A Neural Network Model of Scalar Timing

J. L. Shapiro
John Wearden
Manchester University, Manchester M13 9PL, U.K.

1 Introduction

Numerous experiments have investigated the ability of animals to estimate time intervals and wait appropriately for delayed rewards. A typical type of experiment is the peak procedure experiments. The animal is rewarded for the first response after a given time interval has elapsed. On some trials, however, no reward is given even when the animal responds appropriately. What happens in non-reward trials is that the animal typically will start responding at a certain time, will respond for a period, and then stop responding. Responses averaged over many trials, however, give a smooth curve. The highest response is at the time interval, and there is variation around this. This has been seen in many species, including rats, pigeons, turtles; humans will show similar results if the time intervals are short or if they are prevented from counting through distracting tasks. For reviews of interval timing phenomena, see [4] and [2].

A striking feature of the data on peak procedure and other experiments on time intervals is the so called "scalar property". In its weak form, this states that the standard deviation of the response time scales with the length of the time interval. In other words, the ratio of the standard deviation to the mean response time (the *coefficient of variation*) is a constant independent of the time interval. A stronger form of the scalar property is the following: when the response probability is multiplied by the time interval and is plotted against the relative time (time divided by the time interval), the data from different time intervals collapse on to one curve.

This strong form of the scalar property can be expressed mathematically as follows. Let T be the actual time since the start of the trial and $\tilde{T}$ be subjective time. Subjective time is the time duration which the animal perceives to have occurred (or appears to perceive judging from its behavior). The variation of $\tilde{T}$ for given T can be expressed as the conditional probability, $P(\tilde{T}|T)$. The fact that the data collapses implies this probability depends on T and $\tilde{T}$ in a special way,

$$P(\tilde{T}|T) \approx \frac{1}{T} P_{\mathrm{inv}} \left(\frac{\tilde{T}}{T} \right). \tag{1}$$

Here P_{inv} is the function onto which the data collapses. Thus, time acts as a scale factor. This is a strong and striking result.

A key question which remains unanswered is: what is the origin of the scalar property. Since the scalar property is ubiquitous across experiments and species, it may be revealing something fundamental about the nature of animals' internal clock. In this paper, a simple neural network model of an internal clock is described which gives rise to the scalar property. When coupled with reinforcement learning algorithms [8], like those applied to classical conditioning experiments (e.g. Sutton and Barto [7], and Moore et. al. [5]), it reproduces the results of peak procedure and other experiments.

2 A Connectionist Model of the Clock

We now consider a very simple connectionist model which reproduces the scalar property. It consists of a network of spiking neurons. The neurons are noisy and have linear response. The network encodes time as the quantity of activity, i.e. the number of spikes in the network. This grows over time.

The network consist of N identical neurons. Any neuron can potentially be connected to any other; the actual connectivity pattern is defined by a connection matrix $\mathbf{C}$, where C_{ij} is 1 if neuron j receives input from neuron i and 0 otherwise. Each node is connected to C other nodes on average. Time is divided into discrete units of size τ which are assumed short (10 mS, say). During a time interval, a neuron can emit one or more spikes. There is no limit to the spiking rate of these neurons, so saturation effects are ignored.

The neurons are noisy linear neurons. Let $x_i(t)$ denote the number of spikes emitted by neuron i between times t and $t+\tau$. The number of spikes arriving at that neuron (pre-synaptically) $h_i(t)$ is

$$h_i(t) = \sum_j C_{ji} x_j(t) + I_i(t), \tag{2}$$

where I_i is external input. The assumption is that the number spikes emitted by this neuron by time $t+\tau$ is drawn from a probability distribution with mean $\gamma h_i(t)$, with gain $\gamma < 1$. For example, if there is an independent probability γ of each spike crossing the synapse, then the distribution of spikes emitted by neuron i is binomial.

At each time-step, the number of spikes will grow due to the fan-out of the neurons; each neuron can excite several other neurons. At the same time, the number of spikes will shrink due to the fact that a spike invokes another spike with a probability less than 1. An essential assumption of this work is that these two processes balance each other.

$$C\gamma = 1. \tag{3}$$

This insures that the number of spikes internally excited remains constant over time on average.

Finally, in order for this network to act as an accumulator, it receives internal input during the time interval which is being perceived. The input is assumed statistically the same at each time-step.

This simple network gives rise to the scalar property in timing. Let $P(n|t)$ be the probability of network having total activity n at time t. What is proved is,

$$\lim_{t\to\infty} P(n|t) = \frac{1}{t} P_{\mathrm{inv}}(n/t), \tag{4}$$

with corrections of order $1/t$. Thus, if an interval of time is measured through the activity of this network, the perception of time will obey the scalar property equation (1). The invariant distribution is a gamma-function,

$$P_{\rm inv}(x) = \frac{\exp(-x/b)x^{a-1}}{b^a\Gamma(a)};\ a = \frac{2m_I}{C\sigma_\eta^2};\ b = \frac{C\sigma_\eta^2}{2\tau}, \tag{5}$$

where m_I is the mean external excitation per unit time and σ_η is the standard deviation of the neuron noise.

3 Learning Time Intervals

The above model represents a way of representing and measuring time. How would this be used to produce behavior? We have considered two models for learning time intervals. In the first, time is converted into a spatial code which are associated with time intervals. In the second, the accumulator output is used by the system directly. The spatially encoded model with be considered here. The details will be presented elsewhere [6].

Learning is done with a form of temporal difference learning. In this model, the accumulator network feeds into a set of nodes which encode the activity of the accumulator. Each of the spatial code nodes feed into response nodes which stimulate the responses. The connection between the the ith spatial node and the jth response node learns via a temporal difference learning rule.

This model has some similarities to that of Desmond and Moore [1]. The spatial nodes play the role of the tapped delay-line nodes in that model. However, here they are stimulated by the accumulator rather than each other, and they will follow a stochastic trajectory due to the fluctuating nature of the accumulator. Learning uses an eligibility trace for the spatial nodes and learning is via a simple TD rule. The model also as some similarities to the Grossberg and Merrill model [3]. The spatial nodes respond to different times and the response is associated with time via Hebbian association. However, the spatial nodes require no special tuning curves.

The model has been used to simulate peak procedure. In the simulations, the model is forced to respond for the first set of trials (50 trials in the simulations); otherwise the model would never respond. This could represent shaping in real experiments. After that the model learns using reward trials for an additional number of trials (100 trials in these simulations). The model is then tested on a number of non-reward trials. Figure 1 shows the average responses for non-reward trials, averaged over 100 runs. The results qualitatively match that of animal experiments; the model clearly exhibits the scalar property.

4 Discussion

Previous models of interval timing fail to explain its most striking feature — the collapse of the data when scaled by the time interval. We have presented a simple model of an accumulator clock based on spiking, noisy, linear neurons which produces this effect. The model is analytically tractable and has only three adjustable parameters. The major weakness of this model is that it requires fine-tuning of a pair of parameters, equation (3).

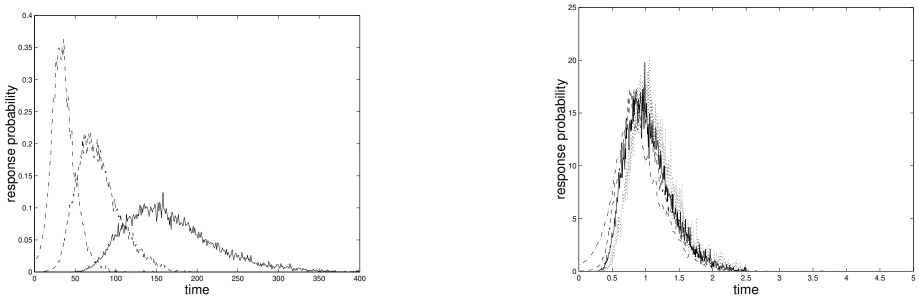

Figure 1: Left figure: Average response of the network for non-reward trials. 100 non-reward trials have been averaged. Right figure: The response probabilities scaled with the time interval and plotted against relative time. The strong scalar property is clearly observed. The time intervals are 40τ (dot-dashed line), 80τ (dashed line), and 160τ (solid line); with the addition of a time interval of 320τ in the right figure.

Once a scalar clock is produced, simple reinforcement learning can be used to associate the clock signal with appropriate responses. Simple temporal difference association between the intermediate nodes at reinforcement and an eligibility trace simulates peak procedure. Another set of experiments, bisection experiments, have also been modelled using this approach.

References

[1] John E. Desmond. In Michael Gabriel and John Moore, editors, *Learning and Computational Neuroscience: Foundations of Adaptive Networks*, A Bradford Book, pages 421–456. The MIT Press, 1990.

[2] John Gibbon and Russell M. Church. *Cognition*, 37:23–54, 1990.

[3] Stephen Grossberg and John W. L. Merrill. *Cognitive Brain Research*, 1:3–38, 1992.

[4] S. C. Hinton and W. H. Meck. In C. M. Bradshaw and E. Szadabi, editors, *Time and Behaviour: Psychological and Neurobehavioural Analyses.* Amsterdam: Elsevier Science, 1997.

[5] J. W. Moore, J. E. Desmond, and N. E. Berthier. *Biological Cybernetics*, 62:17–28, 1989.

[6] Jonathan Shapiro, John Wearden, and Rossano Barrone. Simple models of scalar timing. In preparation.

[7] Richard S. Sutton and Andrew G. Barto. In Michael Gabriel and John Moore, editors, ibid, pages 497–537.

[8] Richard S. Sutton and Andrew G. Barto. *Reinforcment Learning: An Introduction.* A Bradford Book. The MIT Press, 1998.