COMPUTATIONAL MODELS OF CLASSICAL CONDITIONING: A COMPARATIVE STUDY

Christian Balkenius
Jan Morén

christian.balkenius@fil.lu.se
jan.moren@fil.lu.se

Lund University Cognitive Science
Kungshuset, Lundagård
S-222 22 Lund, Sweden

Abstract

We describe computer simulation of a number of associative models of classical conditioning in an attempt to assess the strengths and weaknesses of each model. The behavior of the Sutton-Barto model, the TD model, the Klopf model, the Balkenius model and the Schmajuk-DiCarlo model are investigated in a number of simple learning situations. All models are shown to have problems explaining some of the available data from animal experiments. The ISI curves for trace and delay conditioning for all the models are presented together with simulations of acquisition and extinction, reacquisition, blocking, conditioned inhibition, secondary conditioning and facilitation by an intermittent stimulus. We also present cases where some of the models show an unexpected behavior.

1 INTRODUCTION

In Pavlov's (Pavlov 1927) well known experiment, a dog would repeatedly hear a tone which was always followed by the presentation of food. When the food was presented, the dog started to salivate. After a number of such presentations, the dog had learned that food follows shortly after the tone and would start to salivate already when it heard the sound. The pairing of a signalling *conditioned stimulus* (CS) with a motivationally significant *unconditioned stimulus* (US) defines classical conditioning as an experimental procedure. While the unconditioned stimulus produces an *unconditioned response* (UR), the response produced by the conditioned stimulus is called the *conditioned response* (CR) and is used as an index of the learning process. Initially, no CR is produced, but as learning progresses, the intensity or probability of the CR increases.

Classical conditioning has attracted researchers for many years and many models have been developed. Recently, attention has focused on the real-time dynamics of conditioning. Below, we investigate some well known real-time models of conditioning in a number of computer simulations. The goal is to demonstrate their strengths and reveal their weaknesses.

Although traditionally seen as a very simple phenomenon, classical conditioning has offered unexpected resistance to theoreticians. Still, almost a hundred years after Pavlov's initial experiments, there exist no model capable of explaining the full range of phenomena under this heading. During the reign of behaviorism, classical conditioning was marginalized and considered to be of little significance. It was never given a satisfactory explanation, however. In fact, classical conditioning contains some of the hardest problems in learning today.

It contains real-time aspects, such as inter-stimulus (ISI) effects and the modelling of timing and topology of the conditioned response. The rate of conditioning is optimal at a certain ISI. For example, in classical eye-lid conditioning in rabbit, the optimal ISI is approximately 250 ms and decays exponentially with time (Smith, Coleman and Gormezano 1969, Schneiderman and Gormezano 1964). Also, independently of the ISI, the CR tends to appear slightly before the US (Desmond 1990).

Conditioning is also sensitive to the statistical dependency between events (Rescorla 1968). While it was once believed that the pairing of CS and a US were sufficient for conditioning, it has been shown that this is not the case. In an important experiment, Rescorla (1968) presented dogs with random pairing of CS and US. On some trials the CS was presented alone, on some it was followed by the US, and on other trials the US occurred without the CS. In one experimental group, the CS did not predict the US, while in the other it predicted the US at half of the trials. In the first group, no conditioning occurred, while in the second, the animals produced the CR. The response occurred at a rate proportional to the predictive level of the CS.

Furthermore, learning depends on previous experiences (Kamin 1968). For example, in the blocking paradigm, a CS does not acquire any association when it occurs simultaneously with another

stimulus that has already been conditioned with the US. Another example is reacquisition which is yet another example where learning depends on previous experiences.

The specifics of conditioning is also greatly influenced by sensory modality and the type of response. For example, Garcia and Koelling have shown that some associations are very easy to acquire, for example, the association of taste to subsequent illness, while other association, such as taste to shock, is almost impossible to learn (Garcia and Koelling 1966).

While most contemporary models of conditioning are good in some areas, they are often much worse off in other. The main problem that hampers computational model is that it is very hard to construct a model that cover a large area of experimental conditions. It is not uncommon to find broad claims in the literature about the ability of various models, but at closer examination, all models generate inappropriate prediction in some areas. (See sections 3-7). It is most unsatisfactory that the general opinion is that conditioning is a phenomenon that is simple to model and that has been solved a long time ago.

We believe that the main reason for this situation is the limited amount of detailed studies that have been made of the different models. This is partly because it can be very hard to correctly implement the models. Very seldom are the models specified at sufficient detail to allow other researchers to use the models. When mathematical description are given, they are often incomplete or confusing.

The rest of this paper is an attempt to compare a number of computational models of conditioning. We have chosen a number of popular associative models. These models were selected since they are fairly similar and make similar claims about their abilities.

In the next section, we describe the different experiments used in the study. Section 3 reports simulations with the model proposed by Sutton and Barto (1981). The more recent development of this model, the TD-model (Sutton and Barto 1987) is described in section 4. In section 5, the same experiments are tested with the model described by Klopf (1982, 1988). In section 6, the model developed by Balkenius (1995, 1996, 1998) is tested. Finally, in section 7, we investigate the model proposed by Schmajuk and DiCarlo (1992). The order of the presentation represents the similarities between the different models.

2 THE EXPERIMENTS

Some typical instances of classical conditioning experiments were chosen to test the models. All these experiments were selected on the grounds that at least one of the models claim to produce the effect. The experiment selected where the ones that we considered most basic for the classical conditioning paradigm. Some of the models admittedly handle much more complicated cases than the ones presented, but our main aim here is to see to what extent the different models can account for very simple learning situations.
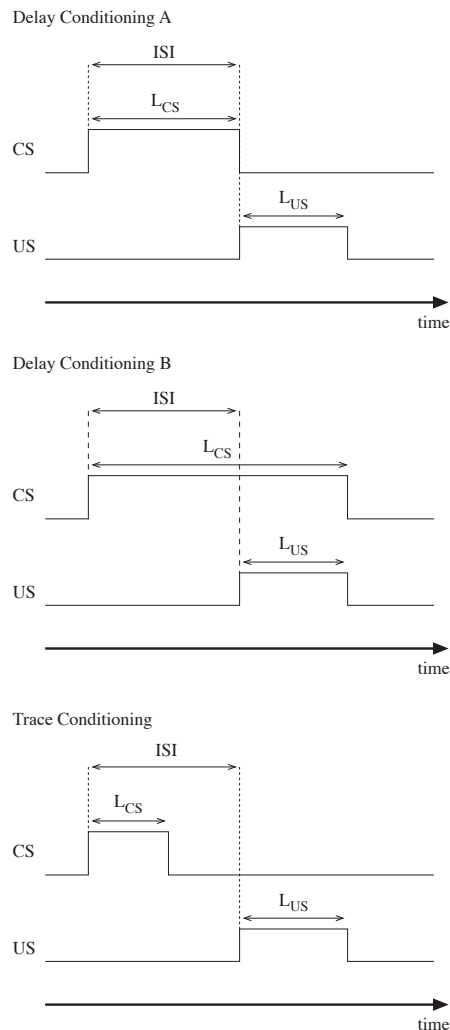

Figure 2.1. Delay and trace conditioning. In delay conditioning of type A, the length of the CS (L_{CS}) is equal to the inter-stimulus interval (ISI), for delay conditioning of type B, the length of the CS is equal to the ISI plus the length of the US, that is the CS and the US terminates simultaneously. For trace conditioning, the CS terminates before the onset of the US.

2.1 Acquisition

The experiment described in the introduction has been termed an acquisition procedure and requires the most fundamental process in classical conditioning: the ability to establish an association when a CS is paired with a US. We use the notation below to describe the experiment in a general sense.

$$CS + US \Rightarrow CS \rightarrow CR$$

A CS is first followed by the US at a number of presentations, and as a result, a CS that is later presented on its own will be able to produce a CR. All models tested can obviously model acquisition to some extent. One would be hard put to claim that a model describes classical conditioning if this was not the case.

A feature of some empirical studies of acquisition of associations is that the response level forms an S-shaped curve similar to a sigmoid function (Schneiderman, Fuentes and Gormezano 1962). There are various explanations of this phenomenon, and many

see it as desirable for a model to exhibit this feature. Of the models tested in this paper only the Klopf model, and to some extent, the Balkenius model show this effect. Interestingly, the Klopf model loses this feature during reacquisition.

2.2 Inter-stimulus interval effects

The effect of the inter-stimulus interval on the response level is one of the primary learning effects described in Sutton and Barto (1990). Empirical studies have been done with both trace conditioning, where the CS terminates before the US is presented; and with delay conditioning, where the CS offset occurs at the onset of the US, or later (see figure 2.1).

We have tested the various models with three types of timing. Two types of delay conditioning and one type of trace conditioning. In the first type of delay conditioning, that we will call delay A, the CS terminates exactly at the onset of the US. In the second type, delay B, the CS continues to be present until the termination of the US. Finally, in trace conditioning, the CS and US have fixed lengths and only the ISI changes. In trace conditioning, the ISI can be both positive and negative.

The empirically determined profile, according to Schneiderman (Schneiderman 1966) and Smith et al. (1969) can be found in figure 2.2. The desirable behavior is for the response level to have a single peak at small positive ISI:s, no response at all for negative ISI:s, and asymptotically declining values as the ISI grows large.

All models tested in this paper are affected by variable ISI:s, albeit for very different reasons, as will be seen below.
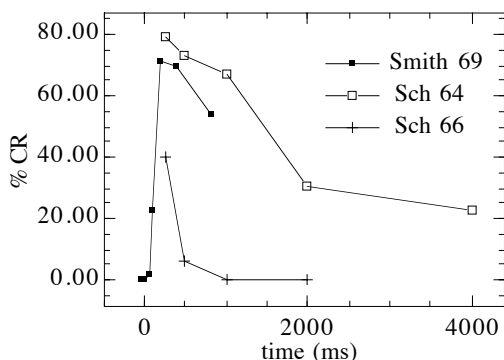

Figure 2.2. The effect of the inter-stimulus interval in classical nictitating membrane conditioning. Delay B conditioning after Scheniderman and Gormezano 1964). Trace conditioning after Smith et al. (1969) and Schneiderman (1966).

2.3 Extinction

In an extinction experiment, an earlier acquired response disappears as the CS is presented on its own without a subsequent US. The training procedure can be summarized as,

$$CS_1 \Rightarrow CS_1 \rightarrow no\ CR$$

This is again a very fundamental ability of any conditioning models since it shows that learning is in some sense reversible. When the CS is presented on its own, it no longer predicts the US and should no longer produce any CR. All the models tested below can handle this situation, which, of course, is not very surprising. It is interesting to note that, like for ISI-effects, the mechanisms behind extinction are somewhat different.

2.4 Reacquisition Effects (Savings effects)

The reacquisition, or savings, effect appears when an animal relearns an association that has previously been extinguished. Learning the second time is faster than during the original learning phase, and it is faster still the third time (Pavlov 1927).

To test for this effect, one first lets the system learn an association by pairing a CS and the US. Then the CS is presented by itself a number of times, until the response disappears, implying that the association is extinguished. When the CS is paired with the US again, the number of trials required to relearn the association should be less than it was the previous time. If so desired, this can be repeated any number of times.

We tested the reacquisition effect by running the association–extinction cycle four times. Of the models described here, only the Klopf and the Schmajuk-DiCarlo models show reacquisition effects, and for the Klopf model, it only appears for the first relearning phase.

A possibly related effect is called spontaneous recovery. It results when the CS of an extinguished association is not presented for a while. When the cue reappears, the CR too may come back. This is often assumed to be the result of some type of passive forgetting during the period without the CS. None of the models presented can model this aspect of conditioning.

2.5 Blocking

Blocking, is a very desirable effect for a model to handle well. It is one of the most well known features of all classical conditioning (Kamin 1968, Rescorla and Wagner 1972). The regime for blocking is:

$$CS_1 + US \Rightarrow CS_1 \rightarrow CR$$

$$(CS_1 + CS_2) + US \Rightarrow CS_2 \rightarrow no\ CR$$

The parenthesis is used to indicate that CS_1 and CS_2 are presented simultaneously. When CS_1 already predicts the US, there will be no conditioning of CS_2 to CR; CS_1 blocks CS_2. All models tested perform well on tests for this effect. The blocking experiment shows that acquisition of an association is not independent of earlier learning and that all the stimuli present influence learning in a non-local way.

2.6 Conditioned Inhibition

In Pavlov's account for extinction, he describes it as the acquisition of an inhibitory association that removes the effect of the excitatory association from the CS (Pavlov 1927). That is, extinction is considered an active process rather than some passive decay of an association. The existence of an inhibitory association can be shown in a conditioned inhibition experiment. In

such an experiment, the animal is first conditioned with a test stimulus CS_2. In a second phase, trials occur where the CS_0 is paired with the US, and trials where CS_0 together with CS_1 do not predict the US. Finally, CS_1 is paired with CS_2. If CS_1 has acquired inhibitory properties, it will be able to inhibit the CR that CS_2 would otherwise produce:

Phase I	$CS_2 + US$
Phase II	$CS_0 + US$
	$(CS_0 + CS_1) + $ no US
Test	$(CS_1 + CS_2) \rightarrow$ no CR

It is assumed that CS_1 takes on its inhibitory properties during the trials where it is presented together with CS_0 that on its own would predict the US.

2.7 Secondary Conditioning

When CS_1 has been associated with an US, CS_1 acquires reinforcing properties and can itself be used to reinforce a CS_2. This is called secondary conditioning. The learning regime for this effect is:

$CS_1 + US \Rightarrow CS_1 \rightarrow CR$

$CS_2 + CS_1 \Rightarrow CS_2 \rightarrow CR$

The effect is typically weak and very dependent on the exact timing of CS_1 and CS_2, as CS_1 will extinguish at the same time that CS_2 is reinforced. All models except the Schmajuk-DiCarlo model can handle secondary conditioning effects (but see section 4 about the TD model). This effect is especially important if one wants to use classical conditioning as a part of an instrumental learning system (Klopf, Morgan and Weaver 1993, Balkenius 1995). In this case, the chaining of responses is often described as secondary conditioning.

2.8 Facilitation by an Intermittent Stimulus

A different type of facilitation of acquisition can be seen when an extra stimulus CS_2 is introduced inbetween CS_1 and the US. If the conditioning to CS_1 is weak due to a long inter-stimulus interval, the extra CS_2 will facilitate conditioning to CS_1. The two cases can be summarized by:

Normal

$CS_1 + US \Rightarrow CS_1 \rightarrow$ weak CR

Facilitated

$CS_1 + CS_2 + US \Rightarrow CS_1 \rightarrow$ strong CR

Note that the interval between CS_1 and the US is the same in both cases. One possible explanation for the facilitation effect is as a form of secondary conditioning. It is therefore interesting to see that the Schmajuk-DiCarlo model that does not demonstrate secondary conditioning still shows the facilitation effect as does all the models tested.

3 THE SUTTON-BARTO MODEL

The Sutton-Barto (SB) model (Sutton and Barto 1981), is an early time-derivative model of reinforcement in conditioning. It works in real-time in contrast to the Rescorla-Wagner model (Rescorla and Wagner 1972), from which it is descended. This model is a precursor to many later computational models, including the TD model, the SBD model (Blazis *et al*. 1986) and the Klopf model (Klopf 1988).

The reinforcement in the SB model is the time derivative of the sum of the stimulus strengths:

$$\dot{Y}_t = Y_t - Y_{t-\Delta t},$$
$$Y = \sum_i CS_i V_i + V_{US}US, \qquad \text{(SB1)}$$

where the V_i is the stimulus strength of CS_i, and V_{US} is the strength of US. With suitably chosen constants, these equations will account for all predictions of the Rescorla-Wagner model.

An eligibility trace mechanism is added to account for temporal relationships:

$$\overline{X}_{i,t+1} = (1-\delta)\overline{X}_{i,t} + \delta X_{i,t},$$
$$X_{i,t} = \begin{cases} 1, & CS_i \text{ present} \\ 0, & CS_i \text{ absent} \end{cases} \qquad \text{(SB2)}$$

The rate of decay δ is in the range $0<\delta<1$. The same trace is also used in the TD model. The reinforcement signal and the eligibility trace are combined as:

$$\Delta V_i = \beta \dot{Y} \times \alpha_i X_i \qquad \text{(SB3)}$$

with α_i and β_i positive constants as in the Rescorla-Wagner model.

We have run a series of simulations covering all learning experiments described in the previous section. During these simulations, the constants chosen were $\alpha_i=0.1$, $\beta=1$, $\delta=0.2$; these are the values used by Sutton and Barto.

As the model is an early one (in fact, it is the oldest one covered in this paper), it is not surprising that it has a number of problems with these simulations.

Like the other models, SB has no problems with acquisition and extinction. Unlike the Klopf and Balkenius models, SB works over a wide range of CS-US intervals due to the nature of the eligibility trace function used. However, SB does not model the S-shaped acquisition curve that Klopf and Balkenius does.

The Inter–stimulus interval effects are poorly modelled by this method. For example, the model shows inhibitory conditioning with ISI close to zero during trace conditioning (figure 3.1 and 3.2). This seems to be inconsistent with empirical data (Smith *et al*. 1969). The model also fails to account for ISI effects during delay conditioning, with no decay of asymptotic learning with longer ISI:s, and with overall learning level dependent on proximity between CS_{off} and US_{off}, as seen in figure 3.1. As Sutton and Barto notes in (Sutton and Barto 1990), this model does not, and is not intended to, account for these effects.
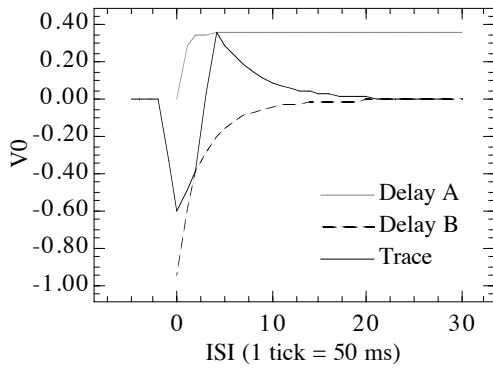
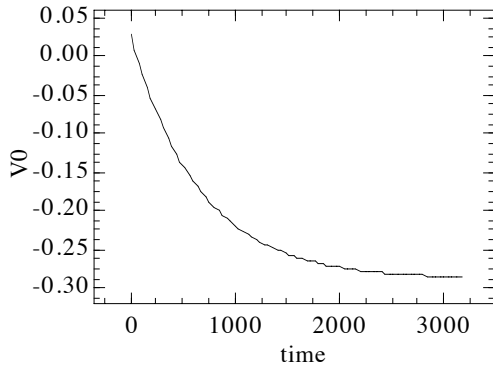**Figure 3.1.** The behavior of the SB model during ISI trials

Figure 3.2. Inhibitory conditioning when ISI=2, $US_{length}=2$, $CS_{length}=4$ during trace conditioning.

No reacquisition effects are modeled by SB. Blocking effects are handled satisfactory, which is not surprising, as the capabilities are a superset of Rescorla-Wagner, and blocking is one of the effects modeled well by that model.

Secondary conditioning with no overlap between CS_0-US or CS_1-CS_0 works as expected; CS_1 is reinforced by CS_0, albeit weakly, as CS_0 is simultaneously extinguished. However, when the inter-stimulus intervals are short enough to make the stimuli overlap, an erroneous inhibitory conditioning is produced. Briefly, the reason is that when CS_0 and US overlap, V_0 is initially small, making $\dot{Y}$ small. Thus, change in X is the primary factor in changing V_0. The greatest change in X is for CS_{off}, and when overlapped with US, this negative change will influence V_0, producing the effect. The analogous effect is achieved for phase two of secondary conditioning, as V_0 is large and V_1 is small.

Facilitation effects are modelled fairly well; the stimulus trace model used directly facilitates this effect.

4 THE TD MODEL

The TD (temporal difference) model (Sutton and Barto 1987) was introduced as an extension of the earlier Sutton-Barto model (Sutton and Barto 1981), which in turn was a real-time extension of the Rescorla-Wagner model (Rescorla and Wagner 1972). According to the TD-model, the goal of conditioning is to predict the temporally discounted value of all future rewards. At time t, this prediction is called $\overline{V}_i$ and is calculated as the sum of all weights V_i for the CS_i:s at time t:

$$\overline{V}_t = \max(\sum V_i CS_i, 0). \qquad \text{(TD1)}$$

Here, the weight V_i represents the contribution from CS_i to the total prediction. During learning, these weights are updated according to the following equation,

$$\Delta V_i = \beta\left(\lambda_{t+1} + \gamma\overline{V}_{t+1} - \overline{V}_t\right) \times \alpha_i\overline{X}_i, \qquad \text{(TD2)}$$

where $\lambda_{\tau+1}$ represents the strength of the US at time t+1, γ is the discount factor, α_i and β are learning rate constants, and $\overline{X}_i$ is a trace of stimulus CS_i. Thus, the first term of (TD1) can be seen as the expected change in US from t to t+1 (called reinforcement by Sutton), while the second term is a function of temporal proximity to the CS. The definition of $\overline{X}_i$ is identical to (SB2) above.

This section describes a number of simulations of the TD model. The actual code was downloaded from R. Sutton's web-site and interfaced with our simulator. The model itself is, thus, identical to Sutton's implementation. Using this implementation, we have successfully reproduced the simulation data as presented in Sutton and Barto (1990).

The TD model is able to reproduce several aspects of classical conditioning. Most importantly, it models the ISI-dependency shown in empirical studies. Figure 4.1 shows the asymptotic value of the weights as a function of the ISI for trace conditioning.

Unlike the other models discussed in this paper, the ISI-curve is an emergent effect of the learning equations and handles a variety of different ISI:s and stimulus lengths.

The TD-model also handles blocking effects in a satisfactory manner, but the presence of secondary conditioning depends on the stimulus length. This appears as a shortcoming of the model. However, the existence of secondary conditioning is somewhat controversial. Some studies have shown the existence of this effect, but in many cases, it has been hard to reproduce (See Klopf 1988). It may be the case that these results could be explained by differences in stimulus length, but this is debatable.
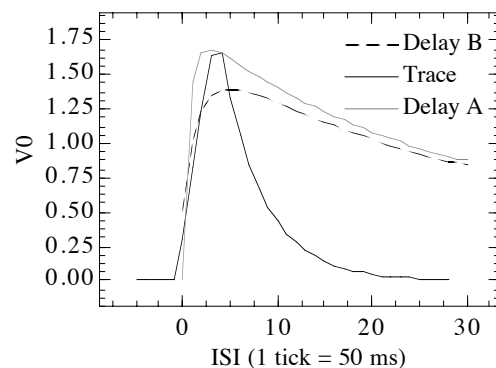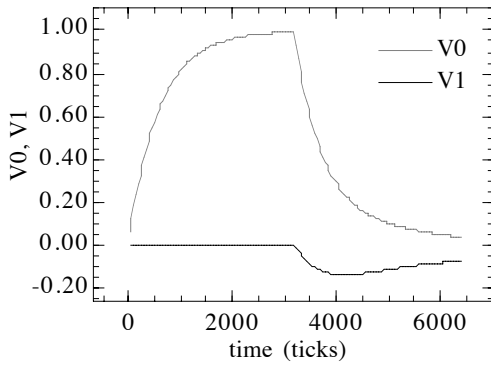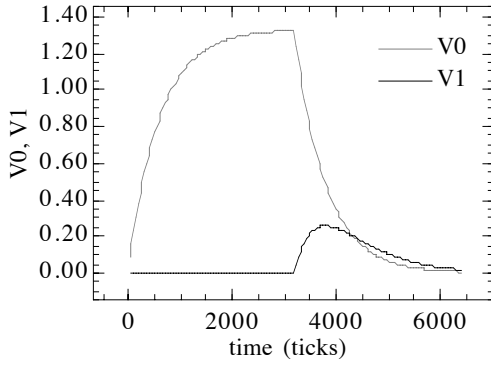

Figure 4.1 The ISI curve for the TD model.

Figure 4.2. Secondary conditioning in the TD model with and without overlapping stimuli. The upper curves show V_1 and the lower curves show V_2. (TOP) With non-overlapping CS_1 and CS_2, the secondary conditioning effect is clearly visible. (BOTTOM) With overlapping conditioned stimuli, V_2 shows inhibition rather than excitation.

Interestingly, the TD model suffers from a similar problem as the Sutton-Barto model for secondary conditioning: overlapping CSs produce inhibition rather than excitation, as shown in figure 4.2. The reason this effect is not apparent for first-order acquisition, is that US level is not taken into account when computing $\overline{V}_i$, in contrast to the Sutton-Barto model.

The results for blocking, inhibitory conditioning and facilitation are the same as for the Sutton-Barto model; the model performs reasonably well. Like the SB model, the TD model does not show the S-curve for acquisition. A related shortcoming is that the TD model does not show any reacquisition effects or spontaneous recovery.

5 THE KLOPF MODEL

The Klopf model was introduced in 1982 as the Drive-Reinforcement (DR) model (Klopf 1982), but we present the model as described in Klopf (1988). In this model, the output at time t, the UR or the CR, is given by the value $y(t)$. This is calculated as the sum of all CS representations x_i multiplied with their corresponding weights w_i. The value θ is a threshold which was set to 0 in the simulations reported by Klopf (1988).

$$y(t) = \sum_{i=1}^{n} w_i(t)x_i(t) - \theta. \qquad (K1)$$

During conditioning the weights change according to the equation,

$$\Delta w_i(t) = \Delta y(t) \sum_{j=1}^{\tau} c_j |w_i(t-j)| \Delta x_i(t-j), \qquad (K2)$$

where c_j are learning constants, $|w_j|$ are the magnitude of the individual weights, and Δx_i the change in the CS representation. It is stated in Klopf (1988) that only *positive* changes should be considered in the equation above. That is, if Δx_i is less than 0 it is set to 0 for the purposes of equation (K2). It should also be noted that there are both excitatory and inhibitory weights in the model. These are treated separately and are constrained to stay on the positive and negative side respectively. It is also required that the weights must be larger than 0. In the simulations in Klopf (1988), a minimum of 0.1 is used.

Equation (K2) illustrates a fundamental assumption of the Klopf model: that changes in the output should be correlated with changes in the inputs to determine whether learning should occur. When both the input and the output changes, the weights should increase. To allow non-zero ISI:s, each CS is assumed to leave an eligibility trace in the system. This is represented by the sum in equation (K2) which is used as a memory that extends backward in time.

The model is able to mirror a complex ISI-curve since an array of learning constants, c_i, are used to determine the shape of the curve. The constants are explicitly chosen to reflect the ISI-curve for classical delay conditioning (Klopf 1988). Figure 5.1 shows the simulation result for the same experiment as the TD model described above.
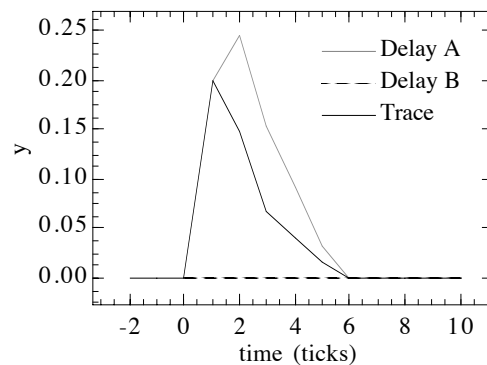

Figure 5.1. The ISI curve for the Klopf model.

Although the Klopf model has the ability to produce a range of ISI-curves depending on the learning rate constants, it has almost no predictive value in this area since the curve is essentially rigged by the constants c_i. A further limitation is the fixed length memory buffer used for the eligibility trace. The use of a memory buffer limits the temporal resolution of the model as well as the length of the ISI.

The model also shows the initially accelerated S-shapes learning curve that is observed in animals. This is the result of the multiplication of w_i in the eligibility trace. This makes learning faster when the weights are larger.

This unusual aspect of the learning equation also results in reacquisition effects. Since the weights are larger after extinction than before any learning has taken place, learning will be faster in the second acquisition phase. Although Klopf (1988) presents a simulation where the reacquisition effect is clearly visible, this effect only occurs for the *first* reacquisition. All reacquisitions after the first one are identical for the parameters used in Klopf (1988). See figure 5.2. This was not illustrated in the original article. Also note that the S-shaped learning curve disappears for the reacquisitions.

The reacquisition effect depends on the learning rule where the weight change depends on the magnitude of the weight. A larger weight will change more than a smaller one for the same reinforcement signal. This aspect of the model is also the reason for the initially accelerated S-curve in acquisition.

Klopf (1988) reports simulation of secondary conditioning which were reproduced by our simulator. It is interesting to note that this experiment results in the inhibitory effect illustrated above when run with the TD model. The Klopf model also handles blocking in an satisfactory manner.

Also facilitation by an intermittent stimulus is modelled correctly. This effect is both visible at an ISI smaller than the length of the eligibility trace, and at larger ISI where the mechanism if effect extends the length of the fixed memory buffer.

6 THE BALKENIUS MODEL

Like the Klopf model, the model presented in Balkenius (1995, 1996, 1998) is based on a neural interpretation of the conditioning mechanism. But contrary to the Klopf model, the Balkenius model uses a network rather than a single node. Like the Klopf model, the Balkenius model separates inhibitory and excitatory learning, but this is made explicit in the formulation of the model. Here, only the positive side of the network is described. The equations for the negative side are identical. The output of the model is given by,

$$CR(t) = x^+(t) - x^-(t), \qquad (B1)$$

$$x^+(t) = \sum_{i=1}^{n} \sum_{j=0}^{\tau} w_{ij}^+(t) CS_i(t-j). \qquad (B2)$$

The extra index j in equation (B2) corresponds to a tapped delay-line for each CS with length τ. This approach differs from the eligibility traces used in the other models but have a similar role. The representation is usually called a multiple-element stimulus trace and has the advantage that it can support more complex associations that a single eligibility trace. On the other hand, it requires many more variables since the number of weights must equal the length of the stimulus trace.

The reinforcement, that is, the weight change, is calculated as,

$$R^+ = \left[US + \delta\left(\Delta x_{t+1}^+ - \Delta x_{t+1}^-\right) - \left(\Delta x_t^+ - \Delta x_t^-\right) \right]^+. \quad (B3)$$

In this equation, δ is the discount factor which is responsible for secondary conditioning. This equation can be compared to equation (TD1) for the temporal difference model. In that model, the absolute values are used in the learning equation. In this model, it is instead the *changes* that contribute to the reinforcement. This is consistent with the idea that the model tries to predict the *level* of the US rather than the integral over it as the TD model.

Like in the Klopf model, changes in the US level are correlated with changes in the CS levels and during conditioning, the weights change according to the equation,

$$\Delta w_{ij}^+(t) = \gamma^+ R^+ \left[\Delta CS_i(t-j) \right]^+, \qquad (B4)$$

where γ^+ is the learning rate.

The Balkenius model differs from the other models in the way the ISI effect is modelled. In the TD and Klopf model, it is the result of the eligibility trace. In this model, it results from secondary conditioning within the stimulus trace. Figure 6.1 shows the ISI curve for the model with the discount factor of $\delta = 0.90$. The curve directly reflects the discount factor since the asymptotic value of the weights converges to δ^{i-1} for an inter-stimulus interval, $i \geq 0$.

A surprising effect of the stimulus representation is that the model predicts the S-shaped learning curves for ISI:s larger than 1 tick. Figure 6.2 shows three simulations with different ISI:s. For an ISI of 1 tick, the learning curve is similar to that of the TD model, but for a larger ISI, the curve is initially accelerating, giving rise to the characteristic S-shape.
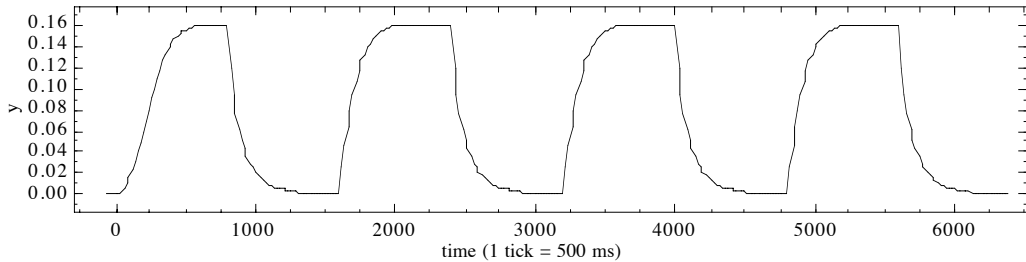
Figure 5.2. The reacquisition effect in the Klopf model. The reacquisition effect only appears the first time.
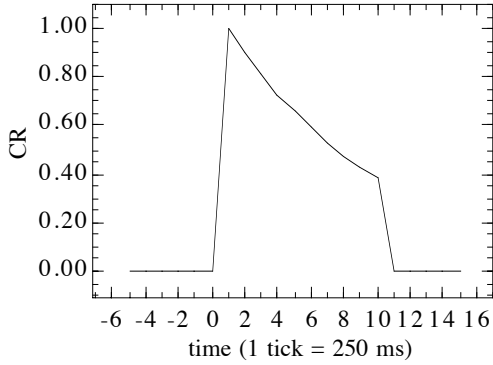
7

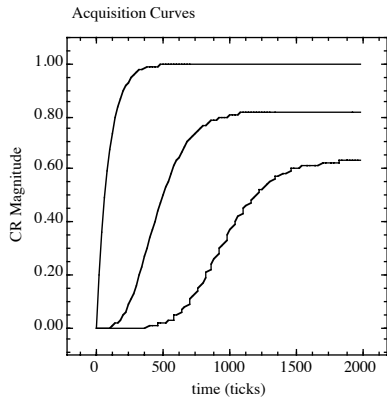**Figure 6.1.** The ISI-curve for the Balkenius model for δ = 0.9.

Figure 6.2. The shape of the learning curve for the Balkenius model depends on the ISI. For an ISI larger than one, the curve shows the characteristic S-shape. An ISI of 1, 5, and 10 ticks are shown in the figure.

This explanation of the S-shaped curve is radically different from the one offered by the Klopf model. In the Klopf model, the learning equation itself is constructed to give the S-shaped acquisition curve, while in the Balkenius model, it is the result of the mechanism for secondary conditioning given by equation (B3). For an ISI of 2 ticks, the initial acquisition curve is quadratic, for an ISI of 3 ticks, it is cubic, and so forth.

Secondary conditioning is also handled by the model. In fact, the Balkenius model is the only model of the three to give robust secondary conditioning on both the experiment described in Klopf (1988) and the one in Sutton and Barto (1990). The TD model did not do well on the Klopf experiment, and the Klopf model was not able to handle the experiment used for the TD model. This however, is done at the cost of ignoring the difference between trace and delay conditioning.

Like the other two models, blocking is handled without problems. It appears that since Rescorla and Wagner (1972), blocking is the first experiment to be tested for any model of conditioning.

A shortcoming of the model is that it is not able to model reacquisition effects. Like in the Klopf model, the weights reflect the fact that earlier conditioning has occurred, but this is not utilized in the learning equation.

A further problem with the Balkenius model is that it does not distinguish between delay and trace

conditioning. Since only positive changes in the CS is used in the model, trace and delay conditioning will appear identical to the learning mechanism. Empirical data suggests that this should not be the case since delay conditioning usually results in faster and stronger conditioning.

Finally, it is possible to raise the same objection to the multiple element stimulus trace in this model as to the eligibility trace in the Klopf model: it limits the temporal resolution and sets a fixed length on the memory for passed events.
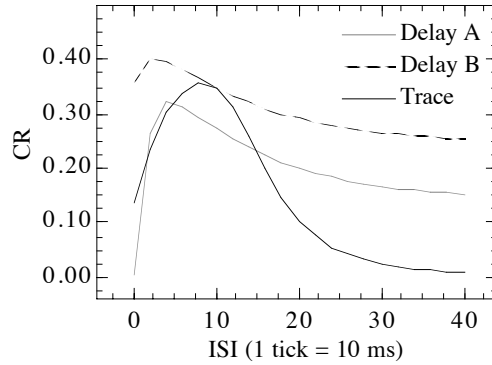

Figure 7.1. The ISI-curves for the SD-model.

7 THE SCHMAJUK-DiCARLO MODEL

The Schmajuk-DiCarlo (SD) model was introduced in 1992 and was shown to model a number of classical conditioning phenomena (Schmajuk and DiCarlo 1992). Especially interesting is its ability to model the effect of various types of configurational stimuli and the effects of hippocampal lesion on conditioning. Here, however, we will only investigate the more fundamental abilities of the model. The complete characterization of the model can be found in Schmajuk and DiCarlo (1992) and we will only describe the equations responsible for the dynamics found in the simulations. A stimulus, CS_i, gives rise to a short-term memory trace X_i that is described by,

$$\frac{dX_i}{dt} = -K_1 X_i + K_2\left(K_3 - X_i\right)CS_i. \qquad (SD1)$$

K_1, K_2 and K_3 are constants that determine the passive decay of the trace, the rate of increase, and the maximum level of the trace, respectively. The associative strength, VS_i changes according to the equation,

$$\frac{dVS_i}{dt} = K_5 K_6\left(1 - |VS_i|\right)EO, \qquad (SD2)$$

where K_5 is an output level constant and K_6 is the learning rate constant. EO describes the error in the prediction of the model. This error is calculated as the difference between the US level and the sum of all stimulus traces multiplied with their respective weights,

$$EO = US - \sum_i K_5 X_i VS_i. \qquad (SD3)$$

In the simulations, we followed the complete description of the model given in Schmajuk and

DiCarlo (1992), and the equations above should serve only as an indication of the dynamics of the model and not as a complete description.

The ISI curves for the model are shown in figure 7.1. Together with the TD model, the SD model is the only one to model the difference between the trace and delay conditioning in a qualitatively correct way.

The SD model also correctly models blocking and conditioned inhibition. Of the models we tested, the SD model is the only one to show a reacquisition effect that increases with each repeated relearning. This is shown in figure 7.2 where four acquisition and extinction phases are presented.

A serious difficulty for the SD model is that it does not allow for secondary conditioning. This is a direct consequence of equation (SD3). Since the change in associative strength depends on the difference between the US and the aggregate prediction given by the sum in equation (SD3), there is no room for secondary conditioning.

It is interesting to see that the SD model is able to model facilitation but not secondary conditioning. This implies that secondary conditioning is not necessarily required for facilitation.

8 DISCUSSION

The models in this study have different objectives and it is thus not surprising that the performance differs dramatically for various experiments. None of the models claim to provide a model of all of classical conditioning phenomena; indeed, they are intended to provide an explanation of only a few specific experimental situations. The experiments themselves were chosen for their appearance as examples in the texts describing the various models, or their widespread use as examples of classical conditioning.

The results should thus not be interpreted as a ranking of the models, but instead be seen as an aid to understanding of the strengths and weaknesses of each model, and to gain a better understanding of the field of conditioning.

Conspiciously absent are statistical models and experiments such as Gallistel (1992). Nor have we included attentional models (Grossberg 1975). We intend to extend these simulations with many more models and experiments; the present study can rightly be seen as a pilot study in preparation for a much more complete work.
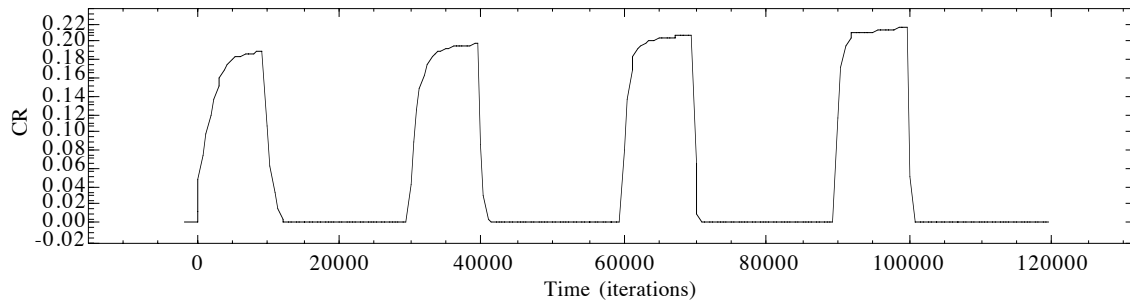

Figure 7.2. The reacquisition effects in the SD-model. This is the only model where each subsequent acquisition is faster than the previous one.

	SB	TD	Klopf	Balken-ius	SD
Trace Conditioning	+	+	+	+	+
Delay Conditioning	-	+	+-	+	+
ISI-curve	-	+	+-	+-	+
S-shaped Acquisition	-	-	+	+	-
Extinction	+	+	+	+	+
Reacquisition	-	-	+-	-	+
Blocking	+	+	+	+	+
Secondary Conditioning	+-	+-	+	+	-
Spontaneous recovery	-	-	-	-	-
Conditioned Inhibition	+	+	+	+	+
Facilitation	+	+	+	+	+

Table 8.1. Summary of the study. + indicates that the model handles this effect, - that it does not, and +- that it handles it in some cases. This table should not be taken as a ranking of the models.

REFERENCES

Balkenius, C. (1995). *Natural intelligence in artificial creatures.* Lund University Cognitive Studies 37.

Balkenius, C. (1996). Generalization in instrumental learning. In Maes, P., Mataric, M. J., Meyer, J.-A., Pollack, J., Wilson, S. W. (Eds.) *From Animals to Animats 4: Proceedings of the Fourth International Conference on Simulation of Adaptive Behavior* (pp. 305–314). Cambridge, MA: The MIT Press/Bradford Books.

Balkenius, C. (1998). A neural network model of classical conditioning I: the dynamics of learning. *in preparation.*

Blazis, D. E. J., Desmond, J. E., Moore, J. W., Berthier, N. E. (1986). Simulation of the classically conditioned nictitating response of a neuron-like adaptive element: a real-time variant of the Sutton-Barto model. *Proceedings of the Eigth Annual Conference of the Cognitive Science Society* . : Erlbaum.

Desmond, J. E. (1990). Temporally adaptive responses in neural models: the stimulus trace. In Gabriel, M., Moore, J. (Eds.) *Learning and computational neuroscience: foundations of adaptive networks* (pp. 421-456). Cambridge, MA: MIT Press.

Gallistel, C. R. (1992). Classical conditioning as an adaptive specialization: a computational model. In Medin, D. L. (Ed.) *The psychology of learning and motivation: advances in research and theory* (pp. 35–67). New York: Academic Press.

Garcia, J., Koelling, R. A. (1966). Relation of cue to consequences in avoidance learning. *Psychonomic Science,* 4, 123-124.

Grossberg, S. (1975). A neural model of attention, reinforcement, and discrimination learning. *International Review of Neurobiology,* 18, 263–327.

Kamin, L. J. (1968). Attention–like processes in classical conditioning. In Jones, M. R. (Ed.) *Miami symposium on the prediction of behavior: aversive stimulation* (pp. 9–31). Miami: University of Miami Press.

Klopf, A. H. (1982). *The hedonistic neuron: a theory of memory, learning and intelligence.* Washington, D. C.: Hemisphere.

Klopf, A. H. (1988). A neuronal model of classical conditioning. 16, 2, 85–125.

Klopf, A. H., Morgan, J. S., Weaver, S. E. (1993). A hierarchical network of control systems that learn: modeling nervous system function during classical and instrumental conditioning. 1, 3, 263–319.

Pavlov, I. P. (1927). *Conditioned reflexes.* Oxford: Oxford University Press.

Rescorla, R. A. (1968). Probability of shock in the presence and absence of CS in fear conditioning. *Journal of comparative physiological psychology,* 66, 1–5.

Rescorla, R. A., Wagner, A. R. (1972). A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In Black, A. H., Prokasy, W. F. (Eds.) *Classical conditioning II: current research and theory* (pp. 64–99). New York: Appleton-Century-Crofts.

Schmajuk, N. A., DiCarlo, J. J. (1992). Stimulis configuration, classical conditioning, and hippocampal function. *Psychological Review,* 99, 268–305.

Schneiderman, N. (1966). Interstimulus interval function of the nictitating membrane response of the rabbit under delay versus trace conditioning. *Journal of comparative and physiological psychology,* 62, 397-402.

Schneiderman, N., Fuentes, I., Gormezano, I. (1962). Acquisition and extinction of the classicaly conditioned eyelid response in the albino rabbit. *Science,* 136, 650–652.

Schneiderman, N., Gormezano, I. (1964). Conditioning of the nictitating membrane of the rabbit as a function as the CS–US interval. *Journal of comparative and physiological psychology,* 57, 188-195.

Smith, M. C., Coleman, S. R., Gormezano, I. (1969). Classical conditioning of the rabbit's nictitating membrane response at backward, simultaneous, and forward CS-US interval. 69, 226-31.

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. (1987). A temporal-difference model of classical conditioning. *Proceedings of the ninth conference of the cognitive science society* . : Erlbaum.

Sutton, R. S., Barto, A. G. (1990). Time-derivative models of Pavlovian reinforcement. In Gabriel, M., Moore, J. (Eds.) *Learning and computational neuroscience: foundations of adaptive networks* (pp. 497-538). Cambridge, MA: MIT Press.