

# A Simple recurrent network model of serial conditioning: Implications for temporal event representation

Michael E. Young  
University of Minnesota  
317 Elliott Hall  
Minneapolis, MN 55455  
young@turtle.psych.umn.edu

## Abstract

Elman (1990) proposed a connectionist architecture for the representation of temporal relationships. This approach is applied to the modeling of serial conditioning. Elman's basic simple recurrent network (SRN) was modified to focus its attention on the prediction of important events (Unconditioned Stimuli, or USs) by limiting the connection weights for other events (the Conditioned Stimuli, or CSs). With this modification, the model exhibited blocking and serial conditioning to sequential stimulus compounds. An exploration of the underlying mechanisms suggests that event terminations (CS offsets) were used in predicting US occurrences following simple trace conditioning and event beginnings (CS onsets) were more important following serial conditioning. The results held true under a series of learning rate and momentum values.

## Introduction

The study of classical conditioning, beginning with Pavlov's demonstration of dogs salivating to bells in the early 1900s, is perhaps the most mature area of contemporary psychology. The Rescorla-Wagner model (Rescorla & Wagner, 1972) is one of the best known attempts to explain and predict classically conditioned behavior. Their model focused on the behavior of subjects at the *trial level*, which differs from recent models (e.g. Desmond, 1990; Grossberg & Levine, 1987; Grossberg & Schmajuk, 1987; Lee, 1991; Sutton & Barto, 1981, 1990) where the focus is on *intratrial* stimulus relationships. To demonstrate the distinction, consider the case of *blocking*. Here we have two concurrent events (CS<sub>1</sub> and CS<sub>2</sub>) that consistently precede a third event (the US). However, the subject has received

---

This research supported by the Center for Research in Learning, Perception and Cognition and the National Institute of Child Health and Human Development (HD-07151).

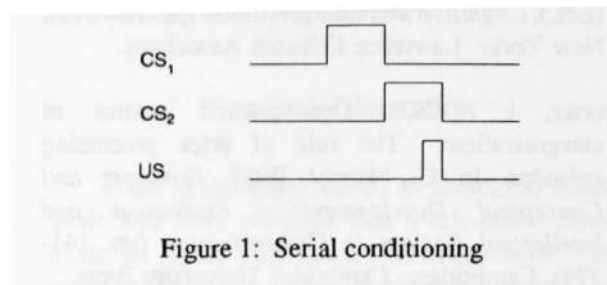


Figure 1: Serial conditioning

prior training in which CS<sub>1</sub> consistently preceded the US. This pretraining retards or *blocks* the subsequent learning of the CS<sub>2</sub>-US relationship. Trial level theories can account for this observation, but they are silent on the role of intratrial variables like inter-stimulus interval (ISI - the temporal distance between CS onset and US onset), the temporal relationships among multiple CSs, and the duration of the CSs and USs. Intratrial models like the present one are designed to address these variables.

The particular classical conditioning paradigm investigated here is that of *serial* or *sequential* conditioning. In serial conditioning, multiple CSs precede the US, but unlike simple compound conditioning where the CSs co-occur, serial CSs are sequentially ordered (see Figure 1) where the first CS is presented in a *trace* relationship with the US. *Trace conditioning* is the term used to describe a situation where the CS is no longer present at the time of occurrence of the US. The second CS serves to provide a mechanism by which the first event acts on the US. In fact, earlier work (e.g. Bolles, Collier, Bouton & Marlin, 1978; Kehoe, Gibbs, Garcia & Gormezano, 1979) has shown that an intervening CS facilitates learned responding to a CS presented in a trace relationship to the US. In essence, it "bridges the gap."

Neural modelers have used a variety of approaches to capture temporal relationships among inputs. The approach adopted in this article is based on the work of Elman (1990). Elman used a simple recurrent network (SRN) to encode time where inputs can have sustained effect via a recurrent delay loop (see Figure 2).

The activations of the hidden units are fed back as input to themselves during the following time step. This permits the model to have some memory of its previous state. Elman trained this type of network to predict the input on the next time cycle: the values presented at the output layer at time  $t$  are identical to the inputs at time  $t+1$ . The primary goal of this study is to evaluate the promise of SRNs for modeling intratrial relationships. The secondary goal is to determine the mechanisms underlying the encoding of these temporal relationships.

## The Model

Earlier work with SRNs demonstrated their ability to model basic excitatory conditioning and phenomena such as blocking (Young, unpublished data). To show blocking it was necessary to treat the two classes of events, the CSs and the USs, differently. Historically, there have been two theoretical approaches to capturing this difference and the relationship between the CSs and the US. Mackintosh (1975) assumes that concurrent CSs must compete for attention. Blocking occurs because the subject learns to attend to the pretrained CS<sub>1</sub> thus interfering with later attending to CS<sub>2</sub>. Alternatively, Rescorla & Wagner (1972) assume a competition among the CSs for US associative strength. Blocking occurs because CS<sub>1</sub> has captured most of the associative strength available from the US. The US is no longer a surprising event (being predicted by CS<sub>1</sub>) and thus does not require any additional predictors.

The software (*tlearn*) used in these simulations was developed at the University of California - San Diego's Center for Research in Language. Tlearn provided a mechanism to encourage CS competition. Figure 2 illustrates the architecture used. The weighted connections between the internal representations (the hidden layer) and the non-US portion of the output layer were limited in value. The limits were chosen as a result of the earlier work with blocking and represent one of the free parameters within this model. These limits constrain the degree to which errors in CS prediction can affect the learning process.

As the weights increase, hidden nodes will be more sensitive to US errors than to CS errors. If the US is being adequately predicted, less error will be propagated back to the internal connections. This process is analogous to the competition among CSs for US associative strength. When any error in US prediction is reintroduced (e.g. a change in salience), the model will be sensitive to these changes, thus allowing the CSs to compete for the ability to predict the "new" US.

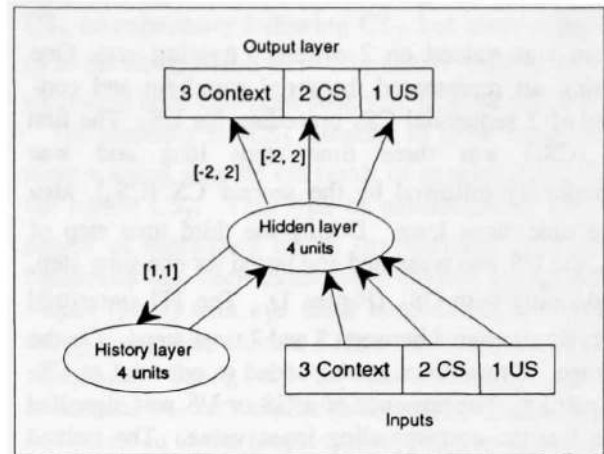


Figure 2: The SRN architecture used in the simulations. Labeled arrows have a limit on connection weights: [lower limit, upper limit]. The [1,1] connections are the copy back links discussed in Elman (1990).

In the first set of simulations, the author tested the model's performance during serial conditioning (see Figure 1). Learning of the CS<sub>1</sub>->US relationship should be facilitated by the presence of the intervening CS<sub>2</sub> as compared to a control without CS<sub>2</sub> (Bolles *et al.*, 1978). In the second set of simulations the roles of the onset and offset of the trace CS following both serial and trace conditioning were tested. The results were compared *qualitatively* to previous empirical research. Hence, no claims as to the correspondence between time in the model and real time will be made. The simulations demonstrated the performance of the model under different parameter settings (learning rate and momentum) to examine their effect on the qualitative results.

The output of the US node is the dependent variable of interest. This is a measure of the model's US expectancy for the following time step on a scale of [0,1]. Most previous modeling work uses the conditioned response (CR) as a dependent variable. Since I am not prepared to deal with the issues of learning vs. performance, I opted for a measure of the model's learning and suggest that the CR is a function of the US expectancy. For comparison purposes, it may be assumed that measures of CR and US expectancy are correlated.

## Simulation 1

The first set of simulations were run to examine performance of the model during serial conditioning.

## Method

Tlearn was trained on 2 different training sets. One training set represented the serial paradigm and consisted of 2 sequential CSs preceding the US. The first CS (CS<sub>1</sub>) was three time steps long and was immediately followed by the second CS (CS<sub>2</sub>), also three time steps long. During the third time step of CS<sub>2</sub>, the US was presented and lasted for one time step, overlapping with CS<sub>2</sub> (Figure 1). The ITI (inter-trial interval) alternated between 5 and 7 time steps. On the average, 12 time steps corresponded to one trial or CS-US pairing. The presence of a CS or US was signalled by a 1 at the corresponding input value. The trained output values, as in Elman's (1990) model, were the input values for the subsequent time step. Learning was accomplished via backpropagation (Rumelhart, Hinton & Williams, 1986). The model also contained 3 additional inputs/outputs that were present for future work regarding the effect of contextual cues on conditioning. For the current simulations, these values were constant with values of [.5, 1, .5]. The second training set represented a control in which the intervening CS<sub>2</sub> was absent. Previous empirical work (e.g. Bolles *et al*, 1978) suggests that learning of the CS<sub>1</sub>->US relationship should be slower following trace as compared to serial conditioning.

The model was run six times for each of three sets of parameter settings. Learning rate, designated  $r$ , and momentum,  $m$ , were set to the following: 1)  $r=.1$ ,  $m=0$ , 2)  $r=.2$ ,  $m=0$  and 3)  $r=.1$ ,  $m=.3$ . After training, performance was measured in response to CS<sub>1</sub>, CS<sub>2</sub> and the CS<sub>1</sub>->CS<sub>2</sub> compound *in the absence of the US*. No learning was permitted during this phase, thus preventing any extinction. The ITI between the end of the last US and the first of the test CSs was longer than that present during training to encourage the model to flush its temporal memory of previously occurring stimuli. The average ITI during training was 6 while that during testing was 12.

## Results

The general results are presented in Figure 3 which represent the model's performance under the  $r=.2$ ,  $m=0$  settings. Qualitative results for the other settings were quite similar and will be described below. Figure 3 illustrates the US expectancy as a function of time since CS onset. Note that for optimal prediction, the peak of US expectancy should occur on the time step *before* presentation of the US. This was true for CS<sub>1</sub> under all

of the parameter settings and after both serial and trace conditioning. Regarding the facilitation of learning the CS<sub>1</sub>->US relationship, the results are mixed. The expectancy appeared to grow faster under serial conditioning during the first few 1000 time steps. This qualitative result is clearest with the  $r=.1$ ,  $m=0$  settings.

Given the small sample size ( $n=6$ ), the only statistically significant difference between serial and trace condition peaks (at 4 time steps following CS onset) was at 30,000 time steps for  $r=.1$ ,  $m=0$  ( $t(10)=2.408$ ,  $p=.037$ ). Most of the other apparent differences at step 4 had  $p$ -values  $< .2$ .

With further training, peak expectancy to CS<sub>1</sub> under both conditions reached approximately the same asymptote (after 40K time steps,  $p=.168$  for  $r=.1$ ,  $m=0$ ;  $p=.225$  for  $r=.1$ ,  $m=.3$ ;  $p=.741$  for  $r=.2$ ,  $m=0$ ). However, note that US expectancy was significantly greater during the immediately preceding time steps under the serial paradigm.

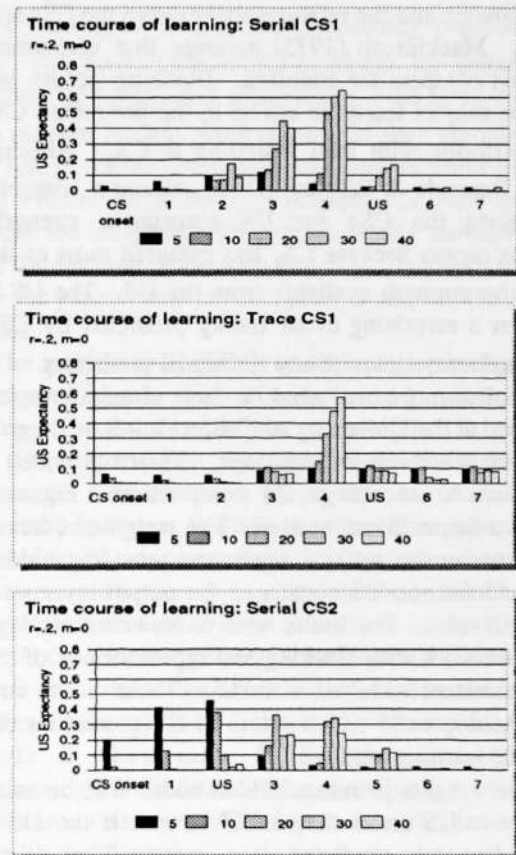


Figure 3: Time course of learning. The bars represent the US expectancy after training the model for 5, 10, 20, 30 and 40 thousand simulation time steps. The x-axis represents time steps during testing, indexed against the onset of the test CS. Onset of US during training (not testing) relative to the CS is noted on each graph.

Regarding CS<sub>2</sub>, the peak expectancy tended to occur later than optimal as training progressed. In the  $r=.1, m=.3$  case, the peak expectancies occurred two steps later than expected considering the CS<sub>2</sub>->US relationship. There was also a trend toward later expectancies as the strength of the CS<sub>1</sub>->US relationship grew. The earliest US expectancy peaks (and highest under all but the  $r=.1, m=.3$  setting) occurred during the first five to twenty thousand time steps (depending on the settings). Latency grew longer with more training.

## Discussion

The first thing to note was the common qualitative results for all three of the parameter settings. There were some differences in degree, but the trends were similar. Given the small sample size, the only conclusion regarding US expectancy peak that can be made is that the mean differences between the two conditions grew smaller with more training. A similar result was observed by Bolles *et al* (1978) in their animal subjects. They compared delay, trace and filler conditions where the filler condition represented serial conditioning. They observed facilitation (as measured by suppression ratio) in the filler condition after 16 CS-US pairings. However, after 64 pairings there were no significant differences among the three groups.

One of the more interesting simulation results is that serial conditioning did result in shorter latencies to CS<sub>1</sub> under all three parameter settings. If the goal of the system is to accurately predict the occurrence of the US, then the trace procedure was more accurate over the long run. However, in an adaptive sense, having a bit more forewarning of the USs' occurrence is beneficial. By that criteria, serial conditioning was superior, although the mechanism for this is unclear.

The graphs illustrate average performance over a number of runs. This conceals a couple of the interesting strategies adopted by some of the networks. In two runs (one at  $r=.2, m=0$ ; one at  $r=.1, m=.3$ ), the system was observed to develop no US expectancy to presentation of CS<sub>1</sub> or CS<sub>2</sub> alone but showed a normal expectancy to the compound (with a peak of approximately .8 to .9). This is evidence of configural learning where the compound is treated differently than the sum of its elements. Configural learning in animals usually results when there is differential reinforcement of the compound and its elements. However, in an experiment involving simple compound (non-serial) conditioning, Kehoe (1986) has found low levels of responding to the elements following conditioning of the compound only. In a similar vein, the model twice (once at  $r=.2, m=0$ ;

once at  $r=.1, m=0$ ) showed US expectancy following CS<sub>1</sub>, no expectancy following CS<sub>2</sub>, but more expectancy to the compound than to CS<sub>1</sub> alone. Similar empirical results have also been observed (e.g. Kehoe, 1979).

The tendency for CS<sub>2</sub> latencies to grow longer with more training may be the result of generalization from the earlier CS<sub>1</sub>. This type of generalization and the degree of supremacy of the first element in a serial compound has been extensively studied by Kehoe & Napier (1991) with eye blink conditioning in rabbits. Using serial pulse stimuli, Kehoe observed that the CR topography during test of later elements of the compound was very similar to the topography expected and observed to the first element of the compound. In Experiment 2, an A->B->C->D compound was presented where the ISI from A to the US was 400 msec. Observed CR peaks to all of the singly presented elements occurred after 400 to 450 msec despite the fact that the ISIs of B, C and D during training were 300, 200 and 100 msec respectively.

In simulation 2, I was interested in exploring the variables that drive the model's US expectancy. Moore, Desmond and Belthier's (1989) model relied on both CS onset and CS offset for its responses. The next simulation investigates the SRN model's dependence on these two variables following both trace and serial conditioning.

## Simulation 2

In this set of simulations, the duration of the test CS was systematically manipulated. If, after equating for CS onset, the latency of US expectancy was the same for all durations of the test CS, then CS onset is determining expectancy. However, if the US latency systematically covaried with the changes in duration (and thus offset), then US expectancy is based on CS offset. The data from delay conditioning in the Kehoe & Napier (1991) studies indicate that the earliest part of a sequence of pulse stimuli commands substantial responding (the *temporal primacy effect*). This might generalize to apply to the earliest part of a single CS. A different result is suggested by Boyd & Levis (1976). Their results demonstrated a greater reliance on the *later* stimuli in the compound following avoidance conditioning. However, there is a significant difference in the CS durations (and hence ISIs) in the two studies. In Boyd & Levis (1976), the CSs were 6 sec. long. Given that they were using a three component compound, the ISI from CS<sub>1</sub> to the US was 18 sec. My hypothesis was that the durations being used in the present simulations would be better approximated by

those used by Kehoe & Napiers (1991) rather than those of Boyd & Levis (1976). Note that this begs the question of optimal ISIs for the two different paradigms, NMR in the former and avoidance responding in the latter.

## Method

The training method was identical to that used in Simulation 1. The testing phase consisted of a systematic variation of the duration of CS<sub>1</sub> including time step lengths of 1, 2, 3 (the one on which it was trained) and 6. The ITI between the last training US and the first testing CS was also systematically varied to investigate any effect on the system's performance. Testing was performed following forty thousand training steps (approximately 3300 trials) for each of the parameter settings used. Two runs at each of the settings were conducted.

## Results

The results are shown in Figure 4 collapsed across parameter settings (there were no significant differences

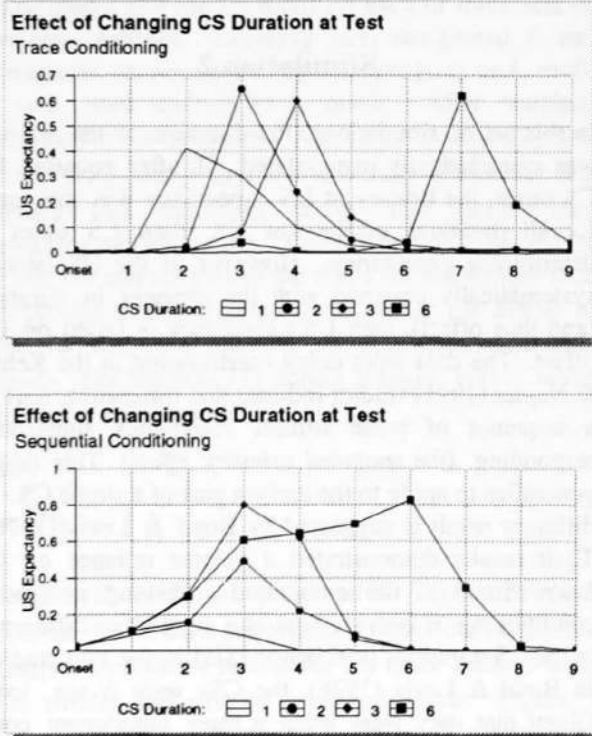


Figure 4: Effect of changing the duration of CS<sub>1</sub> at test. Time since CS onset represented on x-axis.

among results for different parameter values). Note that in the graphs, CS onset is equated across test CS durations. Offsets vary with duration (e.g. duration 6 is on through 5 on the graph and off at 6). Given the training set, a US expectancy that depends on the CS onset should peak at 4 on the graph. A US expectancy that depends on offset should peak at 2 for duration 1, 3 for duration 2, 4 for duration 3, and 7 for duration 6.

The variation of ITIs did have an impact on performance for some networks. Specifically the shortest test ITI (length 7) resulted in the poorest performance while test ITIs of 11 and 16 performed equally well.

## Discussion

It is apparent that the model depends primarily on CS offset for predicting US occurrence during trace conditioning. However during serial conditioning, the CS onset played a major role. All durations longer than 1 resulted in significant expectancy at time steps 3 & 4.

Following serial conditioning, the CS offset played a larger role in signalling when to stop expecting the US rather than in initiating expectancy. The longer the CS was on, the more sustained the expectancy. Hence, the offset tends to attenuate expectancies at time 3, 4, 5, and 8 for durations of 1, 2, 3, and 6 respectively. This helps to explain the lack of a peak at time 3 or 4 for the CS of length 1 and the drop in expectancy from 3 to 4 for the CS of length 2. The apparent drop at 4 for the CS of length 3 was not significant.

Note that the observed dependence on CS onset vs. offset will likely change for different trained CS durations. Longer CSs will drive the system to use the nearer CS offset for US initiation while shorter CSs lessen the burden of reliance on the farther onset.

The fact that the test ITI of 7 performed worst was a surprise considering that the model was trained using an ITI that varied between 5 and 7. The initial reason for testing this independent variable was to insure that the system was not learning about the regularity of US occurrence. Hence, it was a surprise that matching the test ITI to training ITI resulted in worse performance. The short training ITI may have actually retarded the network's learning.

## General Discussion

The model produced very different event representations as the result of serial vs. trace conditioning. Other intratrial models of conditioning (e.g. Grossberg & Levine, 1987; Grossberg & Schmajuk, 1987, 1989;

Klopf, 1988; Lee, 1991; Sutton & Barto, 1981, 1990) have been tested on a wide variety of conditioning paradigms. The SRN model's performance on serial conditioning demonstrates promise and it should be compared to that of the other models. Empirical work can then be planned to resolve the theoretical differences. As a model of conditioning, the current model is not comprehensive. Motivation, drive, habituation, and instrumental training have yet to be explored.

## References

- Blazis, D.E.J. & Moore, J.W. (1991). Conditioned stimulus duration in classical trace conditioning: Test of a real-time neural network model. *Behavioral Brain Research*, **43**, 73-78.
- Bolles, R.C., Collier, A.C., Bouton, M.E. & Marlin, N.A. (1978). Some tricks for ameliorating the trace-conditioning deficit. *Bulletin of the Psychonomic Society*, **11**, 403-406.
- Boyd, T.L. & Levis, D.J. (1976). The effects of single-component extinction of a three-component serial CS on resistance to extinction of the conditioned avoidance response. *Learning & Motivation*, **7**, 517-531.
- Desmond, J.E. (1990). Temporally adaptive responses in neural models: The stimulus trace. In M. Gabriel & J. Moore (Eds.), *Learning and Computational Neuroscience: Foundations of Adaptive Networks*. Cambridge, MA: MIT Press.
- Elman, J.L. (1990). Finding structure in time. *Cognitive Science*, **14**, 179-211.
- Grossberg, S. & Levine, D.S. (1987). Neural dynamics of attentionally-modulated Pavlovian conditioning: Blocking, inter-stimulus interval, and secondary reinforcement. *Applied Optics*, **26**, 5015-5030.
- Grossberg, S. & Schmajuk, N.A. (1989). Neural dynamics of adaptive timing and temporal discrimination during associative learning. *Neural Networks*, **2**, 79-102.
- Kehoe, E.J. (1979). The role of CS-US contiguity in classical conditioning of the rabbit's nictitating membrane response to serial stimuli. *Learning & Motivation*, **10**, 23-38.
- Kehoe, E.J. (1986). Summation and configuration in conditioning of the rabbit's nictitating membrane response. *Journal of Experimental Psychology: Animal Behavior Processes*, **8**, 313-328.
- Kehoe, E.J., Gibbs, C.M., Garcia, E. & Gormezano, I. (1979). Associative transfer and stimulus selection in classical conditioning of the rabbit's nictitating membrane response to serial compound CSs. *Journal of Experimental Psychology: Animal Behavior Processes*, **5**, 1-18.
- Kehoe, E.J. & Napier, R.M. (1991) Real-time factors in the rabbit's nictitating membrane response to pulsed and serial conditioned stimuli. *Animal Learning & Behavior*, **19**, 195-206.
- Klopf, A.H. (1988) A neuronal model of classical conditioning. *Psychobiology*, **16**, 85-125.
- Lee, C. (1991). Modeling the behavioral substrates of associate learning and memory: Adaptive neural models. *IEEE Transactions on Systems, Man, and Cybernetics*, **21**, 510-520.
- Mackintosh, N.J. (1975). A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological Review*, **82**, 276-298.
- Moore, J.W., Desmond, J.E., & Belthier, N.E. (1989). Adaptively timed conditioned responses and the cerebellum: A neural network approach. *Biological Cybernetics*, **62**, 17-28.
- Rescorla, R.A. & Wagner, A.R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A.H. Black & W.F. Prokasy (Eds.), *Classical Conditioning II: Current Research and Theory*. New York: Appleton-Century-Crofts.
- Rumelhart, D.E., Hinton, G.E. & Williams, R.J. (1986). Learning internal representations by error propagation. In D.E. Rumelhart & J.L. McClelland (Eds.), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition* (Vol. 1, pp. 318-362). Cambridge, MA: MIT Press.
- Sutton, R.S. & Barto, A.G. (1981). Toward a modern theory of adaptive networks: Expectation and prediction. *Psychological Review*, **88**, 135-170.
- Sutton, R.S. & Barto, A.G. (1990). Time-derivative models of Pavlovian reinforcement. In M. Gabriel & J. Moore (Eds.), *Learning and Computational Neuroscience: Foundations of Adaptive Networks*. Cambridge, MA: MIT Press.