CONNECTIONISTIC LEARNING IN REAL TIME: SUTTON-BARTO ADAPTIVE ELEMENT AND CLASSICAL CONDITIONING OF THE NICTITATING MEMBRANE RESPONSE

J.W. MOORE, J.E. DESMOND, N.E. BERTHIER, D.E.J. BLAZIS, R.S. SUTTON, AND A.G. BARTO

Departments of Psychology and Computer and Information Science, University of Massachusetts-Amherst

Sutton and Barto's model of connectionistic learning by a neuron-like adaptive element predicts many aspects of classical conditioning, including acquisition, extinction, interstimulus interval effects in trace conditioning, overshadowing, Kamin blocking, and conditioned inhibition (Barto & Sutton, 1982; Sutton & Barto, 1981). The equations governing the learning process, i.e., the moment to moment changes in the "synaptic weights" associated with CSs, depend on two memory processes. The first is a local decaying trace representing the after-effect of input to the element from a CS. The second is the trace or memory of the element's output during the preceding computational epoch. The element's output or response is simply the weighted sum of its inputs, and the equation for modifying weights is basically equivalent to the Rescorla-Wagner model of associative learning. The output trace is interpreted as the element's prediction of its behavior during the current computational epoch.

This report outlines an extension of the Sutton-Barto model to the classically conditioned nictitating membrane (NM) response of the rabbit, a widely adopted "model system" for theoretical and neurobiological studies of learning and memory. The rabbit NM and related eye blink response offer an extensive experimental literature for assessing the performance of the model (Gormezano, Kehoe, & Marshall, 1983). Our approach might be applied to other instances of classical conditioning, but the mathematical details, such as constraints on functions, constants, and parameters, would no doubt vary from one case to the next.

Our studies of the properties of the Sutton-Barto model are concerned with the development of variants of the basic model capable of generating the form or topography of conditioned responses (CRs) as they unfold in real time within trials. Generating CR topography necessitates the imposition of constraints on the variables of the model. Given constraints that model CR topography for comparatively simple protocols, such as acquisition with a single CS in a forward delay paradigm, the question becomes whether the same constraints hold for more complex protocols, such as serial compounds conditioning and conditioned inhibition, in which two CSs are involved. In short, how valid are the model's predictions over a range of training paradigms?

To appreciate why efforts to model CR topography are interesting from the perspective of AI and neuroscience, consider the following rationale: One begins with the

hypothesis that a single adaptive element, which is basically a linear device with multiple input lines and a single graded output, is capable of modeling all of the characteristics of the system under consideration, in this case the NM CR. The next step is to select specific mathematical expressions relating variables and parameters of the model to each other and to time. This step has many earmarks of curve fitting but with the important difference that selected functions must adhere to the broader constraints of the model. The third step is to determine the extent to which the detailed model can describe the expermental literature surrounding the system under scrutiny and make novel predictions. Failures of the model guide subsequent theoretical development. In the domain of connectionistic learning of interest to the authors, these theoretical developments would entail networks of adaptive elements. We believe the problems associated with the development of a comprehensive theoretical rendering of the NM CR may point the way toward a better understanding of how connectionistic learning might be linked to problems of adaptive sensory-motor control encountered by living organisms and intelligent machines.

Space limitations preclude a description of the mathematical details of the Sutton-Barto model that produced the simulations presented in this paper. The constraints imposed on the model were designed to reflect the activity of single neurons of the brain stem with firing patterns related to the NM CR in a forward delay conditioning paradigm and with an interstimulus interval of 350 msec between the onset of the CS and the onset of the UCS. The following are some of the considerations that shaped the contraints imposed on the Sutton-Barto model:

1. CR related neurons in this paradigm rarely have baseline firing rates lower than 10-Hz. Their maximum firing rate rarely exceeds 130-Hz.

2. Following the onset of a CS, spikes are recruited at a slow rate. About 150 msec after CS onset, spike recruitment increases sharply and continues to increase in a negatively accelerated fashion throughout the remainder of the interstimulus interval.

3. Onset of the UCS results in a rapid recruitment of spikes to a rate near the maximum of 130-Hz. This high firing rate persists until UCS offset, after which firing initiated by the UCS declines geometrically toward baseline. The contribution of the UCS to the neuron's firing diminishes with CR acquisition because the conditioning process anticipates the UCS and therefore produces firing rates near the maximum. For this reason, post-UCS firing tends to decline with learning, an observation that might be related to "conditioned diminution of the UCR".

4. Other constraints concerned the changes in the topography of the NM CR over the course of acquisition and extinction trials: (a) CR latency decreases progressively during acquisition and increases during extinction. (b) CR amplitude increases progressively during acquisition and

decreases during extinction. Nevertheless, peak CR amplitude tends to coincide with the onset of the UCS.

5. The final constraint concerns the effect of interstimulus interval on the efficacy of conditioning. The model was required to yield progressively poorer conditioning under a delay conditioning paradigm with interstimulus intervals greater than the recognised optimal interval for the preparation of 250-350 msec. The unconstrained Sutton-Barto model yields an appropriate interstimulus interval function for trace conditioning but not for delay conditioning.

Modifications of the Sutton-Barto model that satisfy the above constraints were arrived at largely by trial and error. The fully implemented model successfully described CR topography and its changes with training. It also lends itself to simulation of firing patterns of single neurons with activity related to the CR.

Figure 1A shows simulated NM CR/UCRs (vertical axis) as a function of acquisition trials (oblique axis). Notice that the latency of the CR decreases over training. Peak CR amplitude increases over trials, yet remains just before the UCR. This process is reversed during simulated extinction trials (Fig. 1B).

Figure 2 summarizes a simulation of conditioned inhibition training. CS1 is reinforced with the UCS, and a compound consisting of CS1 and CS2 is not reinforced. The two trial types, designated CS+ and CS-, respectively, were alternated in simulated training. Figure 2A shows simulated CR topographies for the two trial types at the end of training. Figure 2B shows changes of "synaptic weights" over the course of training. Notice that the weight for CS2 becomes increasingly negative in value. Figures 2C and 2D show simulated cumulative peristimulus-time histograms of neuronal firing for the two trial types.

Simulation experiments also indicate that this variant of the Sutton-Barto model does a credible job of simulating serial-compound conditioning, Kamin blocking and overshadowing, and higher-order conditioning. In addition, simulations have revealed a variety of subtle and largely untested effects on conditioned responding associated with the within-trial timing of onsets and offsets of CSs in serial compound paradigms.

Whether the Sutton-Barto model can be constrained to model other varieties of behavioral conditioning with the same success that it has in the case of the NM CR remains an open question. Also unresolved is whether the structure of the model, and the particular constraints imposed for the NM CR, truly have implications for understanding physiological meachanisms 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

Barto, A.G., & Sutton, R.S. (1982) Simulation of anticipatory reponses in classical conditioning by a neuron-like adaptive element. Behavioural Brain Research , 4, 221-235.

Gormezano, I., Kehoe, E.J., & Marshall, B.S. (1983) Twenty years of classical conditioning with the rabbit. Sprague, J.M. & Epstein, A.N. (Eds.) Progress in Psychobiology and Physiological Psychology , 10, 197-275.

Sutton, R.S., & Barto, A.G. (1981) Toward a modern theory of adaptive networks: Expectation and anticipation. Psychological Review , 88, 135-170.

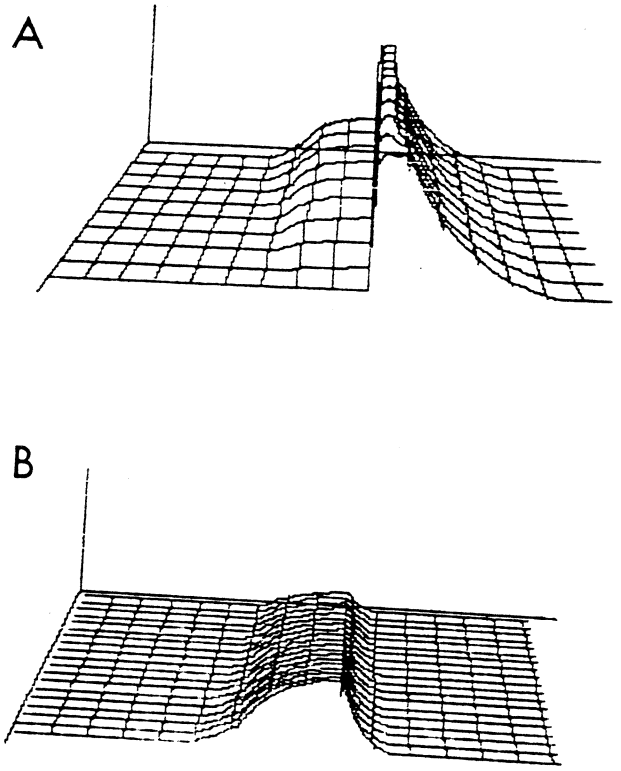

FIGURE 1

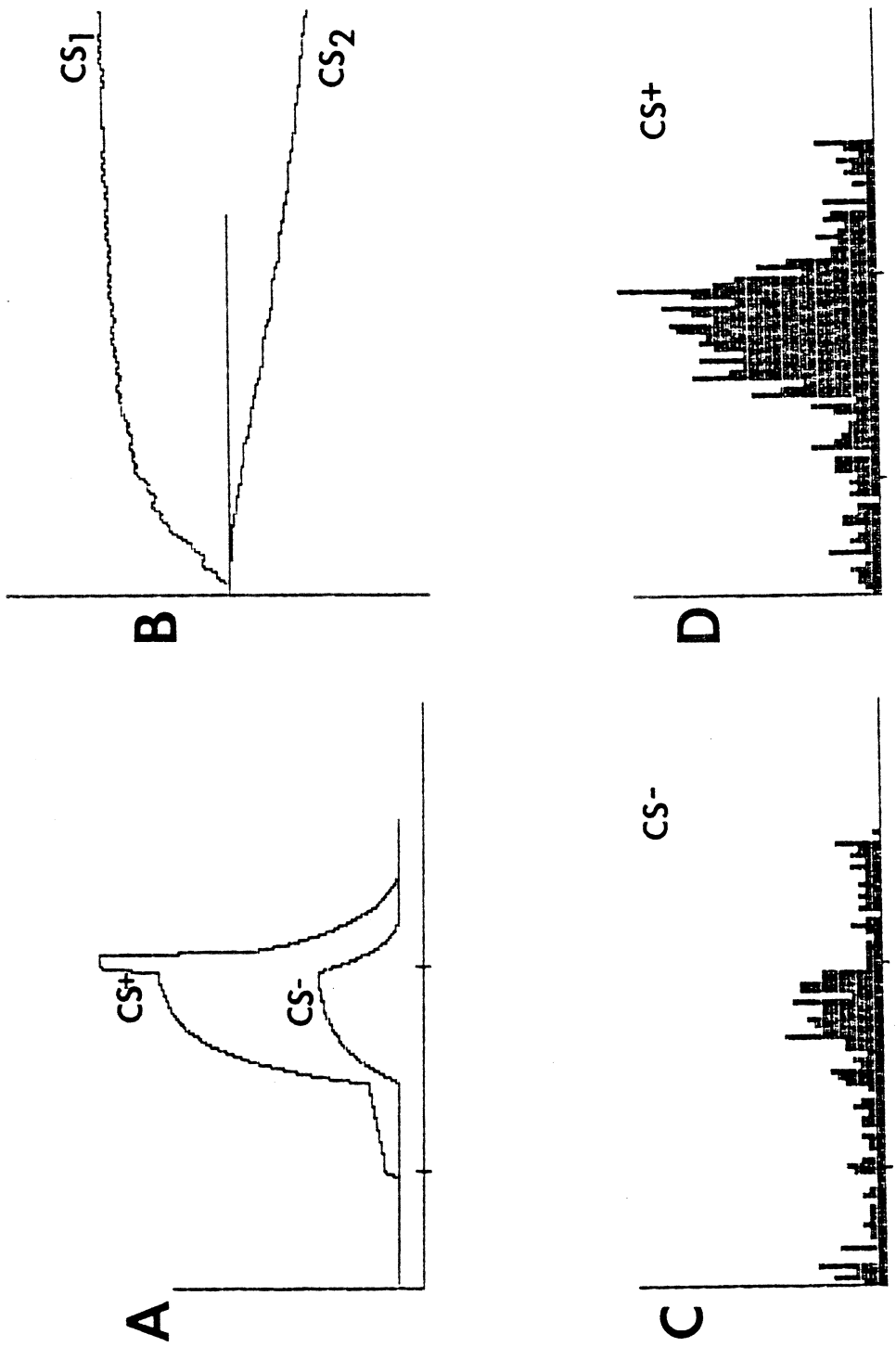

FIGURE 2