Division of Informatics, University of Edinburgh

Institute of Perception, Action and Behaviour

Simulating Classical Conditioning using a Neuro-Connector Net

by

Bridget Hallam

Informatics Research Report EDI-INF-RR-0008

Division of Informatics **March 2000**
http://www.informatics.ed.ac.uk/

Simulating Classical Conditioning using a Neuro-Connector Net

Bridget Hallam

Informatics Research Report EDI-INF-RR-0008

DIVISION *of* INFORMATICS
Institute of Perception, Action and Behaviour

March 2000

Submitted to SAB 2000

Abstract :

Halperin's Neuro-Connector model (1990) has been implemented and used to replicate classical conditioning phenomena. This paper describes two of these experiments, attempts to replicate the effects of pre-exposure to CS and US, and the effect of partial reinforcement. In both cases similarity to animal results exists but has limitations. This indicates that the model may be correct in some of its assumptions, but there is definite scope for improvement.

Keywords : biological modelling, classical conditioning, animal conditioning, Halperin, Neuro-Connector

Simulating Classical Conditioning using a Neuro-Connector Net

Bridget Hallam
Division of Informatics,
University of Edinburgh,
5 Forrest Hill, Edinburgh,
EH1 2QL, UK
<bridget@dai.ed.ac.uk>

Abstract

Halperin's Neuro-Connector model [Halperin, 1990] has been implemented and used to replicate classical conditioning phenomena. This paper describes two of these experiments, attempts to replicate the effects of pre-exposure to CS and US, and the effect of partial reinforcement. In both cases similarity to animal results exists but has limitations. This indicates that the model may be correct in some of its assumptions, but there is definite scope for improvement.

1. Introduction

One of the most important adaptive features performed by any animal or animat is the ability to learn cause-and-effect. In this way the agent concerned can control some enviromental aspects through its own behaviour, and exploit or mitigate the effect of external influences.

Halperin's Neuro-Connector model [Halperin, 1990] is an unusual neural model which claims that timing coincidences drive learning, without reasoning about cause-and-effect. This model appeared promising for use on a mobile robot partly because of this claim, and also because the model uses real time, looked well specified, and had predicted a new form of conditioning which had subsequently been demonstrated in fish [Halperin and Dunham, 1992]. More investigation was indicated.

The example paradigm for cause-and-effect learning is classical or Pavlovian conditioning. Classical conditioning experiments are normally designed so that a previously motivationally neutral stimulus precedes a significant event. The vast body of literature describing various species and genera learning this sequencing gives remarkably similar results once allowances for differences in sensors and actuators are taken into account. This similarity implies that the vastly different hardware involved in computer or robot simulation of conditioning may not be significant, and that an accurate implementation of the mechanism underlying animal conditioning should produce conditioning phenomena in simulation also.

Note that Halperin's model was not designed to account for conditioning data, so any success in this area is extremely significant. An analysis of the ability of the Neuro-Connector model to replicate animal conditioning phenomena in simulation is given in Hallam [2000]. The main conclusions were that the model had both promise and problems. This paper aims to illustrate both, and to indicate that a superficial consideration of results may produce an over-inflated assessment of their worth.

This paper starts by outlining the major features of a Neuro-Connector net. The basics of animal classical conditioning are then outlined and replicated in simulation. Two more complex classical conditioning phenomena – stimulus pre-exposure and partial reinforcement effects – are then considered in more detail.

2. Major Neuro-Connector Net Features

The Neuro-Connector model of learning and motivation is presented in Halperin [1990] and Hallam et al. [1994]. It is described in more cognitive terms in Halperin [1995] and in more mathematical terms in Hallam et al. [1997]. The limitations of the specifications in these publications, and some of the consequences of the various possible alternative in-fills, are described in Hallam [2000]. Here, the main features significant for the reproduction of conditioning phenomena are summarised. For more details see the texts cited above.

The net comprises sensory (S), releaser (R), and behaviour (B) neurons, as illustrated in figure 1. A feedback loop between individual Rs and Bs allows behavioural persistence. Competition between behaviours prevents conflicting motor signals. Learning takes place in unidirectional $S \rightarrow R$ synapses.

$S \rightarrow R$ synapse weights are adjusted according to their *offset* (finishing firing) times. The qualitatively different cases are illustrated in figure 2. The actual or observed difference in offset times t_{obs} is compared with an expec-

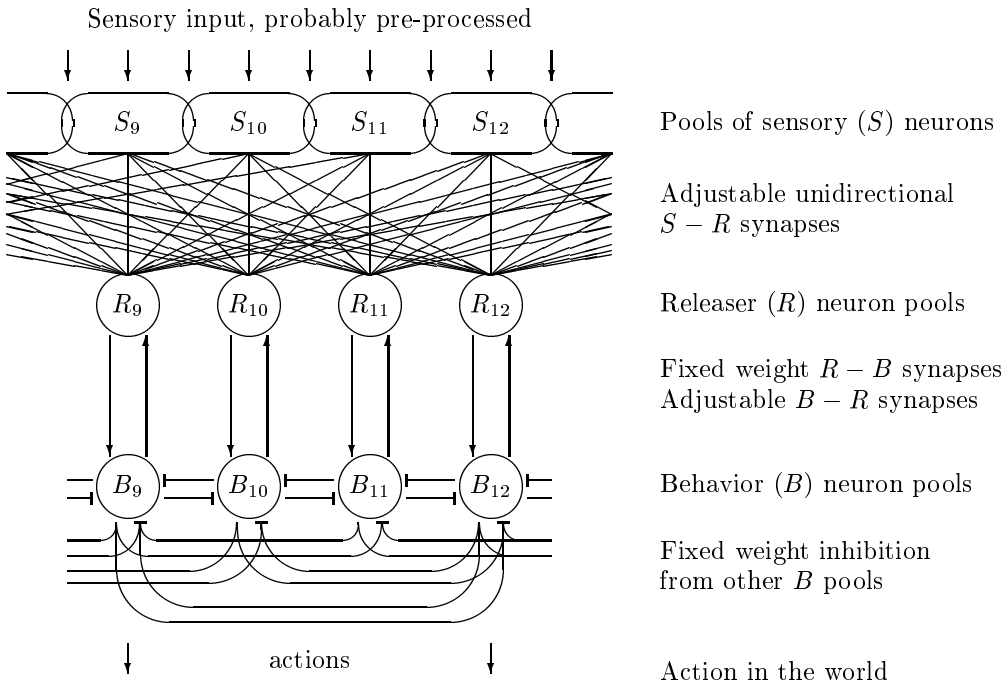

Figure 1: Part of a Neuro-Connector Net.

ted time t_{exp}, which is a synaptic parameter. $S \rightarrow R$ synapse weights increase only if R finishes firing within a small time window around the 'expected' time after S finishes.

Synapse weights decrease if the observed time t_{obs} between R and S offset differs too much from t_{exp}, this decrease being most extreme if either only S fired or else t_{obs} is only just outside the strengthening window shown in figure 2.

Weights increase if firing is correlated as in rule 1. Weights decrease if firing is not correlated as in rule 3 (only S fired) and rule 5 (both fired but uncorrelated). Weights hardly change if only R fires (rule 4), so that multiple stimuli can cause the same response. The most controversial of the rules is rule2, which says that differences in neural onset times are not significant – synapse weight increases if offset times are correlated without reference to previous firing history.

There are many good features supporting the biological plausibility of this model, but two major problems exist: neurons are binary, and there is no inhibition except between B neurons. It also seems [Bitterman, 1975] that both neural onset and offset times contribute to ease of learning in reality.

3. Basic Classical Conditioning

A simplified description of the most basic form of classical conditioning is as follows. The experimenter uses two stimuli, one which produces an innate response and one which produces little response – at most some small orienting response. The first is known as the unconditioned stimulus (US) and the behaviour which it produces is known as the unconditioned response (UR). The second is known as the conditioned stimulus (CS). It is normally shown just before the unconditioned stimulus. After a few trials the agent responds to the CS as if it has learned the CS as a predictor of the US, as if it is expecting the US. This learned response to the CS is called the conditioned response (CR). After showing that the agent can reliably produce the CR, the US is often discontinued. Under these conditions the CR disappears or 'extinguishes' after a few trials.

4. Conditioning a Neuro-Connector Net

To do conditioning in a Neuro-Connector net requires certain assumptions apart from those explicit in the model description. First, that UR and CR are not produced by firing in the same neuron, but are caused by different neurons. This assumption is largely accepted by the conditioning community, behavioural differences in CR and UR having been demonstrated by *e.g.* Zener [1937] and Spence and Ross [1959].

The second assumption is that S_{US} excites various R neurons, not just R_{UR}. Since weight in a Neuro-Connector synapse can only increase when both S and R have fired recently, and S_{CS} is initially insufficient to cause firing in R_{CR} itself, R_{CR} firing must be triggered by S_{US} since S_{US} is the only S neuron guaranteed

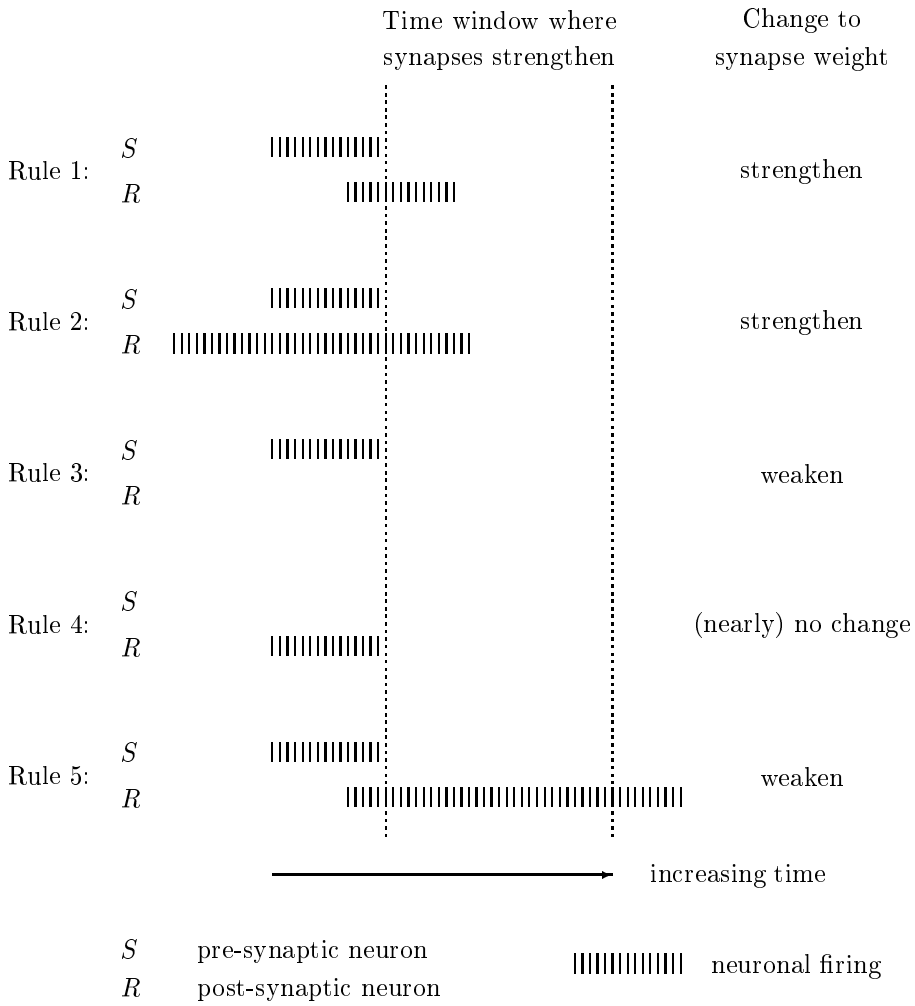

Figure 2: Qualitative Weight Adjustments from Halperin [1990].

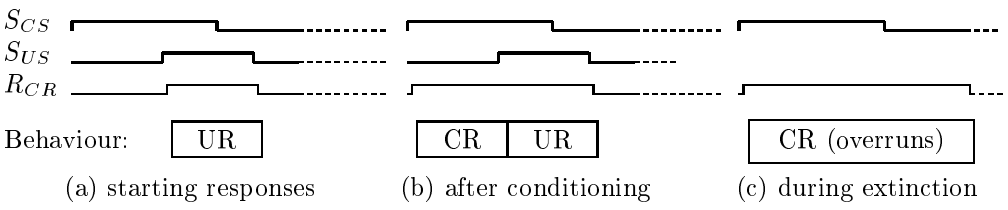

Figure 3: Stimulus Timings for the Classical Conditioning Experiment.

to be firing. This 'potentiation' of R neurons by innately significant stimuli is probably reasonable. The CR will still not be performed since the UR is a more important behaviour and only one behaviour from a system is chosen at any one time.

The third assumption is that R_{CR} firing continues past the time when it would normally have finished if the UR does not appear, because otherwise extinction is not obtainable.

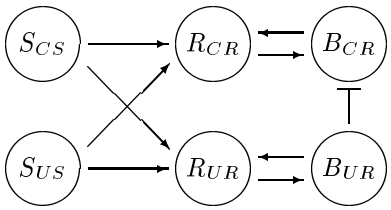

Figure 4: A Small Neuro-Connector Net.

Method The simplest Neuro-Connector net capable of demonstrating classical conditioning is illustrated in figure 4. S_{US} detects the US, and is given high-weight synapses to both Rs, easily sufficient to cause both to fire. S_{CS} detects the CS, and is given low-weight synapses to both Rs. B_{UR} inhibits B_{CR}, so when both Bs have above-threshold input only the UR appears.

The significant neural firings during conditioning, maintenance, and extinction are illustrated in figure 3. Stimulus presentations can be seen by the S_{CS} and S_{US} lines. The firing of R_{CR}, which starts by 'following' S_{US} and learns to also follow S_{CS}, is also shown. During acquisition and maintenance of conditioning S_{CS} and R_{CR} offset times stay constant at values giving a t_{obs} which is designed to be close to t_{exp}, allowing synaptic strengthening. During extinction R_{CR} is held on too long by B_{CR} (firing not shown), so R_{CR} offset time increases, t_{obs} no longer correlates with t_{exp} and synaptic weakening occurs as desired.

Results $S_{CS} \rightarrow R_{CR}$ weight (solid line) and B_{CR} latency (time between S_{CS} onset and B_{CR} onset; dashed line) are given in figure 5. After only five presentations the CR appeared with a latency of 0.33 time units. This latency decreased to a minimum of 0.16. The appearance of the CR after the CS and before the US demonstrates classical conditioning. Extinction is also shown, since the CR fails to appear after thirteen CS-alone presentations.

Weight changes show the same thing in more detail. $S_{CS} \rightarrow R_{CR}$ weight increased from 0.001 to 0.998 during acquisition as shown in figure 5 (solid line). During the 30 extinction trials it fell back down to 0.007 and was still falling. The weight passed the R firing threshold of 0.4 after only five acquisition presentations and, in the opposite direction, after thirteen extinction presentations.

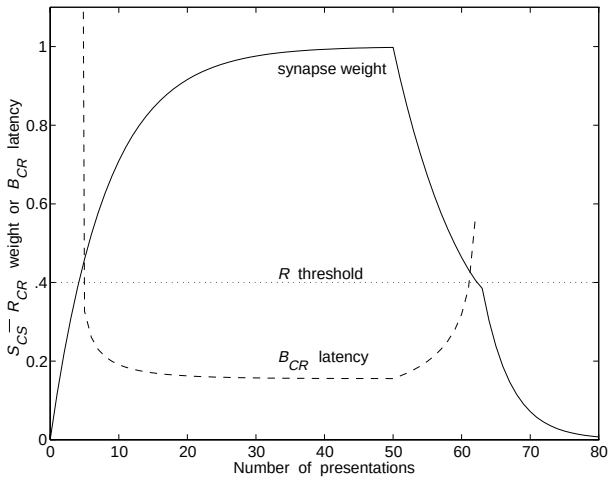

Figure 5: Acquisition and Extinction of Classical Conditioning shown through the change in $S_{CS} \rightarrow R_{CR}$ weight and the change in CR latency.

Discussion Classical conditioning is shown, since a CR develops, remains stable during maintenance trials, then disappears during extinction.

The rate of learning is very fast in that only five presentations were required for acquisition and thirteen for extinction. This is similar to that obtained under optimal conditions in animal experiments and much faster than learning commonly achieved by artificial neural nets. It can be varied to some extent by changing t_{exp} or affecting t_{obs} through changing neural timings.

One difference from animal results is in the details of behavioural latency changes. The CR appears at 0.33 time units after S_{CS} onset then drifts closer until it starts 0.16 after S_{CS}. This drift towards CS onset happens whatever the time difference between CS and US onsets and is in direct contrast to many animal results where the latency of the CR tends to drift towards **US** onset until the response reliably occurs just before the US [Ellison, 1964].

5. The Effect of Pre-Exposure

For animals, pre-exposure to either CS or US 'slows' conditioning. In other words, more presentations of the stimulus sequence (CS–US) are required before learning criteria are met if the animals have seen either CS alone or US alone before the conditioning presentations start than if both are novel stimuli. In simulation, pre-exposure to the CS can slow conditioning but pre-exposure to the US has no effect.

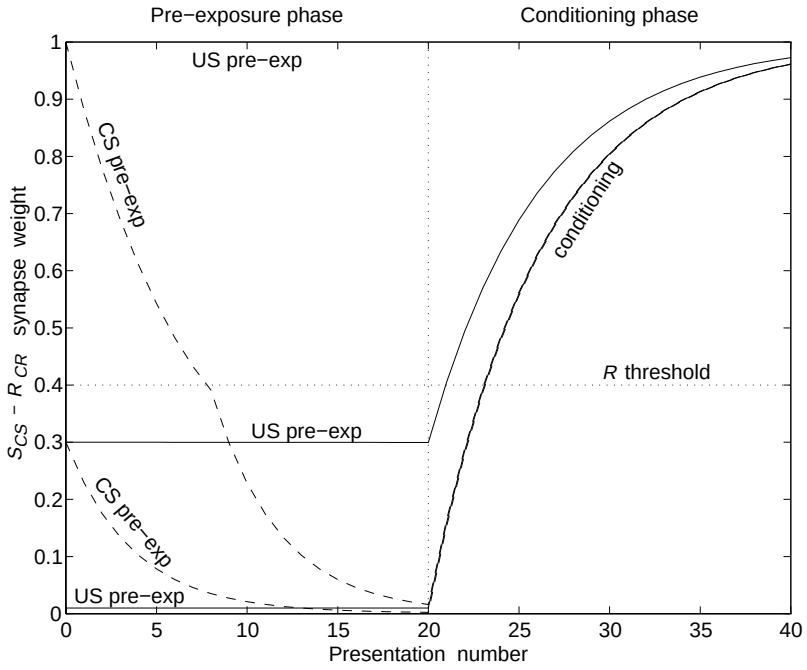

Figure 6: Pre-exposure to CS and US.

Method Various nets were initialised and run as in the basic conditioning experiment described above, except that the starting weight for $S_{CS} \rightarrow R_{CR}$ was set differently. Some nets were given twenty presentations of the CS alone before conditioning, some twenty presentations of the US. A control set were given neither.

Results Figure 6 shows selected combined results. In all cases $S_{CS} \rightarrow R_{CR}$ weights decrease during CS pre-exposure (dashed lines) but not during US pre-exposure (solid lines). Thus pre-exposure to the CS slows conditioning if synapse weights would not normally start at minimum, but US pre-exposure has no effect. Not visible in figure 6 is the control line (no pre-exposure) which starts at presentation 20 and overlays the other conditioning lines exactly, showing that the underlying rate of conditioning is unaffected by pre-exposure to either CS or US.

Discussion The reasons for the difference in pre-exposure effect are found in the intuitive weight-change rules given for the model. During CS pre-exposure S_{CS} is on without R_{CR} which is always a weight-reducing situation. During US pre-exposure S_{US} firing will trigger both R neurons but the CS is not seen. R firing without S causes almost no weight change (rule 4), so the synapses with S_{CS} neurons are unaffected by US pre-exposure. The apparently faster conditioning of the middle 'US pre-exposure' line is due entirely to its higher weight when conditioning starts.

In fact, since the high-weight $S_{CS} \rightarrow R_{CR}$ was given a high enough weight to cause the CR, CS pre-exposure can clearly be seen to be an identical situation to extinction. With a $S_{CS} \rightarrow R_{CR}$ starting weight below R threshold pre-exposure causes the equivalent, behaviourally invisible, retardation of subsequent conditioning found in the animal literature. The animal result implies that, if the Neuro-Connector model is a correct model of animal learning, $S_{CS} \rightarrow R_{CR}$ weights for novel stimuli start somewhere considerably above minimum although below R threshold.

This is a strong claim which may be unlikely due to other characteristics of the model. The problem is that, as other stimuli are sensed, S_{CS} will sometimes fire. This is particularly likely with the distributed representation of stimuli suggested in Halperin [1990]. When S_{CS} fires, either R_{CR} will fire 'correctly' (*i.e.,* so that t_{obs} matches t_{exp}), else incorrectly, else not at all. If $S_{CS} \rightarrow R_{CR}$ weight is above threshold then correct firing is most likely and synapse weight is expected to rise. But if $S_{CS} \rightarrow R_{CR}$ weight is below threshold then R_{CR} is much less likely to fire and synapse weight is likely to fall. Thus high weights are likely to increase and low ones decrease – intermediate weights are unstable. This has wider implications for the model than upsetting the CS pre-exposure effects found here.

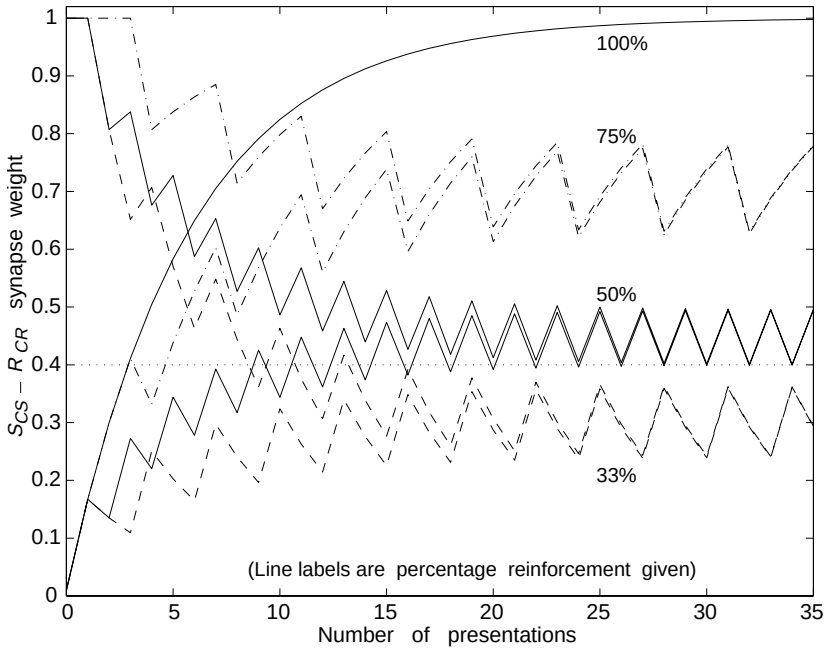

Figure 7: Varying the Reinforcement Schedule.

6. Partial Reinforcement

Instead of reinforcing every presentation of the CS it is possible to reinforce only a proportion of them. In animal classical conditioning, partial reinforcement causes a reduced rate of conditioning and sometimes an increased asymptotic level of response relative to continuous reinforcement.

Various fixed reinforcement ratios were tested, results from 1:3, 1:2, 3:4, and 1:1 are illustrated in figure 7. All experiments were run twice, with $S_{CS} \rightarrow R_{CR}$ starting at high weight as if pre-trained as well as at a low weight.

Results As expected, $S_{CS} \rightarrow R_{CR}$ weights oscillate, with weights increasing when the CS is reinforced and decreasing when it is not. For each reinforcement ratio there is a stable weight cycle which depends upon the ratio and not on the starting weight, shown by synapses with both the trained CS and the novel CS tending towards this cycle and oscillating together. Weights take approximately 25 presentations to achieve a stable cycle.

As the percentage reinforcement increases, so does the average $S_{CS} \rightarrow R_{CR}$ weight finally obtained. 75% reinforcement causes weight oscillations which are completely above the R threshold of 0.4; with 50% reinforcement synapse weight dips fractionally below R threshold on unreinforced trials; and with 33% reinforcement synapse weights are permanently below R threshold under the conditions used.

Discussion The stability of the weight cycles achieved is due to the exponential functions involved in the learning rule: weights jump further towards a more distant target so all weights tend towards the point where n jumps down can be cancelled by m jumps up, where $m : (n + m)$ is the ratio used. This applies wherever the weights start and whatever the reinforcement schedule, causing weights to tend exponentially towards this point and then do their oscillation there.

Enlarging the net to include other behaviours perturbs these otherwise stable oscillations. Whenever a S_{CS} neuron fires but another behaviour has higher priority (so our target response does not occur) the $S_{CS} \rightarrow R_{CR}$ weight will tend towards minimum if R_{CR} doesn't fire. Otherwise, if R_{CR} does follow S_{CS}, $S_{CS} \rightarrow R_{CR}$ weight will tend towards maximum.

The reduction in observed conditioning with reduction in percentage reinforcement is less extreme than found in animal experiments. This difference is further aggravated when the R has input from other S neurons, especially those which also respond to the CS. Since R input is additive each S needs to contribute only a small proportion of the weight needed to exceed R threshold. In the example above, 50% reinforcement caused $S_{CS} \rightarrow R_{CR}$ synapse weights to fall slightly below R threshold on alternate presentations. Under identical conditions but with an extra 0.01 of input for R_{CS}, the synapse weight never falls below R threshold. Put another way, with just two S_{CS}s contributing to R input, even 33% reinforcement gives constant CRs.

Obviously this precise result is an artefact of the R threshold used, but the principle is clear: those R neurons which have several active Ss attached are more likely to fire than those with fewer Ss active. The effect of extra low-weight synapses (whose Ss may be firing in response to completely different stimuli) can be imagined by adding a small amount to each weight shown, or alternatively by reducing R threshold by a little for each other S imagined firing. This may not be a sensible way to order R firing probability, especially since we want to learn precise releasing stimuli and penalise more 'general' S neurons.

Where extra reinforcements are given (*i.e.*, unsignalled US presentations) no weight change occurs on the US-only trials. If the x axis plotted is the number of reinforcements given rather than the number of presentations then the curves for 150% reinforcement and for 100% reinforcement overlie completely.

It can clearly be seen from this discussion that, although the synapse weights obtained reflect the percentage reinforcement given in a manner consistent with animal results, the observed behaviour is not so consistent. A better match would be obtained if the y axis of figure 7 could be labelled 'Response probability' instead of 'Synapse weight', but unfortunately this is not the case.

7. General Discussion

The implementation of the Neuro-Connector model used in this paper was able to demonstrate classical conditioning, in that a conditioned response appeared, was maintained, and extinguished appropriately. This is a significant success for the model since it was not designed to replicate conditioning phenomena but to explain the effect of social isolation on the aggressive display of Siamese fighting fish. However, Halperin claims her model as:

> 'a working hypothesis for the functional mechanism underlying much of vertebrate learning.' [Halperin, 1995, p493]

Any generalised model of vertebrate learning should, in my opinion, be able to replicate at least basic conditioning phenomena.

The effect of pre-exposure to the CS was also consistent with animal results, but only if the assumption were made that untrained synapse weights are significantly above minimum. This is unlikely given the expected effect of random events to force weights to an extreme, either minimum of maximum. So although the result obtained looks laudable, it is unlikely to be obtainable in a fully functioning complete system.

US pre-exposure has no effect on conditioning, in contrast to animal results. This is a direct consequence of rule 4 of figure 2, that R firing without S has no effect on synapse weight. This rule exists because of the necessity of several different stimuli being able to produce the same response, *e.g.* flee. When the US is seen S_{US} fires. We saw on page 4 that this needs to create firing in any potential R_{CR} neurons in order that conditioning be obtainable. During US pre-exposure, therefore, R_{CR} fires – but S_{CS} does not because the CS is not present. This is the case illustrated as rule 4 of figure 2.

The effect of partial reinforcement on synapse weights looks good, in that a lower percentage reinforcement caused a lower synapse weight. However, the implementation does not allow this graded effect to appear in the behaviour but produces binary output from the single, binary, B_{CR} neuron. Halperin specifies binary neurons, but existing in pools. These pools could produce graded output if inter- and intra- pool wiring were organised suitably.

Conclusion Halperin's model has mixed success when attempting to replicate animal conditioning phenomena. This mixed success is common even in models *designed* to reproduce conditioning phenomena [Balkenius and Morén, 1998], [Miller et al., 1995]. The degree of success of Halperin's model is impressive considering that the model was not designed to explain conditioning. Nevertheless, there are improvements which could be made.

A sister paper exploring the effect of timing variations on conditioning a simulated Neuro-Connector net is in production.

Acknowledgements

This work formed part of my doctoral work which was partially funded by SERC grant number 9130847X. The Department of Artificial Intelligence (now part of the Division of Informatics), University of Edinburgh, provided computing facilities.

References

C. Balkenius and J. Morén. Computational models of classical conditioning: A comparative study. Technical Report 62, Lund University Cognitive Studies, Sweden, 1998. A shorter version of this paper is available in 'From Animals to Animats 5', the proceedings of SAB98, published by MIT press.

M. E. Bitterman. The comparative analysis of learning. *Science*, 188:699–709, 1975.

G. D. Ellison. Differential salivary conditioning to traces. *Journal of Comparative and Physiological Psychology*, 57(3):373–380, 1964.

B. E. Hallam. *Simulating Animal Conditioning: Investigating Halperin's Neuro-Connector Model.* PhD

thesis, Artificial Intelligence, University of Edinburgh, 2000.

B. E. Hallam, J. C. T. Hallam, and G. M. Hayes. A dynamic net for robot control. In O. Omidvar and P. van der Smagt, editors, *Neural Systems for Robotics*, pages 227–269. Academic Press, San Diego, CA, 1997.

B. E. Hallam, J. R. P. Halperin, and J. C. T. Hallam. An ethological model for implementation in mobile robots. *Adaptive Behavior*, 3:51–79, 1994.

J. R. P. Halperin. *A Connectionist Neural Network Model of Aggression*. PhD thesis, Department of Ethology, Toronto University, Canada, 1990.

J. R. P. Halperin. Cognition and emotion in animals and machines. In H. Roitblat and J-A. Meyer, editors, *Comparative Approaches to Cognitive Science*, chapter 19, pages 465–499. A Bradford Book. MIT Press, Cambridge, MA, 1995.

J. R. P. Halperin and D. W. Dunham. Postponed conditioning: Testing a hypothesis about synaptic strengthening. *Adaptive Behavior*, 1(1):39–64, 1992.

R. R. Miller, R. C. Barnet, and N. J. Grahame. Assessment of the Rescorla–Wagner model. *Psychological Bulletin*, 117:363–386, 1995.

K. W. Spence and L. E. Ross. A methodological study of the form and latency of eyelid responses in conditioning. *Journal of Experimental Psychology*, 58:376–381, 1959.

K. Zener. The significance of behavior accompanying conditioned salivary secretion for theories of the conditioned response. *American Journal of Psychology*, 50:384–403, 1937.