Prediction in Classical Conditioning: An Adaptive Element Model†

Andrew G. Barto
and
Richard S. Sutton
Department of Computer and Information Science
University of Massachusetts
Amherst, Massachusetts 01003
COINS Technical Report No. 80-02

†This research has been supported by the Air Force Office of Scientific Research and the Air Force Avionics Laboratory through Contract No. F33615-77-C-1191. Some material in this paper is also presented in a paper by the authors to be published in Psychological Review. We are grateful to A.H. Klopf J.W. Moore, and D.N. Spinelli for helpful comments on earlier drafts.

SUMMARY

There is growing evidence that one function of synaptic biochemical mechanisms may be to permit one set of conditions to determine whether a pathway is able to undergo modulation and another set of conditions to determine whether these changes actually occur. We describe a neuron-like adaptive element which uses this principle to model the predictive nature of behavioral conditioning. An important feature of classical conditioning is that the response which occurs after training (conditioned response) usually occurs earlier than the reinforcing stimulus (unconditioned stimulus). The conditioned response therefore predicts, or anticipates, the unconditioned stimulus. This aspect of classical conditioning has been largely neglected by hypotheses that neruons provide single unit analogs of conditioning. The model is strongly supported by behavioral data on prediction as well as other properties of classical conditioning such as interstimulus interval dependency, blocking, and conditioned inhibition. We discuss the possibility that this model is implemented at a cellular level by biochemical mechanisms.

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 synapses 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 the review ref. 2). While a literal form of Hebb's hypothesis has not been vindicated by physiological evidence, there is ample support that forms of nonassociative learning can take place via synaptic modification (e.g., refs. 14, 15), and current research is directed toward a cellular explication of associative learning (e.g., refs. 1, 42, 43).

This paper focuses on a property of classical conditioning which might also be present at the cellular level: this is the predictive aspect of classical conditioning. Classical conditioning effectiveness depends on the interstimulus interval (ISI) which is the time interval between the conditioned stimulus (CS) onset and the unconditioned stimulus (UCS) onset. Asymptotic associative strength between the CS and the conditioned response (CR) is usually an inverted U-shaped function of this interval, being zero at simultaneous presentation, maximal at intermediate values (that depend on the particular response system), and then falling to zero at longer ISIs. Thus, since the 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. 6). Equally well established behaviorally, but less often

noted than ISI dependency, is that the CR usually precedes the UCS (Fig. 1).[22] Being a response to the predictive CS, the CR anticipates, or predicts, the UCS and the unconditioned response (UCR). 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.

Several neural theorists have suggested mechanisms for the ISI dependency of classical conditioning (e.g., refs. 4, 34), but in many cases these mechanisms also prevent the CR from occurring before the UCS. For example, in a Hebbian style 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 (e.g., refs 4, 34). However, this delay also <u>necessarily</u> delays the CR until the time of the UCS occurrence (Fig. 2). It prevents the CR latency from being shorter than the ISI required for conditioning.

We describe here a model of prediction in classical conditioning in the form of a neuron-like adaptive element which learns to increase its rate of response in anticipation of increased stimulation by becoming sensitive to predictive cues. The model also accounts for stimulus context effects, including blocking and conditioned inhibition, and produces several forms of higher order learning suggestive of secondary reinforcement phenomena. Fur-

ther, it solves many stability and saturation problems which have beset other models of neural plasticity. The model was first introduced by the authors in ref. 31 and originated in the work of Klopf[16,17,18] and Sutton[30] on modelling both classical and instrumental conditioning. We briefly discuss some cellular possibilities suggested by the model which are especially relevant due to recent advances in understanding the complexity of synaptic biochemistry.

The model is a combination of two sets of ideas. First, it is closely related to the neural hypothesis by Klopf[16,17,18] that the temporal characteristics of conditioning, both classical and instrumental, can be produced if one set of conditions makes synapses <u>eligible</u> for modification of their transmission efficacies, but <u>actual modifications</u> 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 which does not participate directly in the electrical signalling of the cell such as a transient increase in the concentration of a particular chemical. While there is as yet no direct physiological support for this hypothesis, it does rest on an idea for which there is a growing body of evidence; namely, 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.[20]

The other set of ideas incorporated in the model presented here are due to Rescorla and Wagner. Their descriptive model of the effects of stimulus context in classical conditioning[26] is based on the view that learning

occurs only when expectations are violated. Our model can be thought of in this way but differs from the Rescorla-Wagner model by being able to describe some of the fine temporal structure of conditioning, such as its predictive nature, which is outside the scope of the Rescorla-Wagner model.

THE MODEL

Fig. 3 shows an 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, $i = 1,\ldots, n$. The associative strength of each CS_i is denoted by V_{CS_i} and characterizes the efficacy of the corresponding pathway. The efficacy of the UCS pathway is fixed at some value which we denote by λ. The element output indicates the occurrence of the UCR and the occurrence of the CR.

For each CS_i, $i = 1,\ldots, n$, 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 signal is weighted by the efficacy if its pathway (cf. neuron membrane potential). 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, $i = 1,\ldots, n$, 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 $\bar{x}_i$ 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. 4 shows examples of these traces produced by exponentially weighted averages.†

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 three variables x_i, $\bar{x}_i$, and V_{CS_i} for each input pathway $i = 1,\ldots, n$. In terms of these variables, the model takes the form of a set of difference equations for successively generating the values of the associative strenghts: for each i, $i = 1,\ldots, n$,

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

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

We can describe the process given by Eqs. 1 and 2 as follows: Activity on any input pathway i, $i = 1,\ldots, n$, possibly causes an immediate change in

---

†In the computer simulations which produced the data shown below we generated these traces using the first-order linear difference equations

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

$$\bar{s}(t + 1) = \gamma\bar{s}(t) + \delta s(t)$$

where α, β, γ and δ are positive constants with $0 < \alpha, \gamma < 1$.

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. 4 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.

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 $\bar{x}_i$ which are different from the stimulus-signalling variables x_i. That is, in neural terms, instead of a previous occurrence of a CS being recorded by prolonged reverberatory electrical activity, it is recorded by a synaptically local and electrically non-stimulating trace. The CR is produced by the electrical CS

signal, but learning is governed by the interaction of the non-electrical trace with later cellular 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 the CS. Fig. 5a 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. 4.

The rectangular CS signal causes an increase in the eligibility $\bar{x}$ of the CS pathway which 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. 5b 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. 5c). 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 its 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. 31. Fig. 6, trials 0 - 10, shows a typical acquisition curve plotting the associative strength value after each trial.

Notice from Fig. 5c that the value of s, the output of the element, shows a response to the CS and the UCS. The latter response corresponds to the UCR while the former represents the CR. Note that the CR occurs earlier than the UCS indicating that the model preserves the predictive nature of classical conditioning. In this case the CR latency is zero since we have assumed zero latency for the adaptive element response.

Another fact about the behavior of our 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 equal to the duration of the CS (see ref. 31). Although in animal experiments the optimal ISI is roughly independent of overt CS duration (but see ref. 29), our 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.

STIMULUS CONTEXT BEHAVIOR

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 inhibitions are some of the prominent examples[11].

The results of a simulation of blocking are illustrated in trials 0 - 20 of Fig. 6. For the first 10 trials CS_1 was presented alone and followed by the UCS. The asymptotic associative strength V_{CS_1} quickly rose to the UCS level $\lambda = .6$. 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. 6). Changes in V_{CS_2} were blocked since the signal s did not change while the CS_2 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. 31 for other details).

In a related experimental arrangement our 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. 17).

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 "teacher" inputs. It is not necessary to fix from the start the source of a special signal capable of causing changes in connection weights. Any correlations among the input signals to the element will tend to be reflected in the connection weight values. In particular, since 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 our model can be influenced by activity on any input pathway, our model also exhibits this property. When coupled to the predictive capabilities of our model, several novel consequences appear.

First, the adaptive element tends to find 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. 6. 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 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 orderly. 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 closely related behavior is the ability to chain associations. Fig. 8 shows an experimental arrangement in which four CSs with a particular temporal ordering were paired with a UCS. Also shown in Fig. 8 are the acquisition curves produced by computer simulation. The associative strength of CS_1, the CS immediately preceding the 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. 8, the steady state was achieved in which all four 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. This capa-

bility of our model is not in agreement with all animal behavior data, and we have not yet thoroughly explored this aspect of our model's descriptive validity. We feel that it is significant, however, that this type of higher order learning behavior can be achieved with a single simple mathematical model using the same mechanisms required to model other aspects of classical conditioning.

RELATIONSHIP TO OTHER MODELS

As we have indicated, our model is a variant of Klopf's hypothesis which emphasizes the temporal relationships involved in classical and instrumental conditioning. In the present model we have departed from Klopf's hypothesis in two ways. First, in our model 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, which 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 theory not only provides for stability but also produces the stimulus context effects.

Our model also has strong connections to the Rescorla-Wagner model which was proposed to describe stimulus context effects[26]. 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[30]. Then our

model 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 confirms or contradicts the previously formed expectation or prediction. Thus, our model can be viewed as producing stimulus context effects for essentially the same reason they are produced by the Rescorla-Wagner model. Prediction and ISI dependency, however, are not addressed by the Rescorla-Wagner model since it does not distinguish between times within each trial. In addition, the more complex temporal structure of our model opens the possibility that it may not suffer from all of the shortcomings of the Rescorla-Wagner model (see refs. 8, 24, 44).

Another adaptive element that is closely related to the Rescorla-Wagner model is Uttley's Informon[32,33,34,35,36]. 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 predictive 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 predictive nature of classical conditioning is 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, an example of a gradient descent minimization method, has a long history in applied mathematics and was proposed two decades ago as an adaptive mechanism by Widrow and Hoff[40]. It is also closely related to Rosenblatt's Perceptron[28] as is Uttley's Informon.

Duda and Hart[7] 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 (e.g., ref. 3). This suggests that the mathematical theory may have relevance for understanding the adaptive significance of animal conditioning behavior. Our model is also related to this theory but produces predictive responses and higher order effects. The mathematical aspects of our model are more fully discussed in ref. 31.

CELLULAR MECHANISMS

Despite recent advances, it is still premature to propose a testable cellular model of associative conditioning. However, even though we have emphasized the behavioral validity of our model and have purposefully referred to it as an adaptive element rather than a neural model, we briefly speculate about how cellular mechanisms could provide similar computational machinery.

There is ample evidence that mechanisms can exist within a single neuron for short term stimulus traces as well as longer term memory. Recent 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[14,21,37,38,39]. Many of these phenomena involve very complex interactions between neurotransmitters, cyclic nculeotides, calcium ions, and ionic conductances. For ex-

ample, there is good evidence in a variety of systems that cyclic AMP, cyclic GMP, and calcium ions act as intracellular second messengers which mediate cell activation by extracellular first messengers (for reviews see refs. 23 and 25). Although it is hypothesized that in neurons these second messengers mediate the generation of slow postsynaptic potentials, this hypothesis cannot account for a variety of data[5,25]. It has been suggested 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[9,41].

A biochemically natural way for detecting changes in postysnaptic membrane potential, that is for computing the term $s - \bar{s}$ in Eq. 2, 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. 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. 9). This is similar to the mechanism proposed for the generation of bacterial chemotaxis by which bacteria climb nutrient gradients[19].

This manner of regulating a hypothetical substance X is similar to that proposed for the regulation of intracellular cyclic AMP[12]. According to this view, neurotransmitter-receptor binding activates adenylate cyclase which increases cyclic AMP concentration and also mobilizes calcium. Increased Ca^{2+} concentration within the cell activates cyclic nucleotide phosphodiesterase which brings about a rapid decrease in cyclic AMP concentration to basal levels. 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].

There are also numerous plausible mechanisms for providing the electrically non-stimulating stimulus traces which we have used in our model. For example, the phenomena of post-tetanic facilitation and habituation involve time scales much longer than that of electrical activity[14]. We might postulate that mechanisms which produce post-tetanic facilitation or habituation for some stimulus patterns might provide important record keeping facilities which operate whatever the stimulus characteristics are. This would imply a presynaptic stimulus trace mechanism. A presynaptic site for synaptic modulation is not only in accord with much data on synaptic plasticity but also provides for the locality of the trace required by our model. In a previous paper[31] we suggested that interaction between presynaptic stimulus traces and the reinforcement signal could take place via extracellular feedback involving interneurons and synapto-synaptic contacts.

There is also no shortage of possibilities for postsynaptic stimulus trace mechanisms. The necessity for such traces to remain local to their corresponding synaptic sites suggests that various intracellular ionic pools could be involved. For example, there is evidence indicating that calcium within a cell is not uniformly distributed throughout the cytosol but remains largely confined to mitochondria[25,27]. We might therefore postulate that the synaptically local traces $\bar{x}_i$ required by our model involve subsynaptic mitochondrial calcium concentration.

If we can invoke some of the cellular data to suggest a possible basis for our model, then the required computations could be performed at a cellular level by steps of the following character: 1) Postsynaptic cyclic AMP

concentration reflects changes in postsynapitc potential due to changes in both its rate of formation and its rate of decomposition following synaptic activation. 2) Synaptic activity causes influx of Ca^{2+} which is rapidly taken up by mitochondria to provide a long lasting trace of synaptic activation. 3) Cyclic AMP regulates Ca^{2+} release from mitochondrial pools making Ca^{2+} available in the cytosol. 4) Ionic conductances are modified through membrane protein phosphorylation via the joint action of cyclic AMP dependent protein kinases and a subsequent Ca^{2+}-dependent process.

CONCLUSION

We have clearly gone beyond the available data in suggesting cellular mechanisms for our model of classical conditioning: Associative conditioning has not been observed in a robust form at the cellular level, and the biochemical mechanisms we have enlisted have not all been observed to occur together within any single preparation. However, once these limitations of the above discussion are clearly recognized, there remain several significant observations.

First, the model we have presented is strongly supported by behavioral data as a model of classical conditioning. Although it is not a fully adequate model of classical conditioning, it does account in a simple way for a variety of phenomena, the one most strongly emphasized here being the predictive nature of the CR. Our 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 our model to behavioral data. For example, although our model is closely related to the Rescorla-Wagner model, it has a number of differing implications which have not been fully explored. While there is no direct evidence that the

learning phenomena that are accounted for by our model occur at a cellular level, that possibility exists and we have briefly explored it. We emphasize again, however, that what we have called an adaptive element may not correspond to a single cell.

Second, we have illustrated some of the consequences of a mechanism in which one set of conditions determines whether or not a pathway is modified while another set of conditions actually causes the modification. 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.

Third, we have discussed a form of heterosynaptic facilitation mediated not by postsynaptic discharge but rather by changes in postsynaptic membrane potential. This can be interpreted as providing a mechanism by which learning takes place only when "expectations" are violated and could be mediated by the intracellular diffusion of a postsynaptic chemical. The use of activity changes rather than absolute levels provides for complex temporal relationships and also solves many of the stability and saturation problems which have beset other theoretical models of synaptic plasticity.

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.

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, The MIT Press, 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, 210 (1966) 269-271.

5. Busis, N. A., Weight, F. F. and Smith, P. A., Synaptic potentials in sympathetic ganglia: are they mediated by cyclic nucleotides? Science, 200 (1978) 1079-1081.

6. Dickinson, A. and Mackintosh, N. J., Classical conditioning in animals, Ann. Rev. Psychol., 29 (1978) 587-612.

7. Duda, R. O. and Hart, P. E., Pattern Classification and Scene Analysis, Wiley, New York, 1973.

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

9. Greengard, P. and Kuo, J. F., On the mechanism of the action of cyclic AMP, Advan. Biochem. Psychopharmacol., 3 (1970) 287-306.

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, New Jersey, 1975.

12. Kakiuchi, S., Ca^{2+} plus Mg^{2+}-dependent phosphodiesterase and its activator protein. Advan. 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, Society for Neuroscience, Bethesda, Maryland, 1978.

16. 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: Proceedings of the International Conference on Systems, Man and Cybernetics, IEEE Systems, Man and Cybernetics Society, Dallas, Texas, 1974).

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

18. Klopf, A. H., The Hedonistic Neuron, to be published, Hemisphere Publishing Co., Washington, D. C., 1980.

19. Koshland, D. R. Jr., A model regulatory system: bacterial chemotaxis, Physiological Reviews, 59 (1979) 811-862.

20. Krasne, G. B., Extrinsic control of intrinsic neuronal plasticity: an hypothesis from work on simple systems, Brain Research, 140 (1978) 197-216.

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

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

23. Nathanson, J. A., Cyclic nucleotides and nervous system function. Physiological Reviews, 57 (1977) 157-256.

24. Prokasy, W. F. and Gormezano, I., The effect of US omission in classical aversive and appetitive conditioning of rabbits, Animal Learning and Behavior, 7 (1979) 80-88

25. Rasmussen, H. and Goodman, D. B. P., Relationships between calcium and cyclic nucleotides in cell activation, Physiological Reviews, 57 (1977) 421-509.

26. 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.

27. Rose, B. and Loewenstein, W. R., Permeability of a cell junction and the local cytoplasmic free ionized calcium concentration: a study with Aequorin, J. Membrane Biol. 28 (1976) 87-119.

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

29. Schneiderman, N., Interstimulus interval function of the nictitating membrane response of the rabbit under delay versus trace conditioning, Journal of Comparative and Physiological Psychology, 62 (1966) 397-402.

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

31. Sutton, R. S. and Barto, A. G., Toward a modern theory of adaptive networks: Expectation and prediction. To be published, Psychological Review, 1980.

32. Uttley, A. M., The informon: A network for adaptive pattern recognition. Journal of Theoretical Biology, 27 (1970) 31-67.

33. Uttley, A. M., The informon in classical conditioning, Journal of Theoretical Biology, 49 (1975) 355-376.

34. Uttley, A. M., A two-pathway informon theory of conditioning and adaptive pattern recognition, Brain Research, 102 (1976) 23-35.

35. Uttley, A. M., Simulation studies of learning in an informon network, Brain Research, 102 (1976) 37-53.

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

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

38. von Baumgarten, F. J. and Fukuhara, T., The role of the interstimulus interval in heterosynaptic facilitation in Aplysia Californica, Brain Research, 16 (1969) 369-381.

39. Weight, F. F., Schulman, T. A., Smith, P. A. and Busis, N. A., Long-lasting synaptic potentials and the modulation of synaptic transmission, Federation Proceedings, 38 (1979) 2084-2094.

40. Widrow, G. and Hoff, M. E., Adaptive switching circuits. In 1960 IRE WESCON Convention Record, Part 4, (1960) 96-104.

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

42. 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, Federation of American Societies for Experimental Biology, 33 (1974) p. 399.

43. Woody, C. D., Vassilevsky, N. N. and Engel, J. Jr., Conditioned eye blink: unit activity at coronal-precruciate cortex of the cat, Journal

of Neurophysiology, 33 (1970) 851-864.

44. Zimmer-Hart, C. L. and Rescorla, R. A., Extinction of Pavlovian conditioned inhibition, Journal of Comparative and Physiological Psychology, 86 (1974) 837-845.

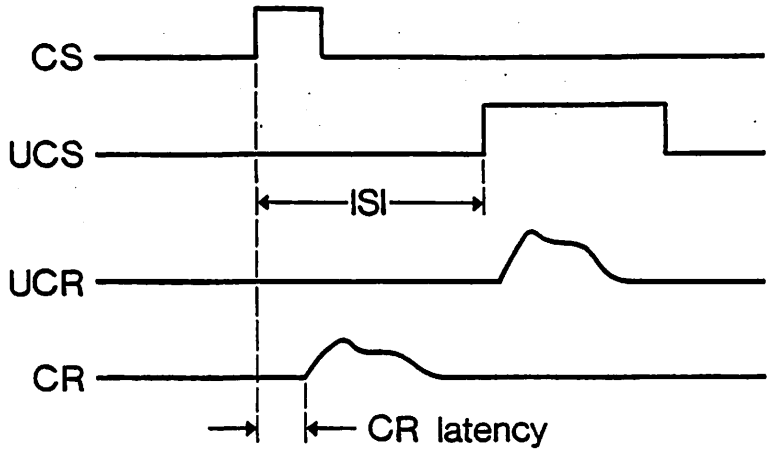


Fig. 1. Time relationships involved in classical conditioning. The interstimulus interval (ISI) is the time between CS onset and UCS onset. The CR latency is the time between CS onset and CR onset. Usually the CR latency is much shorter than the ISI so that the CR occurs earlier than the UCS.[22]

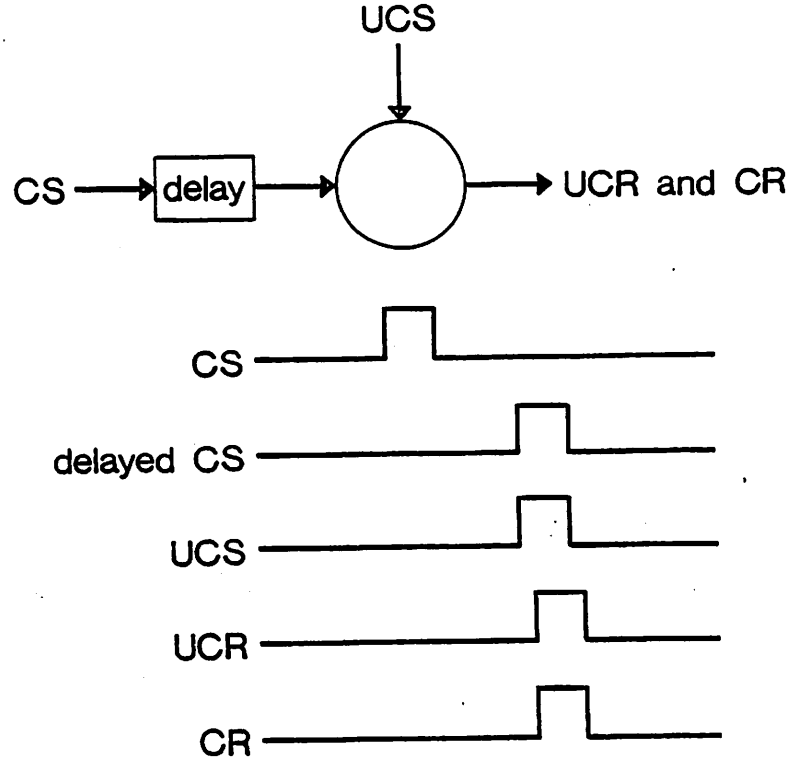


Fig. 2. Hebbian style analog of classical conditioning. If conditioning requires simultaneous apiring of the CS and UCS signals at the neuron, then a delay in the CS pathway can cause conditioning to occur only if the CS precedes the UCS but also prevents the CR from occurring before the UCS.

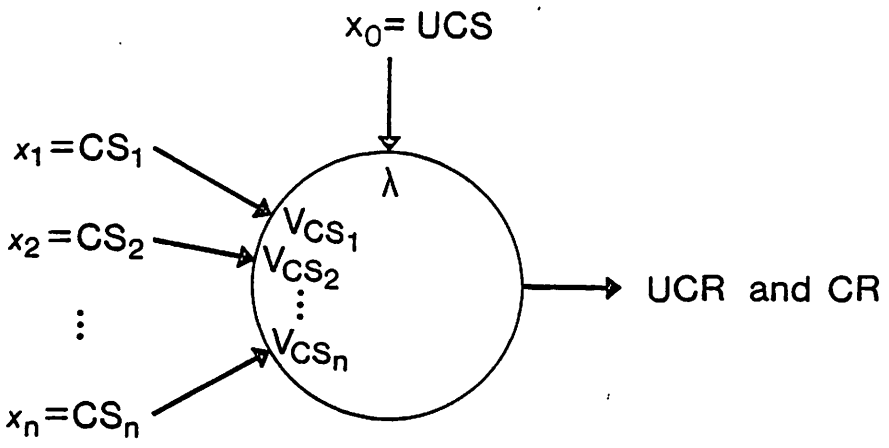


Fig. 3. A neuron-like adaptive element as an analog of classical conditioning. Each input pathway x_i has transmission efficacy V_{CS_i} corresponding to the associative strength of CS_i. The UCS is signalled via a pathway of fixed efficacy λ. Before conditioning, the element output corresponds to the UCR, and after conditioning it corresponds 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.

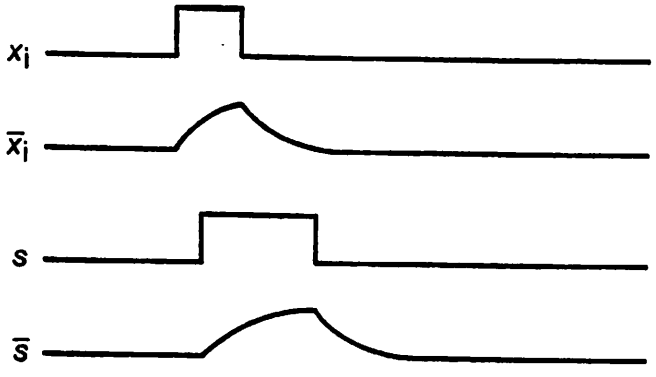


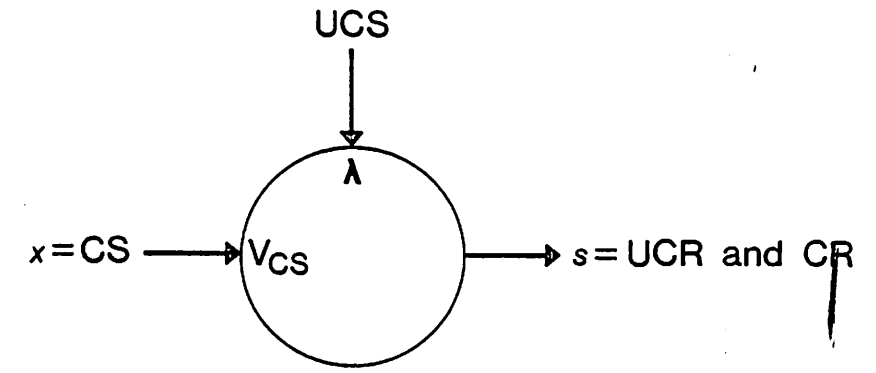
Fig. 4. Traces of the signals x_i and s (the temporal relationship of the signals shown is not intended to have a particular significance). The value of $\bar{x}_i$ at time t is a weighted average of the values of x_i over some preceding time interval, and similarly for $\bar{s}$. Illustrated here are 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.

Fig. 5. Classical conditioning for a single CS.

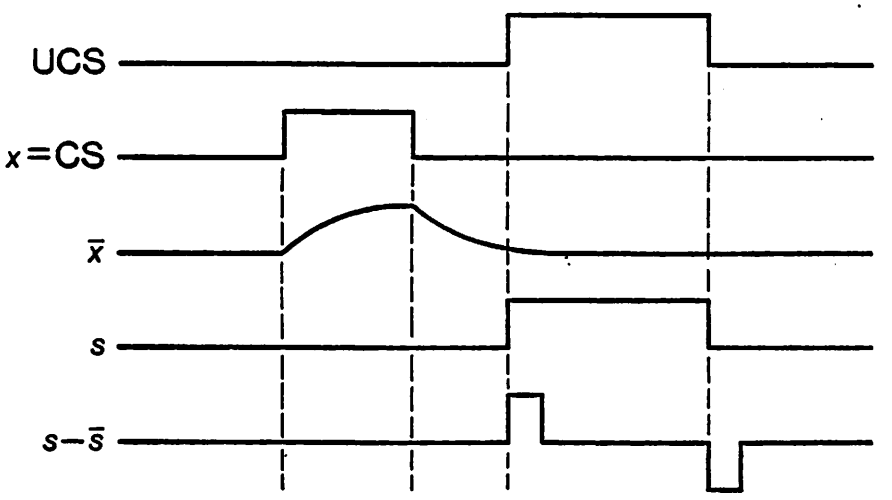
(a) The adaptive element analog. The single CS x has variable associative strength V_{CS}, and the UCS has fixed strength λ. For simplicity, the output is simply the weighted sum of the input signals.

(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.

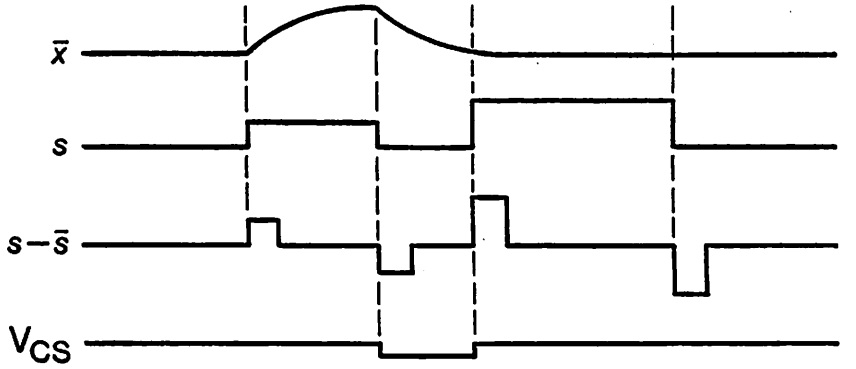
(c) Time courses of the signals after complete training. Since V_{CS} is now positive, CS occurrence causes changes in s. Then CS offset, coinciding with positive eligibility $\bar{x}_i$, causes a decrease in V_{CS}. At equilibrium, this decrease is exactly counterbalanced by the increase caused by UCS onset. Thus, after complete training, V_{CS} continues to change within each trial but undergoes no net change.



(a)



(b)



(c)

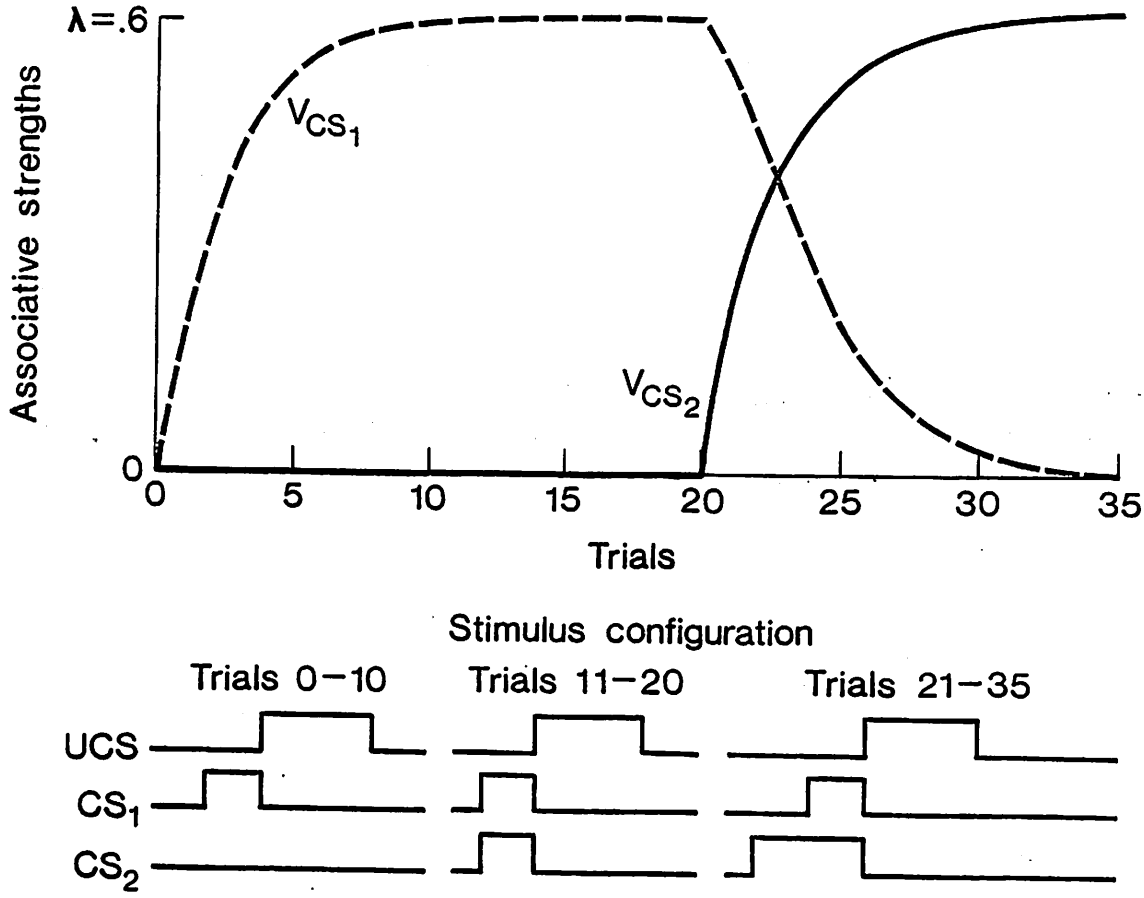


Fig. 6. Results of a computer simulation experiment. The associative strengths at the end of each trial are plotted. Changes in the associative strengths which occur within trials are not shown.

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.

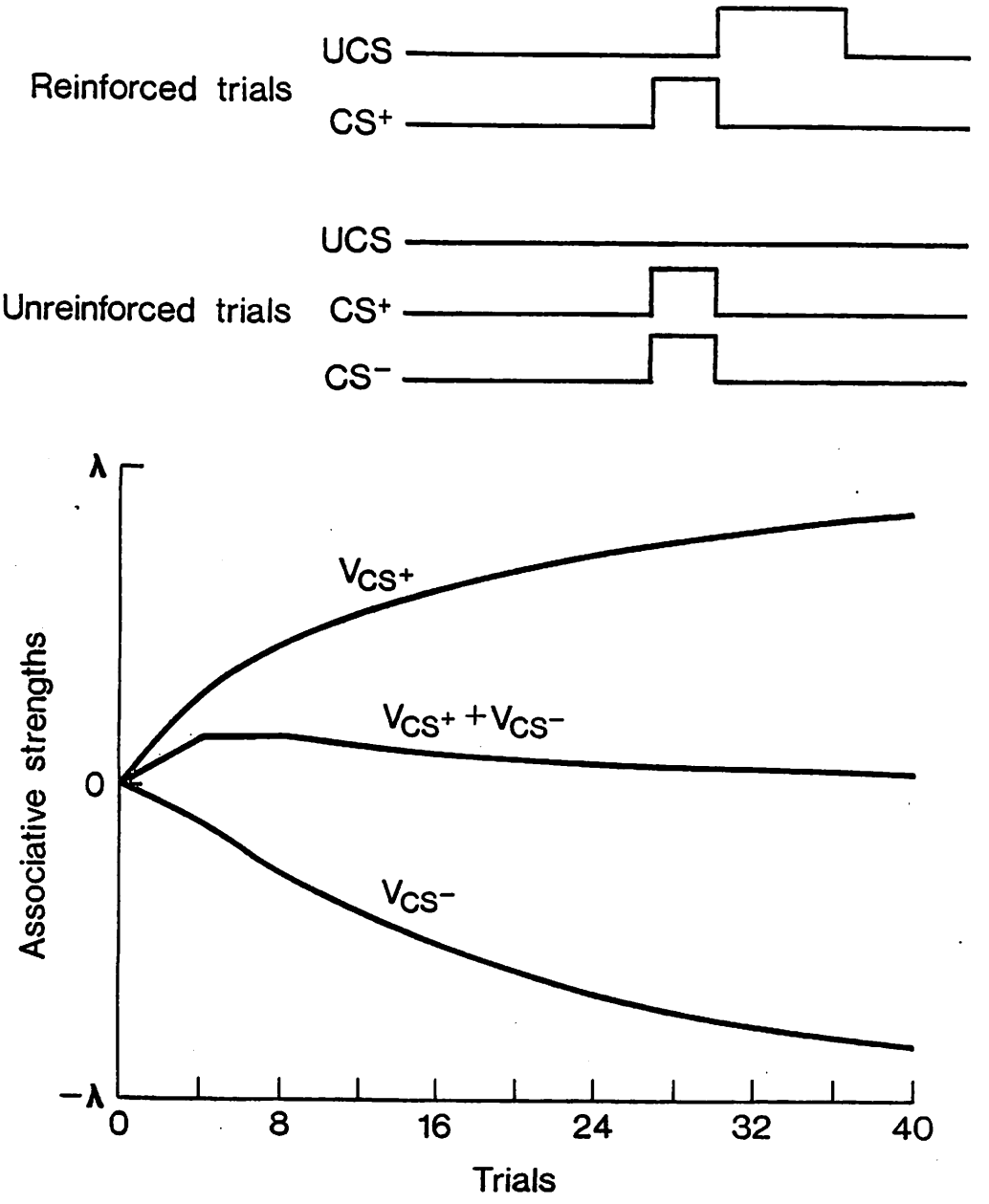


Fig. 7. Conditioned inhibition. CS^+ occurring alone predicted UCS occurrence, but CS^+ and CS^- occurring together predicted its absence. Alternating reinforced and unreinforced trials produced the learning curves shown (plotted points are in groups of four trials). Intratrial changes in associative 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 occurrence of CS^+ and CS^- together. CS^- became a conditioned inhibitor.

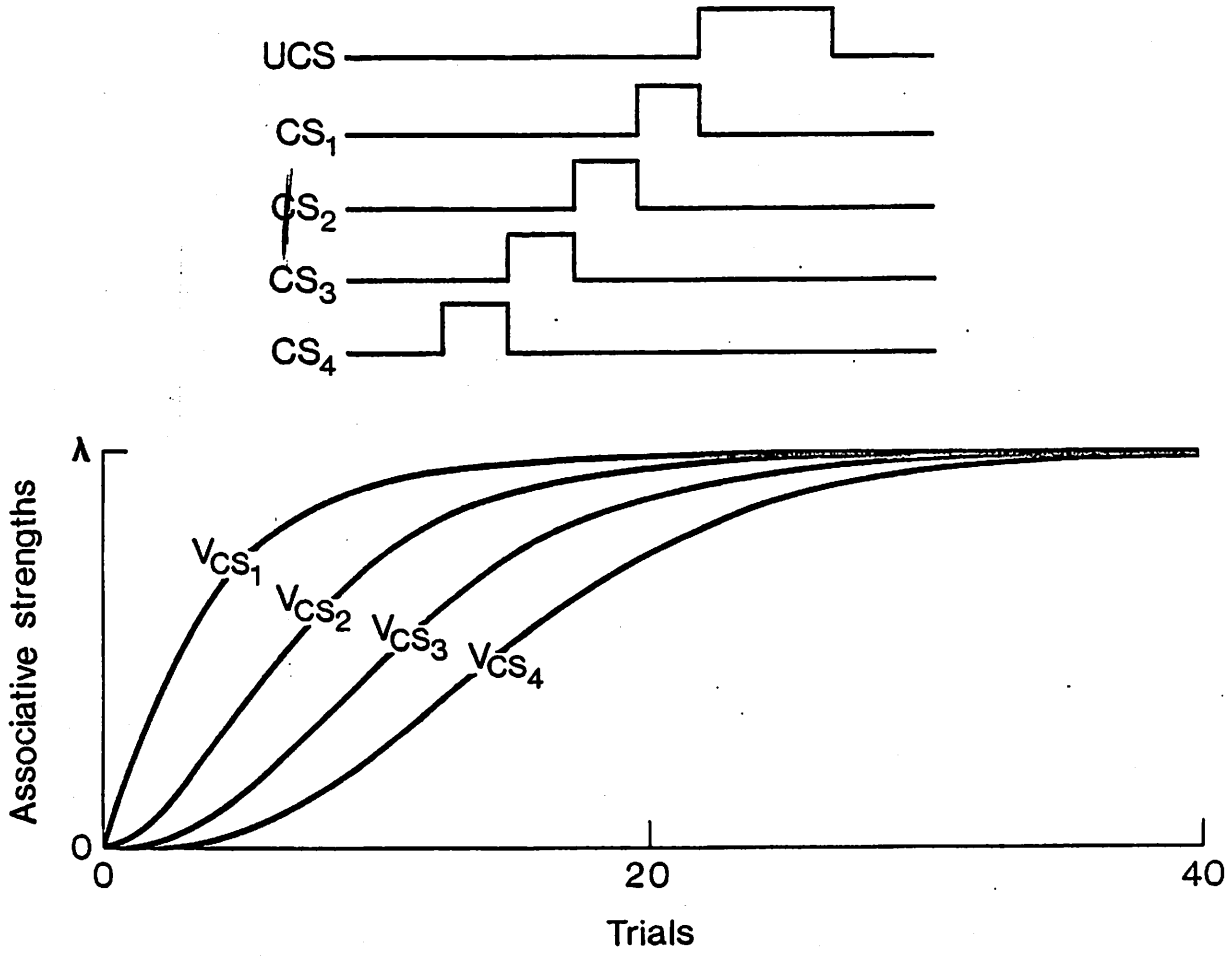


Fig. 8. Chaining of associations. Four CSs were presented to the adaptive element in trials having the temporal arrangement shown at the top of the figure. Conditioning occurred for each CS, with later CSs attaining asymptotic associative strength before the earlier CSs. The onset of a later CS comes 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.

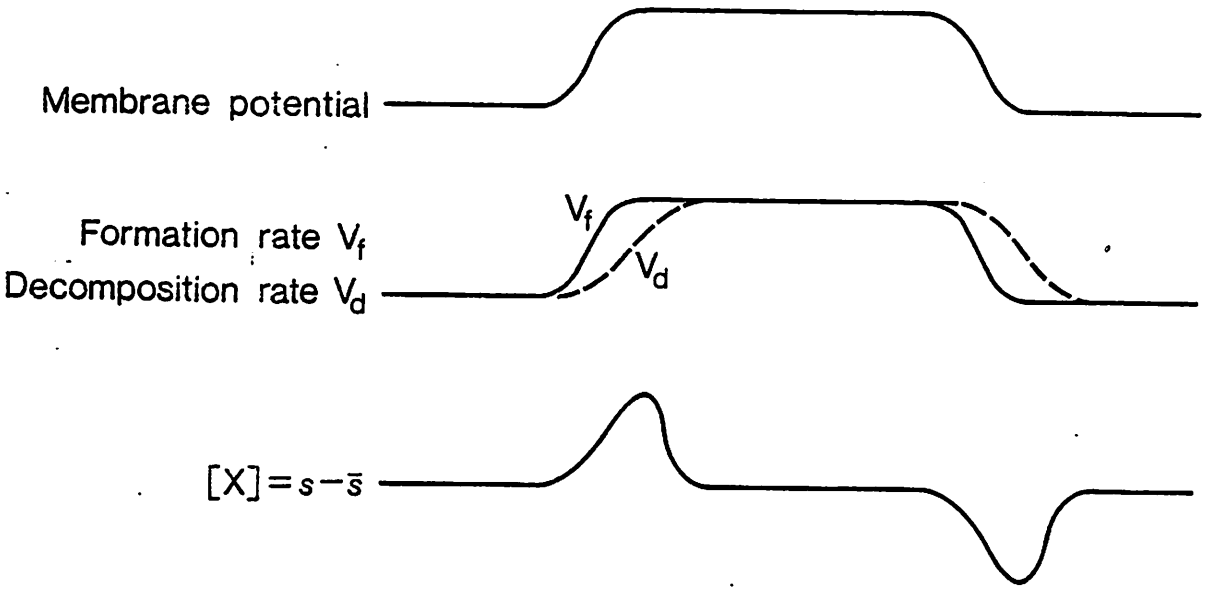


Fig. 9. A hypothetical biochemical mechanism for detecting changes in membrane potential (cf. ref. 19). If the formation rate V_f of substance X increases with depolarization and the decomposition rate V_d increases at a slower rate, then the concentration of X will show a transient increase to a sustained increase in depolarization. In a similar manner the concentration of X will signal the onset of sustained hyperpolarization.