Wagner model since it does not distinguish between times *within* each trial.

Another adaptive element that is closely related to the Rescorla–Wagner model is Uttley's informon [30]. While this adaptive element does produce stimulus context effects and does distinguish between times within trials, it is not a valid model of the intratrial temporal structure of classical conditioning. In particular, it does not produce anticipatory conditioned responses or the appropriate ISI dependency. Clearly it was not Uttley's intention to produce such a detailed model of classical conditioning, and these deficiencies should not detract from his contribution. Nevertheless, we reiterate our position that the anticipatory nature of classical conditioning may be an essential aspect of animal learning.

It is not generally recognized that the Rescorla–Wagner model is essentially identical to a computational method for approximating the solution of a set of linear equations. This method has a long history in applied mathematics and was proposed two decades ago as an adaptive mechanism by Widrow and Hoff [34]. It is also closely related to Rosenblatt's perceptron [26], as is Uttley's informon. Duda and Hart [6] provide a good discussion of the details of this learning rule.

It is remarkable, in our opinion, that the Rescorla–Wagner theory, which was proposed to compactly describe a wide variety of effects observed in animal learning experiments, also provides an important algorithm with strong connections to useful areas of applied mathematics. One type of problem to which these mathematical methods are applied is that of constructing causal models of observed dynamic processes [3]. This suggests that the mathematical theory may have relevance for understanding the adaptive significance of animal conditioning behavior. The model described here is also related to this theory but produces anticipatory responses and higher order effects. These mathematical aspects of the model are more fully discussed in [29].

DISCUSSION

The behavior of the model we have presented has parallels with several aspects of animal behavior in classical conditioning experiments. Although it is not a fully adequate model of classical conditioning, it does account in a simple way for a variety of stimulus context effects, including blocking and conditioned inhibition, and produces several forms of higher order learning suggestive of secondary reinforcement phenomena. It also suggests a basis for some of the anticipatory aspects of classical conditioning. The model clearly does not address higher order modulatory influences such as those produced by attentional or stimulus salience factors. We have also not attempted to relate all of the properties of the model to behavioral data. For example, although closely related to the Rescorla–Wagner model, it has a number of differing implications that have not been fully explored. For example, we have not systematically explored the implications of our model with respect to some of the limitations of the Rescorla–Wagner model that have been pointed out by a number of authors (see, for example, refs. 7, 24 and 38).

While there is no direct evidence that the learning phenomena that are accounted for by our model occur at a cellular level, that possibility cannot be dismissed. There is evidence that mechanisms can exist within a single neuron for short term stimulus traces as well as longer term memory. Studies show that in many preparations, both vertebrate and invertebrate, synaptic modulation can depend on relatively complex temporal factors and that reverberatory activity is not necessary for some forms of short term memory [14, 22, 31–33]. It has been suggested, for example, that in addition to the role cyclic nucleotides and calcium ions may play in simple neurotransmission, they might also carry more indirect messages which might, for example, be involved in learning and memory [8, 35].

There is also evidence suggesting that one set of conditions may determine whether a pathway is able to undergo modulation, while another set of conditions determines whether these changes actually occur [21]. The model we have described illustrates that by separating the functions of stimulus signalling from the storage of stimulus traces, a simple mechanism can extract causal information

from the environment and make that information available early enough to provide a basis for decision making.

The behavior of the model also relies on a mechanism for detecting changes in the level of stimulation. The use of changes in the level of stimulation rather than its absolute level provides for complex temporal relationships and also solves many of the stability and saturation problems which have beset other theoretical models of synaptic plasticity. A cellular interpretation of the model therefore suggests the utility of a means for detecting changes in postsynaptic membrane potential or postsynaptic firing frequency, that is, for computing the term $s - \bar{s}$ in Eqn. 2. A biochemically natural way for detecting this change is to assume that $s - \bar{s}$ represents the level of a substance X which is formed at a rate V_f and decomposed at a rate V_d (cf. ref. 20). If membrane depolarization causes a fast increase in V_f and a slower increase in V_d, then the concentration of X will show a transient increase to any increase in depolarization. Similarly, if hyperpolarization causes a fast decrease in V_f and a slower decrease in V_d, then the level of X will show a transient decrease to any downward change in depolarization (Fig. 6). This manner of regulating a hypothetical substance X is similar to that proposed for the regulation of intracellular cyclic AMP [12]. If this view is correct, then both V_f and V_d for cyclic AMP are linked to membrane potential. There is also evidence that cyclic AMP concentration can be increased by depolarizing agents such as electrical stimulation in the absence of neurotransmitter action [13].

Finally, we have suggested that prolonged changes in synaptic efficacy may depend on the relatively long-term history of synaptic activity and on complex regulatory machinery. If this view is correct, then one would not expect to observe details of cellular associative learning without experimental conditions very precisely defined to control the internal state of the cell and the cellular context of the stimulation.

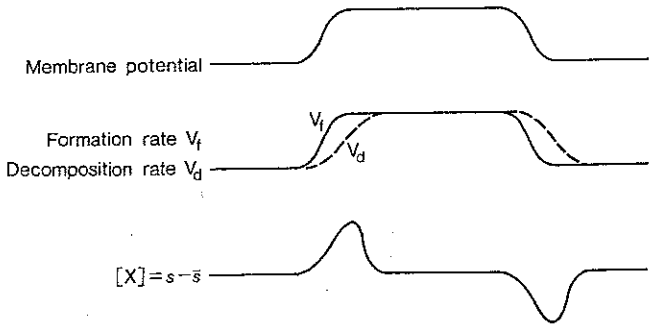

Fig. 6. A hypothetical biochemical mechanism for detecting changes in membrane potential (cf. ref. 20). If the formation rate V_f of substance X increases with depolarization and the decomposition rate V_d increases at a slower rate, then the concentration of X will show a transient increase to a sustained increase in depolarization. In a similar manner the concentration of X will signal the onset of sustained hyperpolarization.

ACKNOWLEDGEMENTS

This research has been supported by the Air Force Office of Scientific Research and the Avionics Laboratory (Air Force Wright Aeronautical Laboratories) through Contract F33615-77-C-1191. Some material in this paper is also presented in a paper by the authors published in *Psychological Review* (ref. 29). We are grateful to A.H. Klopf, J.W. Moore and D.N. Spinelli for helpful comments on earlier drafts.

REFERENCES

1 Alkon, D.L., Voltage-dependent calcium and potassium ion conductances: a contingency mechanism for an associative learning model, *Science*, 205 (1979) 810–816.
2 Arbib, M.A., Kilmer, W.L. and Spinelli, D.N., Neural models and memory. In M.R. Rosenzweig and E.L. Bennet (Eds.), *Neural Mechanisms and Memory*, MIT Press, Cambridge, 1976.
3 Box, G. and Jenkins, G., *Time Series Analysis : Forecasting and Control*, Holden Day, San Francisco, 1976.
4 Burke, W., Neuronal models for conditioned reflexes, *Nature (Lond.)*, 210 (1966) 269–271.
5 Dickinson, A. and Mackintosh, N.J., Classical conditioning in animals, *Ann. Rev. Psychol.*, 29 (1978) 587–612.
6 Duda, R.O. and Hart, P.E., *Pattern Classification and Scene Analysis*, Wiley, New York, 1973.
7 Frey, P.W. and Sears, R.J., Model of conditioning incorporating the Rescorla–Wagner associative axiom, a dynamic attention process, and a catastrophe rule, *Psychol. Rev.*, 85 (1978) 321–340.
8 Greengard, P. and Kuo, J.F., On the mechanism of the action of cyclic AMP, *Advanc. Biochem. Psychopharmacol.*, 3 (1970) 287–306.
9 Gormezano, I., Investigations of defense and reward conditioning in the rabbit. In A.H. Black and W.F. Prokasy (Eds.), *Classical Conditioning II: Current Research and Theory*, Appleton-Century-Crofts, New York, 1972.
10 Hebb, D.O., *The Organization of Behavior*, Wiley, New York, 1949.
11 Hilgard, E.R. and Bower, G.H., *Theories of Learning*, Prentice-Hall, Englewood Cliffs, NJ. 1975.
12 Kakiuchi, S., Ca^{2+} plus Mg^{2+}-dependent phosphodiesterase and its activator protein, *Advanc. Cyclic Nucleotide Res.*, 5 (1975) 163–178.
13 Kakiuchi, S., Rall, T.W. and McIlwain, H., The effect of electrical stimulation upon the accumulation of adenosine 3′,5′-phosphate in isolated cerebral tissue, *J. Neurochem.*, 16 (1969) 485–491.
14 Kandel, E.R., *Cellular Basis of Behavior*, W.H. Freeman, San Francisco, 1976.
15 Kandel, E.R., *A Cell-Biological Approach to Learning*, Grass Lecture Monograph 1, Soc. Neurosci., Bethesda, MD, 1978.
16 Kehoe, J.E., Gibbs, C.M., Garcia, E. and Gormezano, I., Associative transfer and stimulus selection in classical conditioning of the rabbit's nictitating membrane response to serial compound CSs, *Exp. Psychol., Anim. Behav. Proc.*, 5 (1979) 1–18.
17 Klopf, A.H., Brain function and adaptive systems – a heterostatic theory, Air Force Cambridge Research Laboratories Research Report AFCRL-72-0164, Bedford, MA, 1972 (AD 742259). (A summary in Proc. Int. Conf. Systems, Man and Cybernetics, IEEE Systems, Man and Cybernetics Society, Dallas, TX, 1974.)

18 Klopf, A.H., Goal-seeking systems from goal-seeking components: implications for AI, *The Cognition and Brain Theory Newsletter*, Vol. III, No. 2, 1979.

19 Klopf, A.H., *The Hedonistic Neuron: A Theory of Memory, Learning and Intelligence*, Hemisphere Publi., Washington, DC, in press.

20 Koshland, D.R., Jr., A model regulatory system: bacterial chemotaxis, *Physiol. Rev.*, 59 (1979) 811–862.

21 Krasne, G.B., Extrinsic control of intrinsic neuronal plasticity: an hypothesis from work on simple systems, *Brain Res.*, 140 (1978) 197–216.

22 Libet, B., Kobayashi, H. and Tanaka, T., Synaptic coupling into the production and storage of a neuronal memory trace, *Nature (Lond.)*, 258 (1975) 155–157.

23 Mackintosh, N.J., *The Psychology of Animal Learning*, Academic Press, New York, 1974.

24 Prokasy, W.F. and Gormezano, I., The effect of US omission in classical aversive and appetitive conditioning of rabbits, *Animal Learn. Behav.*, 7 (1979) 80–88.

25 Rescorla, R.A. and Wagner, A.R., A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and non-reinforcement. In A.H. Black and W.F. Prokasy (Eds.), *Classical Conditioning II: Current Research and Theory*, Appleton-Century-Crofts, New York, 1972.

26 Rosenblatt, F., *Principles of Neurodynamics*, Spartan Books, New York, 1962.

27 Schneiderman, N., Interstimulus interval function of the nictitating membrane response of the rabbit under delay versus trace conditioning, *J. comp. physiol. Psychol.*, 62 (1966) 397–402.

28 Sutton, R.S., *A Unified Theory of Expectation in Classical and Instrumental Conditioning*, Undergraduate Thesis, Stanford, 1978.

29 Sutton, R.S. and Barto, A.G., Toward a modern theory of adaptive networks: expectation and prediction, *Psychol. Rev.*, 88 (1981) 135–170.

30 Uttley, A.M., *Information Transmission in the Nervous System*, Academic Press, London, 1979.

31 von Baumgarten, F.J., Plasticity in the nervous system at the unitary level. In F.O. Schmitt (Ed.), *The Neurosciences Second Study Program*, Rockefeller University Press, New York, 1970.

32 von Baumgarten, F.J. and Fukuhara, T., The role of the interstimulus interval in heterosynaptic facilitation in *Aplysia californica*, *Brain Res.*, 16 (1969) 369–381.

33 Weight, F.F., Schulman, T.A., Smith, P.A. and Busis, N.A., Long-lasting synaptic potentials and the modulation of synaptic transmission, *Fed. Proc.*, 38 (1979) 2084–2094.

34 Widrow, G. and Hoff, M.E., Adaptive switching circuits. In *1960 IRE WESCON Convention Record, Part 4*, 1960, pp. 96–104.

35 Woody, C.D., If cyclic GMP is a neuronal second messenger what is the message? In D.J. Jenden (Ed.), *Cholinergic Mechanisms and Psychopharmacology*, Plenum, New York, 1976.

36 Woody, C.D., Carpenter, D.O., Gruen, E., Knispel, J.D., Crow, T.J. and Black-Cleworth, P., Prolonged increases in resistance of neurons in cat motor cortex following extracellular iontophoretic application of acetylcholine (ACH) and intracellular current injection, *Fed. Proc.*, 33 (1974) 399.

37 Woody, C.D., Vassilevsky, N.N. and Engel, J., Jr., Conditioned eye blink: unit activity at coronal-precruciate cortex of the cat, *J. Neurophysiol.*, 33 (1970) 851–864.

38 Zimmer-Hart, C.L. and Rescorla, R.A., Extinction of Pavlovian conditioned inhibition, *J. comp. physiol. Psychol.*, 86 (1974) 837–845.