Behavioural Brain Research, 21 (1986) 143–154
Elsevier

BBR 00590

SIMULATION OF THE CLASSICALLY CONDITIONED NICTITATING MEMBRANE RESPONSE BY A NEURON-LIKE ADAPTIVE ELEMENT: RESPONSE TOPOGRAPHY, NEURONAL FIRING, AND INTERSTIMULUS INTERVALS

JOHN W. MOORE, JOHN E. DESMOND, NEIL E. BERTHIER, DIANA E.J. BLAZIS, RICHARD S. SUTTON and ANDREW G. BARTO

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

(Received January 10th, 1986)
(Revised version received May 12th, 1986)
(Accepted May 16th, 1986)

Key words: computational neural model – classical conditioning – nictitating membrane response – rabbit

A neuron-like adaptive element with computational features suitable for classical conditioning, the Sutton-Barto (S-B) model, was extended to simulate real-time aspects of the conditioned nictitating membrane (NM) response. The aspects of concern were response topography, CR-related neuronal firing, and interstimulus interval (ISI) effects for forward-delay and trace conditioning paradigms. The topography of the NM CR has the following features: response latency after CS onset decreases over trials; response amplitude increases gradually within the ISI and attains its maximum coincidentally with the UR. A similar pattern characterizes the firing of some (but not all) neurons in brain regions demonstrated experimentally to be important for NM conditioning. The variant of the S-B model described in this paper consists of a set of parameters and implementation rules based on 10-ms computational time steps. It differs from the original S-B model in a number of ways. The main difference is the assumption that CS inputs to the adaptive element are not instantaneous but are instead shaped by unspecified coding processes so as to produce outputs that conform with the real-time properties of NM conditioning. The model successfully simulates the aforementioned features of NM response topography. It is also capable of simulating appropriate ISI functions, i.e. with maximum conditioning strength with ISIs of 250 ms, for forward-delay and trace paradigms. The original model's successful treatment of multiple-CS phenomena, such as blocking, conditioned inhibition, and higher-order conditioning, are retained by the present model.

INTRODUCTION

Sutton and Barto's (S-B) model of connectionistic learning by a neuron-like adaptive element predicts many facets of classical conditioning, including acquisition, extinction, interstimulus interval (ISI) effects in trace conditioning, blocking, and conditioned inhibition[1,20]. Basic processes of conditioning with which the model might be identified have been discussed in detail in the original articles. Learned connections are portrayed as occurring at synaptic junctions where CSs gain access to the element. The element also receives input from the US. The magnitude of a CR to a CS depends on the value of the corresponding synaptic weight. This report describes simulations of the classically conditioned nictitating membrane (NM) response and correlated neuronal activity based on the S-B model. The NM CR and related eye blink in the rabbit offer an extensive experimental literature[7] for assessing the model's performance. The degree to which the model and its various implementations are capable of describing this body of data provides an index of its neurobiological validity.

Correspondence: J.W. Moore, Department of Psychology, University of Massachusetts, Amherst, MA 01003, U.S.A.

Our primary aim was to develop a variant of the S-B model capable of generating features of the topography of NM CRs as they unfold in real time within trials: CRs anticipate the US, and CR latency decreases progressively during acquisition and increases during extinction; CR amplitude increases progressively during acquisition and decreases during extinction, but peak CR amplitude tends to coincide with the onset of the US. Simulating CR topography necessitated not only constraints on the variables and parameters of the model but some additional assumptions as well. Given these constraints, we required that the model yield appropriate ISI functions, with an optimum at 250 ms and tailing away as ISI increases, while at the same time preserving the model's predictions regarding multiple-CS phenomena, such as blocking, conditioned inhibition, higher-order conditioning, and serial conditioning that were described in an earlier report[1].

Simulations with the S-B model presented in earlier reports yielded appropriate ISI functions for trace conditioning but not for forward-delay conditioning (see ref. 20, p. 147). Furthermore, previous simulations with the model assumed that all computations within a training trial were completed before US offset. This assumption was implemented by specifying a US duration of 30 time steps, each assumed to be 10-ms in duration. Given the assumed rate of decay of the eligibility trace (see below) associated with the CS, the 300-ms US ensured that all computations within a trial were completed prior to US offset.

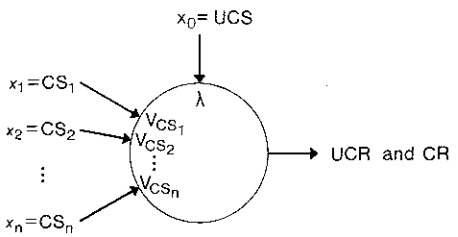

Fig. 1. The original Sutton and Barto neuron-like adaptive element, an analog of classical conditioning. Each input pathway x_i has transmission efficacy V_{CS} corresponding to the associative strength of CS$_i$. The US (labeled UCS) is signalled via a pathway of fixed efficacy λ. Prior to conditioning, the output of the element contributes to the UR (labeled UCR); following conditioning, the element output contributes to both CR and UR (from Barto and Sutton[1]).

Allowing for computations affecting synaptic weights to occur after US offset, as would be the case if the CS's eligibility for modification remained above zero for a significant period of time after US termination, profoundly complicates successful implementation of the model. In this regard it is worth noting that US durations employed in NM conditioning are typically 50 ms in duration or less. Thus, the model presented here assumes that computations occur before, during, and after the US.

ADAPTIVE ELEMENT AND CR RELATED NEURONAL FIRING

With the imposition of certain constraints, the S-B model can simulate CR-related firing patterns of single neurons. Baseline firing rates of single neurons with CR-related activity recorded from the brainstem of awake behaving rabbits is rarely less than 10 Hz, and maximum firing rates during CRs rarely exceeds 100 Hz (ref. 5). With an ISI of 350 ms, spikes are recruited slowly following CS onset. About 150 ms after CS onset, but rarely before 70 ms, spike recruitment increases sharply and may continue to increase throughout the remainder of the ISI. Onset of the US results in a rapid recruitment of spikes to a momentary higher rate that seldom exceeds 200 Hz. This high firing rate associated with the US may persist until US offset, after which firing initiated by the US declines toward baseline.

THE MODEL

Fig. 1 summarizes the S-B adaptive element as applied to classical conditioning. Inputs to the element, denoted x_0 in the case of the US and x_i, i = 1, ..., n, in the case of CSs, multiply with their corresponding synaptic weights to determine the output of the element. These weights are denoted λ for the US and V_i for the respective CSs. The latter weights are modifiable and carry the long-term consequences of training. The element's output or response, denoted s, is simply the weighted sum of its inputs. Despite the theoretical possibility that weights might assume any real value, s is bounded to remain between 0 and 1 in

this implementation. Other departures from the original model are noted below.

Limiting s to the closed unit interval seemed appropriate given our goal of modelling CR topography. Consider, for example, the element's response to a CS with a negative synaptic weight, e.g. a conditioned inhibitor. A negative (rather than 0) output of the element to such a CS implies an action opposite in direction to the CR. Whereas the NM CR consists of eyeball retraction and NM extension (see e.g. ref. 3), the action implied by a negative s is exophthalmus and NM retraction. These CR-opposing responses to conditioned inhibitors are generally not observed in the rabbit NM preparation (see e.g. refs. 15 and 21).

The equations governing learning, i.e. the moment to moment changes in CS synaptic weights, depend on two trace processes. The first, denoted $\bar{x}$, is the so-called eligibility trace. It defines the period and extent to which a CS's weight can be modified. The value of $\bar{x}$ becomes positive at the same time (or shortly after) the CS input, x, reaches the adaptive element. It rises in value gradually with the CS input and decays gradually after the CS is removed. The second trace process, $\bar{s}$, is the trace or memory of the element's output during preceding computational epochs. This variable can be interpreted as the element's prediction or expectation of its output during the current time step. These two trace processes, $\bar{x}$ and $\bar{s}$, have been discussed in detail in the previously published descriptions of the S-B model[1,20].

In order to implement CR topography and ISI effects, the model described below specifies an additional trace process. This is a trace associated with each input to the element, the variable denoted x in the original S-B model, which is distinct from the eligibility trace, $\bar{x}$. Simulations in the original articles describing model behavior assumed that CSs affect the adaptive element with 0 latency and with instantaneous rise and fall times. The present implementation assumes that inputs to the element from CSs are shaped by unspecified mechanisms (preprocessing) in such a way that, instead of reaching the adaptive element instantaneously as a step function, CS inputs have a fixed latency and increase in value gradually to some maximum. In addition to this gradual increase in CS input to the element, the effect of these CS inputs persists for a period after CS offset.

Specifically, the present model assumes that the onset of CS_i causes x_i to increase in value with a latency (justified below) of 70 ms. This increase in the value of x_i continues in an S-shaped fashion until CS offset, at which point it decreases progressively to its baseline value of 0. The value of the corresponding eligibility trace, $\bar{x}_i$, follows the same temporal course, but with a latency relative to x_i of 30 ms. It also decays after CS offset, but at a slower rate than that of x_i. The relative rates of growth and decay of x_i and $\bar{x}_i$ determine how the corresponding synaptic weight, V_i, changes from one time step to the next. As discussed more fully below, these changes in V_i determine CR topography at various stages of training and ISI effects.

Assumptions regarding the US follow the original model in specifying a positive input to the element, λ, related directly to US intensity and with instantaneous onset and rise time. Rather than assuming that the US's input falls instantaneously to 0 at US offset, however, the present model assumes that the US's input to the element decays progressively following US offset. This decaying trace of the US influences post-US computations affecting synaptic weights of CSs. (As previously noted, earlier implementations of the model assumed for simplicity that all computations affecting CS synaptic weights occurred before US offset.) Another important change is that, instead of assuming that λ summates with the weighted sum of inputs from CSs to determine the output, s, the contribution to s of the US in the present implementation is equal to the *difference* between λ and the largest starting weight among CSs presented on that trial. Thus, as CS synaptic weights increase with training, the effectiveness of the US decreases. Although this rule allows for a progressive diminution of the UR as anticipated by other models (see e.g. ref. 6), its main value is in preventing post-US decrements in synaptic weight from canceling the increments accruing from the US.

IMPLEMENTATION

Implementation rules and parameterization of the model were arrived at largely by trial and error and by searching parameter spaces. Computational epochs (time steps) were assumed to represent 10 ms of real time. The S-B model is basically defined by Equations 1 and 2:

Letting t designate a given time step, the synaptic weight of the CS_i, V_i, is modified as in the original model by the rule

$$V_i(t+1) = V_i(t) + c[s(t) - \bar{s}(t)]\bar{x}_i(t), \quad (1)$$

where c is the learning rate parameter such that $0 < c \leq 1$.

Prior to US onset, the output of the element at time t, $s(t)$, is given by

$$s(t) = \sum_{i=1}^{n} V_i(t)x_i(t). \quad (2)$$

As noted, $s(t)$ is confined to the closed unit interval and linear only within this range.

The summation in Eqn. 2 refers only to time steps that include CSs but not the US. When the US occurs, the summation includes an additional term, λ', equal to the difference between λ, which was treated as a parameter between 0 and 1, and the largest positive staring weight among the CSs present at the beginning of the trial. This difference is bounded to the closed unit interval, i.e. $0 \leq \lambda' \leq 1$. Specially, if V_i is the largest starting weight among CSs present on a given trial on which the US occurs,

$$\lambda' = \begin{cases} \lambda - V_i & \text{if } 0 \leq V_i \leq \lambda; \\ 0 & \text{if } V_i > \lambda; \\ \lambda & \text{if } V_i \leq 0. \end{cases}$$

Immediately following US offset, the contribution of λ' to $s(t)$ is replaced by

$$\lambda'(t+1) = 0.9\lambda'(t),$$

which represents a geometrically decaying aftereffect of the US. As in Eqn. 2, at no time was $s(t)$ permitted to exceed 1 or be less than 0; $s(t)$ was linear only within this range.

The element's prediction of its output based on prior activity, $\bar{s}(t)$ in Eqn. 1, is given by

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

where $0 \leq \beta \leq 1$. In the present implementation, $\beta = 0.6$. Other values of β resulted in a variety of predictions that were seriously at odds with NM conditioning data. Therefore, $\beta = 0.6$ might be regarded as constant in the present application of the model.

Equally critical for the model's behavior are the functions adopted for the variables x_i and $\bar{x}_i$. For the initial 7 time steps (i.e. 70 ms) after onset of CS_i, $x_i(t) = 0$. For time steps $t > 7$ until offset of CS_i, $x_i(t)$ is an increasing function of t in the open interval (0,1). In the present simulations,

$$x_i(t) = [\tan^{-1}(0.35t - 5.5) + 90]/180.$$

For time steps following CS_i offset,

$$x_i(t+1) = 0.85x_i(t).$$

These implementation rules for x imply that a CS's effective input to the element begins 70 ms after its onset at the periphery, that it rises in an S-shaped fashion to a maximum of 1, and that it decays geometrically toward 0 following CS offset. This pattern of increase and decline of x determines the form of the CR. The 70 ms delay of CS input to the element was chosen because conditioning of the NM CR in rabbits does not occur with ISIs less than this value[17]. Thus, ISIs of less than 70 ms are ineffective because there is insufficient time for the CS input to influence the element.

The variable $\bar{x}_i(t) = x_i(t-3)$ whenever t is less than or equal to the duration of CS_i (in time steps) plus 3. For t greater than the duration of CS_i plus 3,

$$\bar{x}_i(t+1) = \delta\bar{x}_i(t).$$

The δ factor is given by the expression $\delta = e^{-3/d}$, where d is a variable equal to the duration of the CS in 10-ms time steps, provided that $d \geq 25$. These implementation rules for $\bar{x}$ imply that a

CS's synaptic weight becomes eligible for modification 30 ms after activation of its stimulus trace, x, and that the temporal course of $\bar{x}$ is identical to that of x until 30 ms after CS offset, at which point the decay of $\bar{x}$ depends on CS duration.

These rather complex rules governing the variable rate of decay of CS eligibility traces were dictated by our aim of simulating ISI functions for forward-delay, as well as trace, conditioning. Slowing the rate of decay of $\bar{x}$ for CSs of relatively long duration, as would be encountered in a long-delay conditioning protocol, provides more opportunities for decrementing V during post-US time steps and thereby causes less conditioning than in the case of ISIs nearer the optimal interval. This point is discussed more fully below.

IMPLEMENTATION OF CR TOPOGRAPHY

CR topography is basically determined by the adaptive element's output, s. However, a linear relationship between the amplitude of the CR, as observed at the periphery, and s is unrealistic because of the physical constraints of the system. An NM response occurs passively when the eyeball retracts into the orbit from its resting position, and its amplitude, which depends on the intensity of the eliciting stimulus, has an upper limit[13]. This was one consideration in setting an upper bound of 1 on s for generating NM amplitude at each time step.

The output of the S-B element cannot simulate the progressive decrease in CR latency that occurs over training without additional assumptions. Since s is determined by the input trace from the CS, the variable x in the model, one approach to implementation of decreasing CR latency over training would be to assume that the onset of x, rather than occurring 70 ms after CS onset as described above, begins late in the ISI and somehow migrates over trials toward CS onset. An alternative approach, one that is physiologically easier to justify than the former, is to interpose a threshold between the element's output and motoneurons that generate the peripherially observed response. In the model described here, s had to exceed 0.1 in order to produce a detectable response. Also, in order to smooth the transition from CR to UR, response profiles were generated by the sliding arithmetic mean value of s from the current and two preceding time steps. In sum, NM topography is defined by the arithmetic mean of the element's current output, s, and those of the two preceding time steps. Furthermore, this sliding mean is bounded to remain between 0.1 and 1. This implementation rule tends to smooth response profiles, ensure that NM amplitude is bounded, and yields progressively decreasing CR latency (not less than 70 ms) over training.

IMPLEMENTATION OF NEURONAL FIRING

The S-B model readily lends itself to simulation of neuronal firing on single trials. These can be cumulated over trials into peristimulus–time histograms (PSTHs). Simulated neuronal spikes are derived from the momentary output of the system, s, bounded between by 0 and 1. Because momentary firing rates up to 200 Hz are allowed, the number of spikes that can be assigned to a 10-ms time bin is $k = 0$, 1, or 2. The value of k selected by the implementation rules at time step t was obtained by treating k as a Poisson random variable with parameter $s(t)$. In addition, a random threshold variable, denoted L and obtained from a uniform distribution in the unit interval, governs the value of k selected at time t. Specifically, letting $P[k; s(t)]$ denote the probability of exactly k spikes given the current value of $s(t)$,

$$P[k; s(t)] = \frac{e^{-s(t)}\, s(t)^k}{k!} = P(k,t). \tag{3}$$

The simulation selects L at random, $0 \leq L \leq 1$, and computes the sum over k of $P(k,t)$ for $k = 0,1,2$ until the sum of $P(k,t) \geq L$. If this sum attains L when $k = 0$, the spike count at t is 0; if it attains L when $k = 1$, the spike count at t is 1; if it attains L when $k \geq 2$, the spike count at t is 2.

METHOD

As noted, all simulations assumed that time steps corresponded to 10 ms. Because computations did not involve static contextual cues, it

was sufficient to define a string of computational steps on each trial that was limited to (1) pretrial period, used only for the purpose of generating pretrial neuronal activity with Eqn. 3; (2) an on-trial period that varied in duration depending on the duration of the longest CS and the temporal locus and duration of the US; and (3) a post-trial period of sufficient length to encompass the decaying CS eligibility traces and the after-effects of the US.

CR TOPOGRAPHY

Fig. 2 shows simulated NM CR/UR complexes for single trials of both early (Trial 7) and asymptotic (Trial 30) stages of training. The synaptic weight of the CS, V, was 0 at the outset. Response topography at various stages of training depends on a number of variables and parameters of the model, including the rise and fall of x and $\bar{x}$ associated with CS onsets and offsets. As illustrated in Fig. 2, response topography also depends on the learning rate parameter, c in Eqn. 1, and on the effectiveness of the US, specified by the parameter λ. The paradigm illustrated in Fig. 2 is single-CS training in a forward-delay arrangment with an ISI of 350 ms and a 30 ms US (US onset occurring at CS offset). Various combinations of the parameters c and λ are illustrated. In Fig. 2A, $c = 0.04$ and $\lambda = 0.6$. In Fig. 2B, $c = 0.04$ and $\lambda = 0.9$. In Fig. 2C, $c = 0.40$ and $\lambda = 0.6$. The top row of panels are simulated responses on Trial 7; the middle row shows simulated responses on Trial 30. Associated simulated PSTHs of neuronal firing, discussed below, are shown in the bottom row.

The top and middle rows of Fig. 2 indicate that response topography mimics, albeit ideally, those typically observed in actual experiments (e.g. Fig. 4): peak CR amplitude increasing with training, yet remaining just before the UR, and CR latency decreasing. This process reverses over simulated extinction trials. In addition, UR amplitude became smaller with training, i.e. conditioned diminution of the UR, as reported by some experimenters[6].

Given a near-optimal ISI such as 350 ms, robust CRs during asymptotic stages of training typically blend into the UR without the discontinuity that is evident in the simulated CR/UR complexes in Fig. 2A and B in which $c = 0.04$. They more nearly resemble those in Fig. 2C in which $c = 0.40$. Although response topography

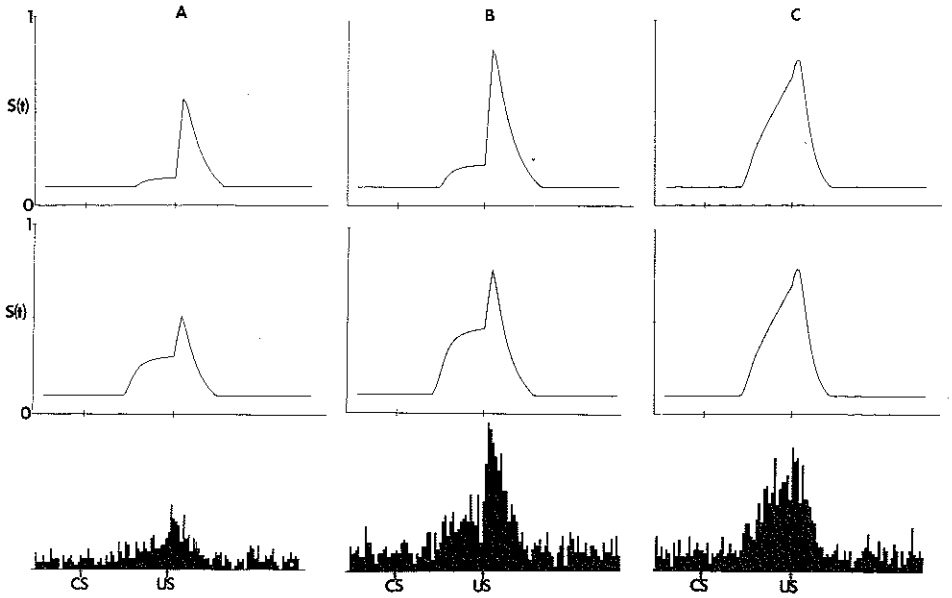

Fig. 2. Simulated CR/UR complexes and PSTHs for single-CS forward-delay training with a 350-ms CS that terminated simultaneously with the onset of a 30-ms US. Top row of panels refer to Trial 7; middle row refers to Trial 30; bottom row are corresponding PSTHs A: $c = 0.04$ and $\gamma = 0.6$. B: $c = 0.04$ and $\lambda = 0.9$. C: $c = 0.4$ and $\lambda = 0.6$.

and rate of learning are greatly affected by the value of c, asymptotic Vs for single-CS training are less strongly affected by this parameter. For example, with $c = 0.04$, the values of V after 30 trials are 0.31 with $\lambda = 0.6$ and 0.47 with $\lambda = 0.9$. With $c = 0.40$, the corresponding values of V after 30 trials are 0.39 and 0.56 for λ values of 0.6 and 0.9, respectively. Unfortunately, values of c that yield realistic topographies, as in Fig. 2C, also cause overly fast acquisition. That is, V increases too quickly over trials to reflect rates of CR acquisition that are typically observed in the laboratory. For example, with $c = 0.40$ and $\lambda = 0.9$, the value of V after only 7 trials is the same as after 30 trials ($V = 0.56$).

There are several approaches to restraining the growth of V while also ensuring realistic response topography at asymptotic stages of training. For example, rather than treating c as a constant, it might be treated as a variable with a low initial value that increases progressively over reinforced trials. Computational models with this feature have been successfully applied to NM conditioning by Moore and Stickney[14]. In the Moore–Stickney model, the rate-of-learning parameter on any trial depends not only on events that occur during and immediately after CS occurrences, but also on the duration of intertrial intervals. Although intertrial interval has a powerful effect on the rate of CR acquisition (see e.g. ref. 18), the present model does not encompass this variable.

With some combinations of parameters, the present model produces non-monotonic acquisition curves. For example, when c is greater than 0.50, λ values less than 0.40 yield acquisition curves whereby V initially increases sharply and then oscillates for a few trials before stabilizing at some value lower than the initial peak. With c values suitably small relative to λ, the problem of non-monotonic acquisition curves does not arise.

One problem regarding CR topography in the present model concerns trace conditioning. As illustrated in Fig. 3, a simulated CR for trace conditioning does not blend into the UR but instead reverts toward baseline within the trace interval. With near-optimal ISIs such as 350 ms, real trace-conditioned CRs tend to increase in amplitude within the trace interval in much the same way as in a forward-delay training paradigm. Premature decline of simulated CR amplitude in trace conditioning is due to the decay of x, which in the present model begins at CS offset. Because x essentially carries CR topography, the problem might be rectified by allowing x to begin to decay at some point in time beyond CS offset. Any change in assumptions about the decay of x would necessitate compensatory changes in rates of decay of $\bar{x}$ and λ'.

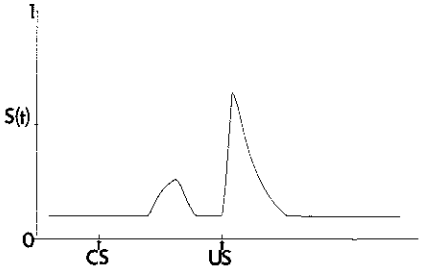

Fig. 3. Simulated CR/UR complex for Trial 48 of trace conditioning with a 250-ms CS followed after 150 ms by a 30-ms US; $c = 0.15$ and $\lambda = 0.9$.

NEURONAL FIRING

The bottom row of Fig. 2 shows simulated PSTHs of neuronal firing cumulated over 30 trials of simulated training as specified in the corresponding panels of the middle row. These PSTHs were generated by the Poisson rule described above in connection with Eqn. 3. Their shape tends to model CR/UR topography. This modelling of CR topography (see Fig. 4A) has been widely reported as characteristic of firing patterns of neurons in a number of brain regions important for NM conditioning: brainstem[5], cerebellum[4,11], and hippocampus[2].

Although CR-modelling neuronal firing similar to that generated by the model is commonplace, other patterns of CR-related neuronal activity in these same brain regions have been reported. For example, CR-related decreases in firing[4,5] yield PSTHs that resemble mirror images of CR topography (Fig. 4B). The model can easily implement such CR-related decreases in firing by interposing a signal inverter between the element's output and the variable used in Eqn. 3 for generating spikes. The challenge, of course, lies in formulating a

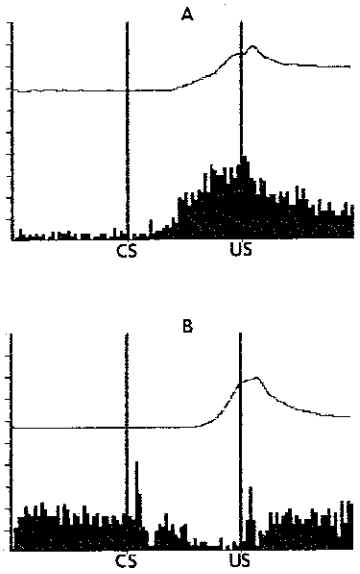

Fig. 4. Average CR/UR complexes and corresponding PSTHs for cells 53 (A) and 17 (B) of Desmond's[5] study of brainstem neuronal activity during NM conditioning in a forward-delay paradigm with an ISI of 350 ms.

model that successfully integrates the contributions of the various types of CR-related firing patterns to conditioned responding and brain processes generally. Such an integration must be deferred pending future insights into the neural basis of NM conditioning.

The simulated PSTHs in Fig. 2 do not reflect the temporal lead over CRs that would be required of a neuron causally related to the behavior. For example, a motoneuron fires several ms before the CR can be detected at the periphery. It is trivially easy to implement this requirement of CR-related neurons by introducing a delay between the generation of neuronal spikes by the variable s and the generation of CR topography.

INTERSTIMULUS INTERVAL FUNCTIONS

Simulated ISI functions appear in Table I. The entries are synaptic weights resulting from 50 simulated reinforced trials ($c = 0.15$) with ISIs ranging from 0.1 to 2 s in forward-delay paradigms and 0.3 to 2 s in trace paradigms. In forward-delay paradigms, the CS terminated concurrently with US onset. In trace paradigms, CS duration was 250 ms for all ISIs, and therefore Table I does not show entries for ISIs of this duration or less. The US was assumed to be 30 ms in duration with λ values of 0.5, 0.7, or 0.9. Appropriately for the rabbit NM response, the maximum weights under forward-delay training were obtained with ISIs of 250 ms. The effect of increasing the effectiveness of the US was to raise and broaden the ISI functions obtained with forward-delay training. Also appropriate for the NM response, forward-delay training yielded higher weights generally than did trace conditioning. With $c = 0.04$ and $\lambda = 0.6$, parameters illustrated in Fig. 2, weights after 50 trials in delay conditioning at ISIs of 100, 250, and 2000 ms were -0.02, 0.33, and 0.07, respectively.

Table I reveals deficiencies in the model's predictions for less-than-optimal ISIs, namely,

TABLE I

Synaptic weights (V) as a function of ISI

The table gives values of V after 50 trials for various combinations of ISI and λ for forward-delay and trace conditioning with $c = 0.15$. Vs for trace conditioning are in parentheses.

	ISI (ms)											
	100	*150*	*200*	*250*	*300*	*350*	*400*	*500*	*700*	*1000*	*1500*	*2000*
$\lambda = 0.5$	−0.07	0.23	0.31	0.33	0.31 (0.30)	0.29 (0.22)	0.28 (0.15)	0.25 (0.06)	0.21 (0.01)	0.17 (0.00)	0.13 (0.00)	0.10 (0.00)
$\lambda = 0.7$	−0.09	0.33	0.43	0.46	0.43 (0.42)	0.41 (0.31)	0.39 (0.21)	0.35 (0.08)	0.29 (0.01)	0.23 (0.00)	0.18 (0.00)	0.14 (0.00)
$\lambda = 0.9$	−0.12	0.42	0.55	0.59	0.56 (0.54)	0.53 (0.40)	0.50 (0.27)	0.45 (0.10)	0.37 (0.01)	0.30 (0.00)	0.23 (0.00)	0.18 (0.00)

the negative weights obtained with ISIs of 100 ms. Although negligible when compared with the positive weights obtained with near-optimal ISIs, there is no experimental evidence to suggest that less-than-optimal ISIs in forward-delay conditioning results in inhibitory learning[17]. The negative weights are due to complex interactions among the rise and fall of the variables x and $\bar{x}$. This issue is discussed more fully below.

Although the present model does a reasonable job of simulating ISI functions, in its present form it cannot account for effects of intertrial interval on the level of conditioning that is possible for a given ISI. In rabbit NM conditioning, for example, high levels of conditioned responding have been obtained in a forward-delay paradigm with ISIs longer than 2 s with intertrial intervals of 24 h or more[9]. Similar interactions between ISI and intertrial interval have been reported in pigeon autoshaping with trace conditioning[8].

MULTIPLE-CS PHENOMENA

Barto and Sutton[1] showed that several multiple-CS phenomena that have attracted the attention of animal learning theorists, e.g. blocking, conditioned inhibition, and higher-order conditioning, are produced by their model. We have confirmed that these predictions are retained in the present model. For example, Fig. 5 summarizes a simulation of conditioned inhibition training similar to that presented in the earlier report (see ref. 1, p. 229): CS_1 was reinforced with the US, and a compound consisting of CS_1 and CS_2 was not reinforced. The two trial types, designated CS + and CS −, respectively, occurred equally often but in an unsystematic sequence. Fig. 5 shows changes of synaptic weights over the course of training. Notice that the weight for CS_2, the conditioned inhibitor, became increasingly negative in value.

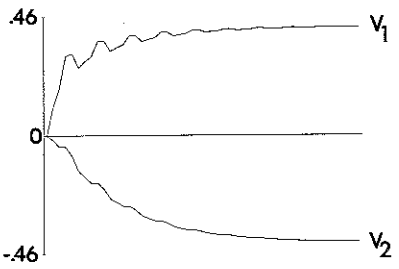

Fig. 5. Synaptic weights (V_1 and V_2) for CS_1 and CS_2 of a simulated conditioned inhibition protocol as a function of trials: 50 trials consisting of CS_1 paired with a 30-ms US in a forward-delay arrangement, with an ISI of 350 ms, were intermixed in a random sequence with 50 trials consisting of CS_1 and CS_2 presented together for 350 ms, but without the US; $c = 0.15$ and $\lambda = 0.7$. Notice that V_1 becomes progressively more positive while V_2 becomes progressively more negative.

Although the present model provides a credible treatment of serial-compound conditioning, simulation studies also indicate that there are a number of subtle effects associated with the within-trial timing of onsets and offsets of CSs. A comprehensive treatment of these paradigms will be the focus of a future report.

DISCUSSION

We have noted a number of differences between the present model, which was designed specifically for NM conditioning, and the original S-B model.

In the original model (see ref. 1, p. 226), x is a step function equal to 0 in the absence of the CS or 1 in its presence. Consequently, s is also a step function, taking the value of V at CS onset and changing only during US onset or offset; V can change only at transition points because it is at these times when the quantity $s - \bar{s}$ in Eqn. 1 takes on non-zero values. In addition, changes of V can only occur during periods of eligibility, i.e. when $\bar{x}$ exceeds 0. Thus, for single-CS delay conditioning, increases in V occur only at US onset, and decreases in V occur only at CS offset. Asymptotic learning occurs when V approaches λ because, at US onset, $s - \bar{s}$ in Eqn. 1 is nearly 0, assuming that β equals 0 as in the the original implementation. During extinction, and non-reinforced trials generally, $s - \bar{s}$ is negative at CS offset, and changes in V are therefore negative.

In the present model, instead of being a step function, x increases monotonically as an S-shaped function of time when the CS goes on and decreases geometrically when the CS goes off. These changes imply that computations affecting V occur in every time step leading up to the US and for a period afterwards depending on

the eligibility trace. When the CS is on, $s-\bar{s}$ is positive throughout the rising portion of the x function. Consequently, V increases from moment to moment within this interval, the contribution of the $s - \bar{s}$ term increasing as eligibility increases. After CS and US offset, $s - \bar{s}$ is negative in every time step until x decays to 0. This results in a series of decrements in V, these decrements growing progressively smaller as the eligibility trace decays. In acquisition, asymptote is achieved when the post-US decrements in V all but equal the increment in V occurring earlier in the trial. To ensure extinction in the absence of the US, the net change in V for a CS with positive value must be negative.

Because in the present model x increases gradually in the ISI and does not approach its asymptotic value until 250 ms after CS onset, it was necessary to introduce a mechanism to prevent asymptotic Vs in simple forward-delay paradigms with ISIs between 70 and 250 ms from becoming too large. Since x is less than 1 at US onset with the shorter ISIs, the term $s - \bar{s}$ in Eqn. 1 at US onset is greater than it would be with a CS of long duration. This causes large net increments in V. To mitigate this problem, we introduced a lag of 30 ms in the eligibility trace, $\bar{x}$, with respect to x. This lag compensates for the large increments of V with ISIs less than 250 ms in two ways. First, the lag causes fewer increments in V during pre-US time steps. Second, with the lag, eligibility remains high during post-US time steps, causing post-US decrements in V that are larger than they would be otherwise. Although lagging $\bar{x}$ made it possible for the model to yield ISI functions with a maximum at 250 ms, it contributed to the inappropriate negative weights noted in Table I for ISIs of 100 ms.

The present model is able to predict a fall-off of Vs with ISIs greater than the 250-ms optimum by assuming that the rate of decay of the eligibility trace, $\bar{x}$, varies inversely with CS duration. In forward-delay paradigms, CS duration and ISI are typically equal. During post-US time steps, when $s - \bar{s}$ is negative, a slowly decaying $\bar{x}$, associated with a long ISI, leads to a greater accumulation of decrements of V than in the case of a more rapidly decaying $\bar{x}$, associated with a shorter ISI. In forward-delay paradigms with long ISIs, this device slows CR acquisition over trials and causes a lower asymptotic V than in the case of shorter ISIs. It also results in comparatively faster extinction with long-duration CSs. However, in order to retain the ability to restrain V with less-than-optimal ISIs, we stipulated that the rate of decay of $\bar{x}$ for CS durations (and hence ISIs) less than 250 ms could not exceed that of a 250-ms CS.

When applied to trace conditioning, the assumption that the rate of decay of $\bar{x}$ varies inversely with CS duration implies that, for a sufficiently long constant trace interval (i.e. the time between CS offset and the US), asymptotic V is an increasing function of CS duration. For example, after 50 simulated trace conditioning trials with a trace interval of 300 ms ($\lambda = 0.9$; US duration = 30 ms; $c = 0.15$), V with a CS duration of 250 ms is 0.08, whereas V with a CS duration of 1000 ms is 0.17. The lower V with the 250-ms CS seems paradoxical because the nominal ISI of 550 ms (250 plus 300) would normally be expected to yield stronger conditioning than would the 1000-ms CS with its nominal ISI of 1300 ms. There is no paradox under forward-delay conditioning with these parameters, i.e. when the duration of the trace interval equals 0 ms. The corresponding Vs for ISIs of 550 and 1300 ms are 0.43 and 0.25, respectively. We know of no directly relevant NM conditioning data on whether trace-CS duration affects conditioning in the way predicted by the present model. The situation is further complicated by stimulus coding processes not encompassed in the model, whereby stimuli are defined by the offset rather than onset of some energy source. For example, a tone that occupies most of the intertrial interval, but which terminates 250 ms before the US, yields levels of NM conditioning (to tone offset) that scarcely differ from those obtained by the more usual procedure of having the tone off during intertrial intervals and on for the 250 ms preceding the US[10].

Another difficulty with the present model regarding trace conditioning is that CS durations of 50 ms can support NM conditioning at optimal ISIs[19]. Because x does not begin to increase

before 70 ms after CS offset, no learning is possible to brief, pulse-like CSs. This problem might be corrected by permitting x to increase for a period beyond CS offset, as suggested above in connection with Fig. 3.

The original S-B model is closely related to the Rescorla–Wagner (R-W) model[16]. The present model bears a less obvious relation to the R-W model and differs in at least two important ways. First, unlike the R-W model, the asymptotic value of V for single-CS forward-delay training can never match the value of λ, the principal parameter specifying US effectiveness. The two variables cannot be equal because of the post-US decrements in V that drain off some of the increments cumulated during earlier time steps. For example, with an ISI of 250 ms and the parameters described in the preceding paragraph, V after 50 simulated trials in the present model is 0.59, roughly 2/3 the value of λ. The same ratio holds for other λ values at this ISI. With longer ISIs the ratio declines further. With an ISI of 500 ms, for example, asymptotic V is roughly 0.5 λ.

The present model also differs from the R-W model by *not* predicting extinction of negative V over a series of simulated CS-alone trials. The prediction of such extinction of conditioned inhibition has been a major sticking point in the R-W model because it has been repeatedly demonstrated that conditioned inhibitors do not lose their effectiveness by being presented without the US (see e.g. ref. 12). The present model does not predict 'extinction of conditioned inhibition' because the element's output, s, cannot be less than 0. Therefore, from Eqn. 1, net increments in V in the absence of the US (or some other excitatory input) are not possible. The model can therefore be viewed as an advance on the R-W model because it nevertheless describes substantiated features of conditioned inhibition such as the acquisition of negative V shown in Fig. 5.

In sum, the present model departs from the original S-B model in a number of non-trivial ways. Not only was it necessary to impose special assumptions about the rise and decay of eligibility traces so as to yield reasonably appropriate ISI functions, it was also necessary to restructure the input of the US to the adaptive element. Instead of being constant over trials, the effective US in the present model is the difference between λ, a constant, and the largest positive starting V among CSs present on that trial. This rule was imposed so that post-US computations, which generally produce decrements of V, do not completely cancel the increments in V that generally precede US offset. Departures such as this from the original S-B model increase the computational burden to be born by a single neuron.

The S-B model is, in principle, consistent with any number of neuronal schemas of sensory-motor integration and learning-schemas ranging in complexity from single classic neurons with many input channels and a single output channel, as in the original S-B model, to those composed of multi-layered networks of heterogeneous computational elements that perform specialized functions. The basic computations of the S-B model occur at the level of synapses of single neurons, neurons presumably lying at the confluence of the sensory-motor pathways mediating the target behavior. The fact that such strategically placed neurons with CR-related activity have been observed in brain regions thought to be essential for NM conditioning lends credence to a single-neuron schema, e.g. cerebellar Purkinje cells of hemispheric lobule VI (ref. 4) and pontine reticular formation[5].

As noted above, however, single neurons with firing patterns related to NM conditioned responding have been reported for several brain regions. Moreover, CR-related firing patterns are not homogeneous. Although most CR-related firing patterns are positively correlated with the behavioral CR (see Fig. 4A), other types of CR-related activity has also been reported (see e.g. Fig. 4B). The fact that CR-related neurons exist in many brain regions and produce a variety of CR-related firing patterns introduces complexities that seem inconsistent with a single-neuron schema, and indeed this may well prove to be the case. Nevertheless, there is at present no compelling evidence that more than one type of neuron and firing pattern are *essential* for learning and performing the NM CR. The evident diversity of CR-related firing patterns and wide anatomical distribution of CR-related neurons might be

attributed to their role in any number of functions that, although important and interesting in their own right, are not directly related to the long-term modification of connection weights. These functions include stimulus coding, response shaping, and conditioning of concomitant responses, to mention but a few. CR-related neuronal activity might also be associated with various compensatory mechanisms not primarily related to learning.

In light of the encouraging preliminary success of the present model in capturing detailed features of the NM conditioning, it seems premature to reject the single-neuron representation of the S-B model in favor of one of greater complexity. Whether the S-B model can be applied with equal success to other varieties of behavioral conditioning is an open question. Also unresolved is whether the structure of the model and the particular constraints imposed by NM conditioning truly have implications for understanding physiological mechanisms of learning and memory. We believe this to be a distinct possibility.

ACKNOWLEDGEMENT

This research was supported by Grants AFOSR 830215, NSF BNS 8317920, and USPHS 1 F31 MH08951.

REFERENCES

1 Barto, A.G. and Sutton, R.S., Simulation of anticipatory reponses in classical conditioning by a neuron-like adaptive element, *Behav. Brain Res.*, 4 (1982) 221–235.
2 Berger, T.W., Alger, B. and Thompson, R.F., Neuronal substrates of classical conditioning in the hippocampus, *Science*, 192 (1976) 483–485.
3 Berthier, N.E., The role of the extraocular muscles in the rabbit nictitating membrane response: a re-examination, *Behav. Brain Res.*, 14 (1984) 81–84.
4 Berthier, N.E. and Moore, J.W., Cerebellar Purkinje cell activity related to the classically conditioned nictitating membrane response, *Exp. Brain Res.*, in press.
5 Desmond, J.E., *The Classically Conditioned Nictitating Membrane Response: Analysis of Learning-Related Single Neurons of the Brainstem*, Ph. D. dissertation, University of Massachusetts, 1985.
6 Donegan, N.H., Priming-produced facilitation or diminution of responding to a pavlovian unconditioned stimulus, *J. Exp. Psychol.: Anim. Behav. Proc.*, 7 (1981) 295–312.
7 Gormezano, I., Kehoe, E.J. and Marshall, B.S., Twenty years of classical conditioning with the rabbit. In J.M. Sprague and A.N. Epstein (Eds.), *Prog. Psychobiol. Physiol. Psychol.*, 10 (1983) 197–275.
8 Kaplan, P.S., Explaining the effects of relative time in trace conditioning: a preliminary test of the comparator hypothesis, *Anim. Learn. Behav.*, 13 (1985) 133–238.
9 Levinthal, C.F., Tartell, R.H., Margolin C.M. and Fishman, H., The CS-US interval (ISI) function in rabbit nictitating membrane response conditioning with very long intertrial intervals, *Anim. Learn. Behav.*, 13 (1985) 228–232.
10 Liu, S.S. and Moore, J.W., Auditory differential conditioning of the rabbit nictitating membrane response: IV. Training based on stimulus offset and the effect of an intertrial tone, *Psychon. Sci.*, 15 (1969) 128–129.
11 McCormick, D.A. and Thompson, R.F., Cerebellum: essential involvement in the classically conditioned eyelid response, *Science*, 223 (1983) 296–299.
12 Miller, R.R. and Spear, N.E. (Eds.), *Information Processing in Animals: Conditioned Inhibition*, Lawrence Erlbaum Associates, Hillsdale, NJ, 1985.
13 Moore, J.W. and Desmond, J.E., Latency of the nictitating membrane response to periocular electrostimulation in unanesthetized rabbits, *Physiol. Behav.*, 28 (1982) 1041–1046.
14 Moore, J.W. and Stickney, J.W., Formation of attentional-associative networks in real time: role of the hippocampus and implications for conditioning, *Physiol. Psychol.*, 8 (1980) 207–217.
15 Moore, J.W., Yeo, C.H., Oakley, D.A. and Russell, I.S., Conditioned inhibition of the nictitating membrane response in decorticate rabbits, *Behav. Brain Res.*, 1 (1980) 397–409.
16 Rescorla, R.A. and Wagner, A.R., A theory of pavlovian conditioning: variations in 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.
17 Salafia, W.R., Lambert, R.W., Host, K.C., Chiaia, N.L. and Ramierz, J.J., Rabbit nictitating membrane conditioning: lower limit of effective interstimulus interval, *Anim. Learn. Behav.*, 8 (1980) 85–91.
18 Salafia, W.R., Mis, F.W., Terry, W.S., Bartosiak, R.S. and Daston, A.P., Conditioning of the nictitating membrane response of the rabbit (*Oryctolagus cuniculus*) as a function of length and degree of variation of intertrial interval, *Anim. Learn. Behav.*, 1 (1973) 109–115.
19 Smith, M.C., Coleman, S.R. and Gormezano, I., Classical conditioning of the rabbit's nictitating membrane response at backward, simultaneous, and forward cs-us intervals, *J. Comp. Physiol. Psychol.*, 69 (1969) 226–231.
20 Sutton, R.S. and Barto, A.G., Toward a modern theory of adaptive networks: expectation and prediction, *Psychol. Rev.*, 88 (1981) 135–170.
21 Yeo, C.H., Hardiman, M.J., Moore, J.W. and Russell, I.S., Retention of conditioned inhibition of the nictitating membrane response in decorticate rabbits, *Behav. Brain. Res.* 10 (1983) 383–392.