Surprising Predictions

Draft – Do Not Distribute

October 29, 1994

Abstract

This paper describes a computational framework for theoretical models of classical conditioning. It is based on the engineering technique of Kalman filtering, which describes how systems can infer regularities in noisy and changing environments based on the results of partial experiments. The resulting class of conditioning theories has similarities with three existing (and apparently incompatible) ones; those due to Rescorla and Wagner (1972), Mackintosh (1975) and Pearce and Hall (1980). As in the Rescorla-Wagner rule, learning is dependent on errors in prediction of the unconditioned stimulus, but, as in the other rules, the amount of learning that accrues to each conditioned stimulus is a function of its own conditioning history.

1 Introduction

Modern theories of classical conditioning focus on inference or prediction. Animals are supposed to learn facets of the causal texture of their environments (Dickinson, 1980) and then use this information to choose (or learn to choose) actions appropriate to those environments. Prediction in various forms is clearly adaptive, but it is also informationally feasible, since actual outcomes are available for training. Prediction is a well-founded computational and engineering notion – the field of adaptive control is about how systems can build models of their own and their environments' behaviours and use these models to choose actions.

In this paper, we formalise the inference problem faced during conditioning in such a way that we can apply a specialisation of one of these engineering techniques. This gives computationally well grounded conditioning theories, which turn out to have close relationships with existing suggestions. We do not attempt to address the details of how animals make moment-by-moment predictions (but see Montague *et al*, 1994) but accept a conventional trial-based view of conditioning.

To introduce terms and notation, table 1 shows the course of one popular conditioning paradigm, called blocking (Kamin, 1969). A, B and C are lights, tones, somatosensory stimuli, *etc*. They are called the *conditioned* stimuli (CSs), and at the start of the experiment are assumed not to have been presented to the animal before, or at least not to have been

Group	Phase		Test
	1	2	C
control	A $\rightarrow$F	BC $\rightarrow$F	$+$
blocking	B $\rightarrow$F	BC $\rightarrow$F	$\cdot$

Table 1: Experimental design for blocking. The $+$ and $\cdot$ indicate that B acquires respectively a positive association with F for the control group and none for the blocking group.

predictive to the animal of things in its environment before. F is called the *unconditioned* stimulus (US), may be appetitive (food to a hungry animal, water to a thirsty one) or aversive (foot shock, forced confinement in a bright arena), and is assumed to have this significance to the animal from the outset.[1] It is standard to assume that each CS and US is represented by the activity (firing) of a unit or units in the animal's brain, and that CSs can predict USs if their respective units are linked. The potency of the prediction is related to the (variable) strength of the connection, and it is these strengths that change during conditioning. Quantitative theories of conditioning put forward learning rules that govern how these strengths change over trials, and use behavioural assays for confirmation.

As shown in the table, there are two learning phases to a blocking experiment, each of which consists of a moderate number of individual trials. During the first, one of the CSs (B, for the blocking group) is delivered to the animal just before the US.[2] At the heart of classical conditioning is the effect that this establishes that CS as a predictor of that US. During the second phase, the same CS (B) and a new one (C) are shown in conjunction with an identical amount of the same US. After (or possibly during) this phase, C is tested to see how much presentation of it alone induces the animal to expect F to be delivered.[3] Blocking is the phenomenon that C does not come to predict F despite having been shown in a perfectly good predictive relationship with F, if it is only shown in conjunction with another CS (B in the table) which has already been established as a predictor of F. By comparison, at the start of the second phase for the control group, B and C are both novel, and for these animals, C would generally acquire a positive association with F.[4]

[1] We do not address the difference between aversive and appetitive USs, or between appetitive USs and the absence of aversive ones.

[2] The precise timing is irrelevant for this paper, since we do not address moment-by-moment predictions. However, timing is responsible for a number of important and subtle experimental results some of which are discussed below.

[3] This can be tested in a variety of ways. For instance, if F is an aversive foot shock, then a suppression-based test may be used. For this, the extent to which C predicts F is assessed by seeing how much less frequently an animal presses a bar to get food when C is presented than when it is not.

[4] The second phase of the control experiment can be subject to a phenomenon called *overshadowing*, in which, for some stimuli, B can block learning to C even though it has not been established as a prior predictor. The inherent salience to the animals of the stimuli is thought to be responsible, and this effect can

Blocking demonstrates the inadequacy of naive associational theories, in which temporal and spatial proximity are the only determiners of learning. Indeed varieties of blocking have been major driving forces behind the development of more sophisticated conditioning theories.

To describe the conditioning rules mathematically, we call $\mathbf{x}^t$ the activity of all the CSs on trial t, where, again for simplicity, $x_i^t = 1$ if CS i is presented on trial t, and $x_i^t = 0$ if it is not. r^t similarly represents whether or not the US was presented on trial t (for most of the experiments there is just one US, so r^t can be treated as a scalar). $\mathbf{w}^t$ contains the connection strengths (and thus the prediction strengths) for all the CSs.

Different conditioning theories make different hypotheses about the form of the learning rule, which specifies $\Delta\mathbf{w}^{t+1} \equiv \mathbf{w}^{t+1} - \mathbf{w}^t$. Our discussion follows Dickinson (1980), although he uses a different notation.[5] Various theories agree that:

$$\Delta w_i^{t+1} = f_i(\mathbf{x}^t, \mathcal{H}) g(r^t, \mathbf{x}^t, \mathbf{w}^t, \mathcal{H}) \tag{1}$$

where f_i and g describe how much processing CS i and the US get as a function of which CSs are actually present and the conditioning history $\mathcal{H}$ (the collection $\{(\mathbf{x}^s, r^s), s < t\}$ of previous trials experienced by the animal). The theories disagree about the forms of f and g, and about the 'sufficient statistics' (the way that $\mathcal{H}$ is summarised so that the animal is not assumed to have to remember every individual trial). Dickinson (1980) motivates the multiplicative form of equation 1 by describing the amount of learning as being dependent on the extent to which the CSs and the US are *processed* by the animal. f_i is also called the *associability* of CS i as it is the extent to which the CS is able to enter into associations.

There are two broad schools of conditioning. One school (Rescorla & Wagner, 1972) is based on net prediction – CS processing is essentially constant across trials for all the CSs that are present, irrespective of learning history, but the amount that the US is processed is a function of how poor the prediction of it was, where that prediction is based on all the CSs present. If $y^t = \mathbf{w}^t \cdot \mathbf{x}^t$ is the summed prediction of the US for trial t based on the CSs $\mathbf{x}^t$, then the Rescorla-Wagner (RW) rule tries to make $y^t = r^t$. It changes the weights according to:

$$\Delta w_i^{t+1} = \alpha_i x_i^t \delta^t \tag{2}$$

where $f_i = \alpha_i$ is a fixed learning rate or associability and $\delta^t = r^t - y^t$ is the error in the current prediction. RW exhibits blocking since $\delta^t = 0$ during the second phase of learning, once B correctly predicts F, and therefore the weight for C never gets elevated.

Sutton and Barto (1981) pointed out that this is a form of the delta rule, a familiar rule

be controlled for by reversing the roles of B and C in the blocking group.

[5] He calls the CSs, E1, and the US, E2. This is convenient, since under various theories (Wagner, 1978; 1981), CSs can predict each other as well as USs.

from engineering (Widrow & Stearns, 1985), and there is a large body of theory about the circumstances under which it converges appropriately.

The other school (Mackintosh, 1975; Pearce and Hall, 1980) has its roots in the notion that animals have limited processing capacities, and have to allocate them adaptively. If this were true of a putative learning processor, then animals might choose to ignore some CSs in favour of others. This puts the emphasis for explaining conditioning phenomena on the f_i rather than the g.

Mackintosh (1975) reasoned that the capacity should be allocated preferentially to those stimuli that have been predictive of USs in the past. This seems rational if there are some stimuli that never have interesting consequences and so should be ignored. In our notation, Dickinson's (1980) formalisation of Mackintosh's theory uses the learning rule:

$$\Delta w_i^{t+1} = \alpha_i^t x_i^t (r^t - w_i^t) \tag{3}$$

where the associabilities α_i^t are now functions of the conditioning history. Note that the US processing terms are different for all the CSs, and are arranged so that all the stimuli that predict the US will have the *same* asymptotic strength.[6] According to Mackintosh $\alpha_i^{t+1} > \alpha_i^t$ if CS i is a better predictor of the US on trial t than all the other CSs present, and $\alpha_i^{t+1} < \alpha_i^t$ otherwise. What is less clear in the Mackintosh rule is how the w_i^t for the various CSs present on a trial should be combined during testing to generate a net prediction for the US.[7]

Pearce and Hall noted that Mackintosh's rule has an unfortunate consequence – it suggests that animals should devote all their processing resources to contingencies which are very familiar to it, whereas in fact it is things associated with changes that are ecologically more important. They performed an experiment (Hall and Pearce, 1979, shown in table 2) in which they first established a CS as an excellent predictor of a US (a small foot shock), and then showed the same CS in conjunction with a different US (a large foot shock).[8] The CS learned more slowly during the second phase than a novel CS, rather than faster, as Mackintosh's rule would predict. They suggested that stimuli gain associability if they are followed by unexpected events, and Dickinson (1980) formalised their associability rule (PH) as:

$$\alpha_i^{t+1} = |\delta^t| \tag{4}$$

[6]This reveals the roots of this rule. CSs that are in the right temporal relationship with the US should learn to predict it entirely *by themselves*, it is just that the learning rates α_i for some of them can be set to 0 by attentional processes. The goal of RW learning is different – it is constructing the correct *summed* predictions.

[7]One possibility is the so-called noisy-or function first popularised in Bayesian belief networks, Pearl (1986). Here one would interpret w_i as being like the probability that if $x_i = 1$ then $r = 1$. Then the overall probability that $r = 1$ would be $1 - \prod_i (1 - x_i w_i)$.

[8]It is irrelevant in the present context that this could be interpreted as a more intense version of the same US.

where δ^t is the error in the predictions at time t. He also noted that 'an elegant feature of the Pearce-Hall theory lies in the fact that it does not need to assume that [US] processing changes at all during the course of learning. In the case where the animal is exposed to simple [CS-US] pairings, no further increments in learning will occur as soon as [the CS] perfectly predicts [the US] because, when this point has been reached, [the CS] will no longer receive the processing necessary for associative learning.(p156)' The overall PH rule can therefore be written as:

$$\Delta w_i^{t+1} = \alpha_i^t x_i^t r^t \tag{5}$$

Grossberg's conditioning theory (Grossberg and Schmajuk, 1987; Grossberg and Levine, 1987) has some similarities with PH, in that learning is prompted by surprise – however its mechanisms are somewhat more complicated.

In this paper, we cast the learning problem of conditioning into the slightly broader context of any system trying to make inferences about the regularities in uncertain and mutable environments. We follow the analysis of Sutton (1992a;b) in deriving the form of learning in equation 1 from the computationally well-grounded notion of Kalman filtering (Kalman, 1960), and show that the resulting two learning rules that he suggested have elements of both schools of conditioning. Since prediction is the goal of Kalman filtering, learning is driven by failures in predictions (as in the RW school). However it behoves systems to learn about the relationship between CSs and USs as fast as possible, but no faster, and this implies that different stimuli should support different amounts of learning, *ie* should be associable to different degrees. It turns out that one of the learning rules fails the test of the conditioning data, and we investigate the other empirically.

The next section develops the Kalman filtering based account of conditioning, following Sutton (1992a;b), and discusses two different on-line versions which constitute conditioning theories in the standard sense. Section 3 compares them with the existing theories, and shows one in operation in a version of the Hall and Pearce (1979) experiment. The discussion evaluates the rules, suggests some avenues to explore for how they might be implemented in the vertebrate brain, and comments on some of the evidence that they ignore.

2 Kalman Filtering and Conditioning

Kalman filtering (Kalman, 1960) is a popular technique in engineering for handling uncertainty in cases where systems have some information about the way their environments change, but have only imperfect data on its current state. The degree of uncertainty changes as the system makes partial observations of the world. We follow Sutton (1992a;b) and Gluck *et al* (1992) in their application of Kalman filtering to conditioning theory.

We take there to be an actual state of the world at any time which is set by the experimenter and is expressed as a set of true weights $\mathbf{u}^t$. The animal's inference task is to make its version of the weights $\boldsymbol{w}^t$ as close as possible to their true values (in a sense described below), using the learning trials as evidence for $\mathbf{u}^t$. In a single learning trial, the experimenter presents CSs $\mathbf{x}^t$, and the animal expects that the amount of US r^t presented will be determined according to:

$$r^t = \mathbf{x}^t \cdot \mathbf{u}^t + n^t \tag{6}$$

where n^t represents noise in the trial. Kalman filtering is easiest to describe in the case of continuous rather than discrete r^t, $\mathbf{u}^t$ and n^t, even though for most conditioning experiments the presence of the US is all or nothing. We comment on this aspect of the model in the discussion.

A key part of Kalman filtering is that the animal has a model which sanctions the possibility that $\mathbf{u}^t$ can change from one trial to the next. It must assume that $\mathbf{u}^t$ *can* change, to allow for the vagaries of an inconstant world, but it is less clear what assumptions it might make. The simplest model is that:

$$\mathbf{u}^{t+1} = \mathbf{u}^t + \mathbf{m}^t \tag{7}$$

where $\mathbf{m}^t$ is (multi-dimensional) noise, and expresses the animal's ignorance as to the exact change.

The inferential difficulty for the subject can readily be appreciated if there is more than one CS. Not only does the animal need some way of averaging out the noise n^t over trials to avoid drawing conclusions too hastily, but also, unless the experimenter is benevolent and arranges for each trial to involve exactly one CS, every trial gives it only one single piece of information r^t about all the components of $\mathbf{u}^t$ for which the CSs are provided. In fact, every experimental set up involves contextual cues as well as the explicit experimental CSs, and so this is not merely a moot point.

In the most straightforward form of Kalman filtering, the task for the animal is taken to be to maintain in $\boldsymbol{w}^t$ the least mean square estimate of $\mathbf{u}^t$ in the case that the initial uncertainty in $\mathbf{u}^0$ and the noise processes $\mathbf{m}^t$ and n^t are Gaussian. We will develop the theory as if the animal could store and process an entire covariance matrix, representing its uncertainty as to $\mathbf{u}^t$. However, this is an unreasonable requirement for an animal, and the actual conditioning rules developed from the general theory are based on more tractable approximations.

If

$$\mathbf{u}^t \sim \mathcal{G}\left[\hat{\mathbf{u}}^t, \Sigma^t\right], \;\; \mathbf{m}^t \sim \mathcal{G}\left[0, \mathcal{M}\right], \;\; n^t \sim \mathcal{G}\left[0, \mathcal{N}\right],$$

where Σ^t is the covariance matrix at trial t, then, after observing $\mathbf{x}^t$, but before knowing the rewarding outcome:

$$\begin{pmatrix} \mathbf{u}^{t+1} \\ r^t \end{pmatrix} \sim \mathcal{G}\left[\begin{pmatrix} \hat{\mathbf{u}}^t \\ \mathbf{x}^t \cdot \hat{\mathbf{u}}^t \end{pmatrix}, \begin{pmatrix} \Sigma^t + \mathcal{M} & \Sigma^t \mathbf{x}^t \\ [\mathbf{x}^t]^T \Sigma^t & [\mathbf{x}^t]^T \Sigma^t \mathbf{x}^t + \mathcal{N} \end{pmatrix}\right] \tag{8}$$

This formula shows how the uncertainty at trial t propagates into trial $t+1$.

The animal then observes r^t in the experiment, and is in a position to update its estimates of the state of the world $\mathbf{u}^t$ in the light of this. In Kalman filtering, it calculates the marginal distribution of $\mathbf{u}^{t+1}$ given r^t, namely:

$$\mathbf{u}^{t+1}|r^t \sim \mathcal{G}\left[\hat{\mathbf{u}}^t + \frac{\Sigma^t \mathbf{x}^t}{[\mathbf{x}^t]^T \Sigma^t \mathbf{x}^t + \mathcal{N}}(r^t - \mathbf{x}^t \cdot \hat{\mathbf{u}}^t),\ \ \Sigma^t + \mathcal{M} - \frac{\Sigma^t \mathbf{x}^t [\mathbf{x}^t]^T \Sigma^t}{[\mathbf{x}^t]^T \Sigma^t \mathbf{x}^t + \mathcal{N}}\right] \tag{9}$$

and since this distribution is still Gaussian, the inference process can be repeated.

This distribution suggests a learning rule: that the weights should be set to be the mean of the distribution $\boldsymbol{w}^t = \hat{\mathbf{u}}^t$ and that the matrix Σ^t of covariances should be maintained accordingly, *ie:*

$$\Delta w_i^{t+1} = \frac{[\Sigma^t \mathbf{x}^t]_i}{[\mathbf{x}^t]^T \Sigma^t \mathbf{x}^t + \mathcal{N}}(r^t - \mathbf{x}^t \cdot \boldsymbol{w}^t) \tag{10}$$

The rule for the update of the weights suggested by equation 10 is an instance of the general form of equation 1, where f_i, the CS associability terms, are $[\Sigma^t \mathbf{x}^t]_i / ([\mathbf{x}^t]^T \Sigma^t \mathbf{x}^t + \mathcal{N})$, and g, the US processing term, is δ^t, the prediction error from the RW rule, that is the difference between the amount of US that was actually delivered and the amount that the animal would have predicted in the absence of noise and if it believed its existing weights.

As it stands, the rule in equation 10 is unattractive for two reasons. First, it requires the animal to store the full covariance matrix – which has entries for the joint uncertainty of every pair of CSs. Second, updating the covariance matrix requires full knowledge of $\mathcal{M}$ and $\mathcal{N}$ and execution of a complicated matrix operation. There is no evidence that uncertainties about different stimuli actually combine, and we therefore follow Sutton (1992a;b) in using simpler, diagonal, approximations to the covariance matrix which devote only a single uncertainty measure to each stimulus and are thus easier to accommodate neurobiologically. In the learning rules using this uncertainty measure, the associability of each stimulus is proportional to its diagonal element.

In Sutton's (1992a;b) first rule, called K1, he took:

$$\Sigma^t = \text{diag}\left[e^{\beta_1^{t+1}}, e^{\beta_2^{t+1}}, \ldots, e^{\beta_n^{t+1}}\right]$$

in equation 10, and set the β_i^t using a form of stochastic gradient descent in the prediction error, *ie:*

$$\begin{aligned} \Delta \beta_i^{t+1} &\propto -\frac{\partial}{\partial \beta_i}(r^t - \mathbf{x}^t \cdot \boldsymbol{w}^t)^2 \\ &= (r^t - \mathbf{x}^t \cdot \boldsymbol{w}^t) x_i^t \quad h_i^t \end{aligned}$$

where h_i^t is a decaying sum of past weight changes. As Sutton explains, this boosts the amount of learning that happens to a CS if the current recommended weight change is well correlated with recent weight changes, on the grounds that had the weight been changed more on the previous occasion the current prediction error would be less. This is a similar heuristic to the ones adopted by Jacobs (1988), Fahlman (1988), Tollenaere (1990), and Silva & Almeida (1990) for speeding the course of gradient descent learning. Note that once the system stops making errors, *ie* once $r^t = \mathbf{x}^t \cdot \boldsymbol{w}^t$, the learning rates do not regress towards 0 (which would require $\beta_i^t \to -\infty$) but instead remain at whatever values they hold.

The second rule that Sutton defined, K2, starts from the definition that:

$$\Sigma^t = \mathcal{E}\left[(\mathbf{u}^t - \boldsymbol{w}^t)(\mathbf{u}^t - \boldsymbol{w}^t)^T\right]$$

Taking $\delta^t = r^t - \mathbf{x}^t \cdot \boldsymbol{w}^t$, one can deduce that:

$$\begin{array}{rrcl} & \mathcal{E}[(\delta^t)^2] & = & \mathcal{N} + [\mathbf{x}^t]^T \Sigma^t \mathbf{x}^t \\ \text{so if} & \Sigma^t & = & \text{diag}\,[s_{11}^t, s_{22}^t, \ldots, s_{nn}^t] \\ \text{then} & (\delta^t)^2 & \sim & \mathcal{N} + \sum_i s_{ii}^t\,[x_i^t]^2 \end{array} \tag{11}$$

Therefore, for K2, the associabilities of the stimuli should be proportional to their correlation with the total amount of surprise, as measured by the squared prediction error.[9] In practice, under K2, previous learning trials are used to estimate the values of s_{ii}^t, and so the associabilities are functions of the correlations with past poor predictions. We updated s_{ii}^t according to:

$$\Delta s_{ii}^{t+1} = \gamma \left[x_i^t\right]^2 \left((\delta^t)^2 - \mathcal{N} - \sum_i s_{ii}^t \left[x_i^t\right]^2\right). \tag{12}$$

where γ is a fixed learning rate. *A priori,* $s_{ii}^t > 0$ (in either its covariance or learning rate interpretations), and so we enforced this directly during the update.[10]

There is one remaining oddity that needs to be cleared up. In equation 11, it appears as if the contribution to the total uncertainty of the noise $\mathcal{N}$ is inherently different from the learned contributions from the stimuli. In fact, if one imagines that there are contextual cues in every experimental set-up, then the animal can learn to associate these with $(\delta^t)^2$ too, just by having a bias term in the prediction of $(\delta^t)^2$.

3 Comparing the Rules

We have now defined a total of five different conditioning rules – RW, Mackintosh, PH and the two Kalman filter rules, K1 and K2. In this section, we compare and contrast

[9] Since x_i^t is either 0 or 1, $(x_i^t)^2 = x_i^t$.

[10] Sutton (1992b) used a more complicated procedure (recursive least squares) to determine s_{ii}^t.

them. K1 and K2 bear close relations to the existing rules, effectively combining certain of their properties.

Both Kalman filter rules share their US processing term ($g = \delta^t$ in equation 1) with the RW rule. This means that they exhibit blocking partially for the same reason as the RW rule – if $\delta^t = 0$ then $\Delta \boldsymbol{w}^{t+1} = 0$ too, and so it is only on trials in which errors are made that learning happens. However, RW cannot model the results of experiments that suggest that CSs are processed differently according to the conditioning history, whereas K1 and K2 can. In this they are closer to the Mackintosh and Pearce and Hall (PH) rules.

As an example of the effects of changing CS associability, consider Mackintosh's explanation of blocking. During the first trial in the second phase in table 1, the novel stimulus (C) should learn about the US (F) in the usual way based on whatever the starting value of its associability α_C^1. However, on that trial, it will be apparent to the animal that B has already been established as a better predictor of F than C, and therefore α_C^2 will be reduced towards 0. Hence, using the learning in equation 3, C will support little learning to F and will therefore appear to be blocked.

There is no direct analogue in the Kalman filter rules for the Mackintosh rule formalised as in equation 3. However there are some points of similarity. First, the CS associability term $[\Sigma^t \mathbf{x}^t]_i / ([\mathbf{x}^t]^T \Sigma^t \mathbf{x}^t + \mathcal{N})$ exactly embodies the notion of competition for the learning on trial t between the different CSs present. If we make the diagonal approximation to the covariance matrix, then if Σ_{ii}^t is largest component and $\mathbf{x}_i^t = 1$, then CS i attracts the most learning, and also inhibits the learning to the other CSs. However, the diagonal component Σ_{ii}^t is not clearly a function of how predictive CS i is of the US, but is a surrogate for this quantity. Rules K1 and K2 offer different surrogates.

According to rule K1, Σ_{ii}^t will be large if stimulus i consistently attracts weight changes of the same sign. This will tend to be true if the animal expects a comparatively large amount of noise on each trial, as $\mathcal{N}$ appears in the denominator in equation 10 forcing all the associabilities to be low. This means that, as is evident in standard learning curves, it will take a number of trials for a newly predictive stimulus to reach asymptote for its US, and during that time, its weight will always change in the same direction. Therefore, during these times, Σ_{ii}^t can indeed be interpreted as a measure of how predictive CS i has recently been for the US, and the overall K1-based rule is quite similar to the Mackintosh rule. Figure 1 below shows this effect.

K1 has one further similarity to the Mackintosh rule – as mentioned above, once the animal has learnt successfully the association between the CS and US, the associability of the CS does not decrease to 0, but rather stays at whatever value it last adopted. This highlights the remaining major difference between K1 and the Mackintosh rule. For K1, this final value can be anything, whereas for Mackintosh it is when the CS perfectly predicts that US that its associability will be highest. The issue of terminal associability was the focus of experiments by Hall and Pearce (1979) which falsified the prediction of the Mackintosh rule and also straightforward versions of K1.

Table 2 and the associated graph show this experiment. To generate the graphs, stimulus A was shown to the animals at various stages during the second phase of the experiment. The more strongly A predicted a shock, the less the animals were prepared to press a bar for food in its presence and so the lower the suppression ratio. The A-alone group shows latent inhibition (*eg* Lubow, 1974; Reiss & Wagner, 1972). This is the effect that presentation of a CS without any US consequence retards subsequent learning to that CS. The A-shock group was the focus of the experiment. During the first phase, A is demonstrated to be a good predictor of something, namely the mild shock (that this is not very consequential to the animal is irrelevant). Mackintosh's rule would suggest that this should establish A as a good predictor, and therefore enhance its associability. However in the second phase, learning to A for the A-shock group is slower than for the B-shock group, for which A was novel. For the PH rule, the associability of A decreases as its predictions in the first phase become established.

According to rule K2, Σ^t_{ii} will be large if stimulus i has recently been associated with unpredicted events. This is exactly the driving force behind the PH rule. In fact, Dickinson's form for the associability terms in the PH rule in equation 4 is particularly close to K2 if there is only one CS, since they both make α_i depend in similar ways on $|\delta|$, the surprise on the preceeding trial. To put it another way, if nothing surprising happens, *ie* if δ^t tends towards 0 (or rather towards $\mathcal{N}$, the intrinsic noise level for the experiment), all the associabilities decline.

We implemented K2 and simulated it both in a standard conditioning paradigm and in the Hall and Pearce experiment. The exact updates (from equations 10 and 12) were:

$$\Delta w_i^{t+1} = \frac{s^t_{ii} x^t_i}{\mathcal{N} + \sum_i s^t_{ii}(x^t_i)^2}(r^t - \boldsymbol{x}^t \cdot \boldsymbol{w}^t), \tag{13}$$

$$\Delta s^{t+1}_{ii} = \gamma \left[x^t_i\right]^2 \left((\delta^t)^2 - \mathcal{N} - \sum_i s^t_{ii} \left[x^t_i\right]^2\right) \tag{14}$$

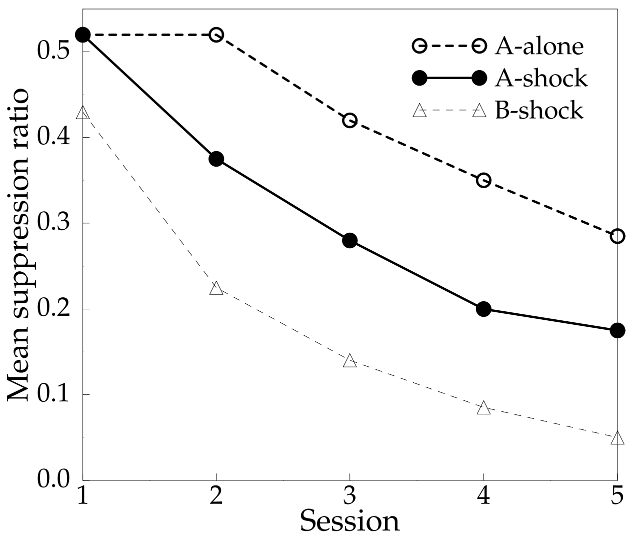
If, in equation 14, s^{t+1}_{ii} would be less than 0, then we enforced $s^{t+1}_{ii} = 0$. This preserves the positivity of s^t_{ii}.

Figure 1 shows how s^t develops as a proportion of s^0 for a simple conditioning task in which there is only one CS (we drop the CS index i) and either one (US$^+$) or no (US$^-$) US. The values of the parameters and the initial weights are given in the legend to the figure. In the US$^+$ condition, s^t first rises, as the system learns that the CS is unexpectedly predictive, and then it falls as the predictions improve. The early rise is the explanation under this model of effects consistent with the Mackintosh rule; the late fall demonstrates PH. In the US$^-$ case, the associability only falls, since the CS is never predictive.

Dickinson's implementation of the Pearce-Hall rule (equations 4 and 5) would not show

Group	Phase					
	1			2		
A-alone	A			A	→	S
A-shock	A	→	s	A	→	S
B-shock	B	→	s	A	→	S

Table 2: Experimental design of Hall and Pearce's (1979) experiment. 's' is a mild shock; 'S' a larger one.



The suppression of lever-pressing for food for presentation of A during the second phase of the experiment to the left. After Hall and Pearce (1979).

---

the steady rise in the first trials of the US$^+$ case. From that one would expect a steep rise, followed by a slow decay as the predictions become correct.

The K2 model of the Hall and Pearce experiment is based on this graph, where the first phase in table 2 is taken to be equivalent to about 30 trials. At the start of the second phase, for the A-alone group, A has an associability of 0 like the US$^-$ group after learning. For the A-shock group, A has an intermediate associability, since A has been predictive of something, whereas for the B-shock group, A has the initial associability since, for these animals, A is initially unfamiliar. Figure 2 shows the development of the weights over the course of the second phase. Since Hall and Pearce (1979) used a suppression test, these curves should be interpreted as the inverse of those in the graph attached to table 2 and thereupon are similar. The starting value of the weights are taken to be 0, since even for the A-shock group, the first phase of learning is designed to manipulate the associability of A rather than the strength of its actual association with the shock.

Comparing the graph associated with table 2 and figure 2, we see that K2 captures the learning behaviour quite closely, except for the very first session.

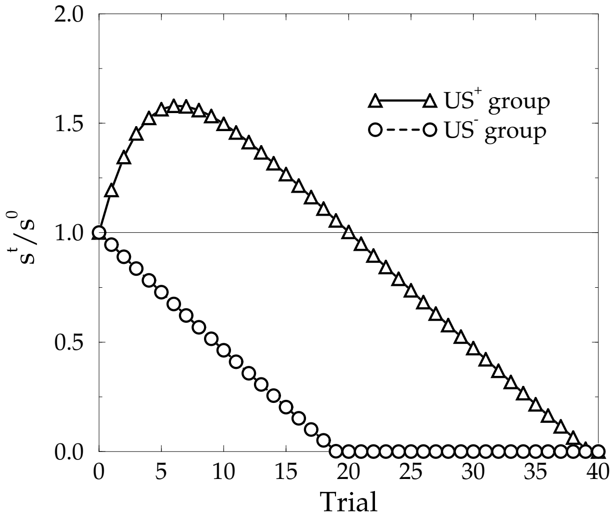


Figure 1: Pavlovian Conditioning. The course of learning for the associabilities under a conventional Pavlovian conditioning paradigm with (US^+) and without (US^-) a unit strength US. The learning rate for the associabilities $\gamma = 0.005$, the starting weights are 0, the starting associabilities are 0.02, the minimum possible associability is 0, and the assumed noise level $\mathcal{N} = 0.2$, although there is actually no noise and the experiment is deterministic.

4 Discussion

We started from a computationally well-founded principle of the goal of conditioning learning, namely to use the CSs to make as good predictions as possible of the USs. Fast and appropriate learning has been the driving force rather than optimising the use of a limited learning capacity. We derived from this notion Kalman filtering rules K1 and K2 that have strong similarities with certain parts of all three of the major conditioning theories that we have discussed – the RW, the Mackintosh and the PH rules. In particular, K2 is based on a delta rule, just like RW, and also has associabilities that change according to learned correlations between CSs and unpredicted happenings, which is just the motivation behind the PH rule. K2 is also supported by experiments that support PH. Of course the matches with PH and the data are not perfect, and there are a number of important details that have been neglected.

In this section we discuss how these rules might be implemented, and comment on inadequacies and possible extensions of the Kalman framework.

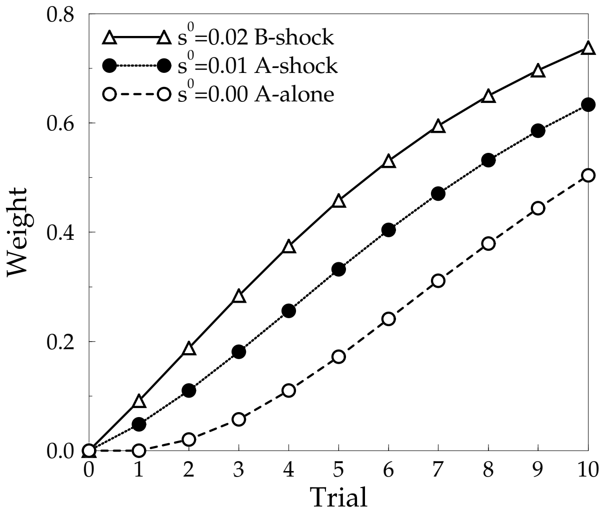


Figure 2: The development of the weights for the K2-based model of Hall and Pearce (1979). The same parameters were used as in figure 1 except that the starting associabilities were $s^0 = 0.02, 0.01$ and 0.00 for the B-shock, A-shock and the A-alone groups respectively.

Neurobiological Substrate

The empirical evidence somewhat favours K2 over K1, because of its similarity with the PH rule. The natural next question is whether it is possible to identify in the vertebrate brain the signals that K2 would require for its implementation and the associabilities that are its consequent. This is relevant, since theoretical impasses such as the ones we have described below over conditioned inhibition *versus* downwards unblocking may have explanations in the details of an implementation. We do not expect the neural substrate exactly to reflect the quantities the rules require.

Dopamine acts as a neuromodulator in the vertebrate brain and has long been associated with the processing of and learning about rewarding events. Schultz and his colleagues (*eg* Ljunberg *et al*, 1992) have made electrophysiological recordings from dopamine projecting cells in the ventral tegmental area (VTA) of monkeys that were performing (classical and instrumental) conditioning tasks. The VTA is a sub-cortical nucleus that has dopaminergic projections to the prefrontal cortex and parts of the limbic system (Oades & Halliday, 1987). Using this evidence, we (Quartz *et al*, 1992; Montague *et al*, 1994; Dayan, 1994) suggested a concrete hypothesis based on a rule called the temporal difference (TD) rule (Sutton, 1988; Sutton and Barto, 1987; 1989; McLaren and Dickinson, 1990; Konorski, 1948) for the role the VTA might be playing in appetitive conditioning learning. Having identified the link between RW and the delta rule, Sutton and Barto developed the TD rule to account for aspects of the moment-by-moment time course of conditioning that are not apparent in conditioning based on trials. TD bases learning on on-going prediction errors, the temporally sensitive version of the δ^t that is used in the RW rule and K1 and K2. In Schultz's conditioning tasks, the VTA dopamine cells seem to be reporting just these prediction errors. We noted that learning rules such as the one in equation 2 have a conventional correlational or Hebbian nature, except the correlation is between the pre-synaptic activity x_i^t and the error signal δ^t rather than the post-synaptic activity.

Norepinephrine (NE) is another neuromodulator in the vertebrate brain whose cortical progenitor is a different mid-brain nucleus, the *locus coeruleus* (LC). The LC is small ($20,000$ neurons a side in the human (ref)), but its NE output cells have massive axonal fields, having $2,000,000$ cortical synapses (ref). It also has substantial sub-cortical projections. Whereas the VTA appears to respond to rewards, and stimuli that predict rewards in the appropriate way, the LC responds to *surprising* events, as assessed in a variety of experimental paradigms (see Robbins *et al*, 1985). We suggest that LC neurons are involved in reporting the equivalent of $(\delta^t)^2$, the amount of surprise or prediction error on trial t, in the same way that the VTA is involved in reporting the amount of rewarding US on a trial. This quantity is just what is required for K2.

Whereas manipulations of the cortical action of dopamine have significant effects on appetitive conditioning, lesions of the dorsal noradrenergic ascending bundle (Robbins *et al*, 1985), a main route for NE into the cortex, appear to have little effect on blocking or latent inhibition. Since these are two of the main conditions in which the effects of different CS learning rates are apparent, it is possible that the cortical NE projection sets associabilities indirectly and not directly.

As described so far, K2 just requires information about the total amount of surprise. However the VTA does not simply provide information about the total amount of reward, but rather produces a temporally sensitive prediction error. If a CS at time s_1 predicts that a reward will be delivered at time $s_2 > s_1$, then, after learning, the VTA fires at s_1 and not at s_2 – the activity at the later time is predicted *away* by the earlier activity. Similarly K2 could be implemented in such a way that the output of the LC reflects ongoing differences between the amount of prediction error that the system expects, on the basis of the stimuli that it observes, and the amount of prediction error that it actually registers (Schmidhuber, 1991). We call the extended K2 rule the rule in which the amount of US processing is dependent not only on δ^t, but also on the extent to which the animal is expecting to be surprised at the time of the US.

This has theoretical implications. Consider the well known partial reinforcement effect. In this, CSs that are associated with variable USs are much slower to extinguish than CSs that are associated with constant USs, even though their asymptotic levels of association are higher. Extinction happens when the CSs are presented on their own without the

Group	Phase		Test
	1	2	B
conditioned inhibition	A $\rightarrow$ 2S	AB $\rightarrow$ S	—
downwards unblocking	A $\rightarrow$ S$\sqcup$S	AB $\rightarrow$ S	+

Table 3: Experimental design for inhibitory learning and unblocking. For downwards unblocking (Mackintosh, Bygrave and Picton, 1977), S represents a shock, 2S represents twice as much shock, and $\sqcup$ represents a 10 second delay between presentations of that shock. The — indicates that B comes to be a conditioned inhibitor of S.

USs with which they have been associated. During extinction, the CSs appear to lose this association. Under the conventional K2 rule, there will appear to be higher levels of observation noise $\mathcal{N}$ in the experiments with variable rewards than those with constant rewards. This translates into slower extinction learning because of the presence of $\mathcal{N}$ in the denominator of the learning rule for the weights (equation 10), which reduces the amount of learning on every trial. This effect will be balanced against the effect that the CSs themselves will be associated with higher levels of noise, and would therefore, other things being equal, have higher learning rates, on the basis of equation 11. This, contrary to the partial reinforcement effect, would tend to make learning faster to the variable CSs. However, in the extended K2 rule, the CSs predict that the US will be variable, and so the lack of delivery of the US during extinction receives less processing because it is less surprising. One would therefore predict slower learning.

Theoretical Omissions

A remaining theoretical problem in K2 is the use of the delta rule – *ie* the use of $g = \delta^t$ in equation 10 rather than $g = r^t$ as in equation 5. Using δ^t is very attractive computationally (if the goal is to use CSs jointly to make predictions of a US then it seems eminently reasonable to use the failure in their current, joint, prediction, to drive learning). Indeed, Dickinson's (1980) formulation of the PH rule uses exactly this same error to drive changes to learning rates. Unfortunately the cases which show up the difference between rules based on the error and rules not based on it are fraught with other theoretical conflicts, making it difficult to use them to reject rules unambiguously. We will give two examples based on the experiments in table 3.

The first row shows an experiment designed to induce inhibitory conditioning to stimulus B. During the first phase, A is associated with the presentation of 2 foot shocks, one right after the other. During the second phase, A and B together are followed only by a single foot shock. The error δ^t is negative during the early part of the second phase of learning (since A predicts too much shock), and so under the delta rule, one might expect B to come to predict the absence of a predicted shock. Indeed, there is some evidence for this (Holland, 1988),[11] which contradicts the PH rule. For that, after the first trial of the second phase, the animal should realise that $|\delta^t|$ is large, boost α_B^{t+1} accordingly, and then, using equation 5 during the second trial, assign B a large association with the shock ($r^t > 0$ as a single shock is still presented). The failure of PH suggests that the animal is using errors in its predictions to decide on what learning to do as well as how much learning to ascribe to particular stimuli.

However, the downwards unblocking paradigm in the second row of table 3 suggests just the opposite. Here, A is also followed by two presentations of a shock, but now there is a 10 second interval between them. Once again, in the second phase of learning, B signals the removal of a predicted shock, but now, rather than coming to predict this absence, as in the conditioned inhibition case, experimentally it actually predicts the *presence* of a shock. This supports PH and contradicts the delta rule. Holland (1988) has done careful studies of the time course of these effects, and also the nature of the predictions being made by each stimulus. It is clear that something about the moment-by-moment predictions is causing the difference, but a theoretical understanding of this is still somewhat elusive. In fact the relationship between excitatory and inhibitory conditioning is a major focus of the difference between attentional and delta-type rules. Resolution of this seems to require advances in understanding the neural basis behind conditioning.

The use of an on-line rule like K2 avoids the animal having to know or estimate the covariance matrix $\mathcal{M}$ for the way that the true weights change over time. In fact, one might expect animals to model this second order effect too. A variety of experiments with higher primates show learning set effects – *ie* they can learn not only that different CSs are associated with different USs, but also that the associations on one trial will change in some regular way (*eg* be reversed) on the next trial. Unfortunately this makes the resulting Kalman filter non-linear and therefore intractable, even if the starting uncertainty and the noise corrupting each experiment were still Gaussian. Non-linear Kalman filtering is much less well understood than the linear variants. This same concern has prevented us from using more realistic models for the noise, the true weights, and the update. In most experiments, US delivery is discrete rather than continuous – the animal gets an integral number of pellets of food, or a fixed quantity of sugar water. This makes our Gaussian model for the weights and the noise in partial reinforcement contexts rather unreasonable. In simple cases, it should be possible to understand empirically the effect of these changes on the learning rates, but one would not expect a qualitative difference.

Another major flaw of the model is the way in which issues of the precise time-course of conditioning have been ignored. We have seen in the difference between downwards

[11] Alternating trials in which 'A $\rightarrow$2S' and 'AB $\rightarrow$S' are presented is actually a more reliable way of producing a conditioned inhibitor.

unblocking and conditioned inhibition that the moment-by-moment unfolding of the delivery of the USs is important.

Gallistel (1990) points out many of the problems mentioned above and others for standard models of conditioning. He discusses a somewhat more complicated conditioning theory that is also computationally well-grounded and is specifically designed to handle complicated time-courses of CS and US delivery. It does this by performing spectral analysis on the time series, and using information from this to set weights. Gallistel's procedure involves the equivalent of a matrix inversion to work out which CSs are most reliably predictors of the US when considered over the whole time-course of the experiment. Although the Kalman rules do the equivalent of this in an on-line manner, unlike Gallistel's rule, they are sensitive in different ways to recent conditioning trials *versus* distant ones. Although, as Gallistel points out, his rule is attractive in having no free learning parameters, this is not necessarily a good thing. Gallistel's computational model is not designed for cases in which the true state of the world can change over time – it lacks the wherewithal to forget stale associations as fast as can be necessary. This theory would need to be augmented with a model of environmental change.

So far, the only role accorded to contextual cues is that of being associated with the intrinsic amount of noise in the conditioning experiments. In fact context plays a more subtle role. For instance, in many circumstances latent inhibition does not transfer from one context to another (Lovibond *et al*, 1984). If a CS is shown repeatedly in one experimental chamber so that the animal no longer pays attention to it, but before conditioning that CS to a US, the animal is moved to an evidently different chamber, then learning to the CS can be normal. This suggests that the associabilities are not functions just of stimuli but also of the context. This makes sense – stimuli are bound to be differentially predictive in different environments – but we have not modelled it.

Stimulus-Stimulus Prediction

In his SOP model, Wagner (1978; 1981) suggested a different way of looking at CS processing. He suggested that as well as predicting USs, CSs would come to predict each other, in both cases using a type of RW rule. In this case, in the same way as in blocking that once a US is completely predicted, it supports no learning, a CS, once predicted, would also support no learning. The usual instrument of predicting CSs would be the various contextual cues that are present in every experiment and are known to exert significant influences. For the experiments we have so far discussed, Wagner's rule has a similar effect to PH (and therefore K2), since the predictability of a CS and the predictability of its consequences have always been tied – however SOP also entails that the animal builds a stimulus-stimulus model of aspects of its world that are not directly reinforcing.

SOP does not provide a complete explanation of the apparent changes to the associability of stimuli during conditioning, as is clear from the removal of latent inhibition by surprising, but irrelevant, consequences. Latent inhibition was mentioned above as the effect that a CS that is repeatedly shown without US consequence supports less learning than an unpresented CS. If, instead of providing no consequences, different, novel, but unreinforcing, stimuli are presented following the CS, then the latent inhibition effect goes away. SOP would predict that the novel stimuli should not have this effect because the CS *itself* still becomes familiar, it is only its consequences that change.

However, we certainly believe that the stimulus-stimulus predictions are constructed. A wealth of experiments has demonstrated that animals have the capacity to build models of their worlds, and it is would be attractively parsimonious if one could show that similar principles govern this learning as govern stimulus-reward/punishment learning. One major trouble with stimulus-stimulus prediction is that there are far too many stimuli that could potentially do the predicting and that could be predicted and some way of choosing between them is required. A (recursive) notion is that stimuli that predict surprising things or rewarding/punishing things that are the ones that are processed and, like secondary reward conditioning, are deemed worthy of prediction themselves.

The extent to which stimulus-reward conditioning can, under certain circumstances, be mediated partly by stimulus-stimulus conditioning is currently moot. For instance, a CS like a light could predict the size, shape, and smell of a food pellet (thus accessing their known relationship with reward) as well as or instead of predicting its reward characteristics directly. Indeed, there are suggestions from lesion studies that these two classes are neuropsychologically dissociable. We envisage both sets of processes going on in tandem, with learning rates being set by coerulear computations.

References

[1] Dayan, P (1994). Computational modelling. *Current Opinion in Neurobiology,* **4**, 212-217.

[2] Dickinson, A (1980). *Contemporary Animal Learning Theory.* Cambridge, England: Cambridge University Press.

[3] Fahlman, SE (1988). Faster-learning variations on backpropagation: An empirical study. In DS Touretzky, GE Hinton & TJ Sejnowski, editors, *Proceedings of the 1988 Connectionist Models Summer School.* San Mateo, CA: Morgan Kaufmann, 38-51.

[4] Gabriel, M & Moore, JW, editors (1989). *Learning and Computational Neuroscience,* Cambridge, MA: MIT Press.

[5] Gallistel, CR (1990). *The Organisation of Learning.* Cambridge, MA: MIT Press, Bradford Books.

[6] Gluck, MA, Glauthier, PT & Sutton, RS (1992). Dynamically modifiable stimulus associability in network models of category learning. *Proceedings of the 14th Annual Conference of the Cognitive Science Society.*

[7] Grossberg, S & Levine, DS (1987). Neural dynamics of attentionally modulated Pavlovian conditioning: Blocking, inter-stimulus interval, and secondary reinforcement. *Applied Optics,* **26**, 5015-5030.

[8] Grossberg, S & Schmajuk, NA (1987). Neural dynamics of attentionally modulated Pavlovian conditioning: Conditioned reinforcement, inhibition and opponent processing. *Psychobiology,* **15**, 195-240.

[9] Hall, G & Pearce, JM (1979). Latent inhibition of a CS during CS-US pairings. *Journal of Experimental Psychology: Animal Behavior Processes,* **5**, 31-42.

[10] Holland, PC (1988). Excitation and inhibition in unblocking. *Journal of Experimental Psychology: Animal Behavior Processes,* **14**, 261-279.

[11] Jacobs, RA (1988). Increased rates of convergence through learning rate adaptation. *Neural Networks,* **1**, 295-307.

[12] Kalman, RE (1960). A new approach to linear filtering and prediction problems. *Transactions of the ASME Series D: Journal of Basic Engineering.* **82**, 35-45.

[13] Kamin, LJ (1969). Predictability, surprise, attention and conditioning. In BA Campbell & RM Church, editors, *Punishment and Avoidance Behavior,* 276-296. New York, NY: Appleton-Century-Crofts.

[14] Konorski, J (1948). *Conditioned Reflexes and Neuron Organisation.* Cambridge, CUP.

[15] Lovibond, PF, Preston, GC & Mackintosh, NJ (1984). Context specificity of conditioning, extinction, and latent inhibition. *Journal of Experimental Psychology: Animal Behavior Processes,* **10**, 360-375.

[16] Ljunberg, T, Apicella, P & Schultz, W (1992). Responses of monkey dopamine neurons during learning of behavioral reactions. *Journal of Neurophysiology,* **67**(1), 145-163.

[17] Lubow, RE (1973). Latent inhibition. *Psychological Bulletin,* **79**, 398-407.

[18] Mackintosh, NJ (1975). A theory of attention: Variations of the associability of stimuli with reinforcement. *Psychological Review,* **82**, 276-298.

[19] Mackintosh, NJ, Bygrave, DJ & Picton, BM (1977). Locus of the effect of a surprising reinforcer in the attenuation of blocking. *Quarterly Journal of Experimental Psychology,* **29**, 327-336.

[20] McLaren, IPL & Dickinson, A (1990). The conditioning connection. *Philosophical Transactions of the Royal Society of London, B,* **329**.

[21] Montague, PR, Dayan, P & Sejnowski, TJ (1994).A computational approach to diffuse ascending systems. In preparation.

[22] Oades, RD & Halliday, GM (1987). Ventral tegmental (A10) system: neurobiology. 1. Anatomy and connectivity. *Brain Research,* **434**, 117-65.

[23] Pearce, JM & Hall, G (1980). A model for Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review,* **87**, 532-552.

[24] Pearl, J (1988). *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference.* San Mateo, CA: Morgan Kaufmann.

[25] Quartz, S, Dayan, P, Montague, PR, Sejnowski, TJ (1992) Expectation learning in the brain using diffuse ascending connections. *Society for Neuroscience Abstracts,* **18**, 1210.

[26] Reiss, S & Wagner, AR (1972). CS habituation produces a "latent inhibition effect" but no active "conditioned inhibition". *Learning and Motivation,* **3**, 237-245.

[27] Rescorla, RA & Wagner, AR (1972). A theory of Pavlovian conditioning: The effectiveness of reinforcement and non-reinforcement. In AH Black & WF Prokasy, editors, *Classical Conditioning II: Current Research and Theory,* 64-69. New York, NY: Appleton-Century-Crofts.

[28] Robbins, TW, Everitt, BJ, Cole, BJ, Archer, T & Mohammed, A (1985). Functional hypotheses of the coeruleocortical noradrenergic projection: A review of recent experimentation and theory. *Physiological Psychology,* **13**, 127-150.

[29] Silva, FM & Almeida, LB (1990). Acceleration techniques for the backpropagation algorithm. In LB Almeida & CJ Wellekens, editors, *Neural Networks,* Lecture Notes in Computer Science **412**, Berlin: Springer-Verlag, 110-119.

[30] Schmidhuber, JH (1991). *Adaptive Curiosity and Adaptive Confidence.* Technical Report FKI-149-91, Institut für Informatik, Technische Universität München.

[31] Sutton, RS (1988). Learning to predict by the methods of temporal difference. *Machine Learning,* **3**, 9-44.

[32] Sutton, RS (1992a). Adapting bias by gradient descent: An incremental version of delta-bar-delta. In *Proceedings of AAAI-92.*

[33] Sutton, RS (1992b). Gain adaptation beats least squares? In *Proceedings of the 7th Yale Workshop on Adaptive and Learning Systems.*

[34] Sutton, RS & Barto, AG (1981). Toward a modern theory of adaptive networks: Expectation and prediction. *Psychological Review,* **88** 2, 135-170.

[35] Sutton, RS & Barto, AG (1987). A temporal-difference model of classical conditioning. *Proceedings of the Ninth Annual Conference of the Cognitive Science Society.* Seattle, WA.

[36] Sutton, RS & Barto, AG (1989). Time-derivative models of Pavlovian reinforcement. In Gabriel & Moore (1989).

[37] Tollenaere, T (1990). SuperSAB: fast adaptive back propagation with good scaling properties. *Neural Networks,* **3**, 561-573.

[38] Wagner, AR (1978). Expectancies and the priming of STM. In SH Hulse, H Fowler & WK Honig, editors, *Cognitive Processes in Animal Behavior.* Hillsdale, NJ: Lawrence Erlbaum Associates, 177-210.

[39] Wagner, AR (1981). SOP: a model of automatic memory processing in animal behaviour. In NE Spear and RR Miller, editors, *Information Processing in Animals, Memory Mechanisms,* Hillsdale, NJ: Lawrence Erlbaum.

[40] Widrow, B & Stearns, SD (1985). *Adaptive Signal Processing.* Englewood Cliffs, NJ: Prentice-Hall.