A Simple Model Exhibiting Scalar Timing

J. L. Shapiro, John Wearden, and Rossano Barone
Manchester University
Manchester, M13 9PL, U.K.
jls@cs.man.ac.uk

Abstract

Animal data on delayed-reward conditioning experiments shows a striking property — the data for different time intervals collapses into a single curve when the data is scaled by the time interval. This is called the scalar property of interval timing. Here a simple model of a neural clock is presented and shown to give rise to the scalar property. The model is an accumulator consisting of noisy, linear spiking neurons. It is analytically tractable and contains only three parameters. When coupled with simple reinforcement learning it simulates peak procedure experiments and temporal bisection.

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 called the peak procedure. In this, the animal is trained to respond after a given time interval has elapsed. There is no penalty for responding before the time has passed, but the animal only receives a reward at the first response after the time interval has passed. On most trials, the first response ends the trial.

On some trials, however, no reward is given even when the animal responds appropriately. This is to see when the animal stops responding. What typically happens in non-reward trials is that the animal starts responding at some time, responds for a period, and then stops responding. Responses averaged over many trials give a smooth curve. The highest response is at the time interval, and there is variation around this. Typical curves are illustrated in figure 1. 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 [8] and [6].

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 variation in the response of the animal, as measured by the standard deviation of the response time, scales with the length of the time interval. Since the mean response time also scales with the length of the time interval, the ratio of the standard deviation to the mean response time (the *coefficient of variation*) is a constant independent of the time interval. This is shown in the left side of figure 2. The right side of figure 2 shows a stronger form of the scalar property. 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.

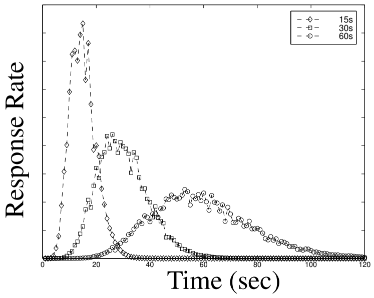

Figure 1: The typical appearance of the average response probability during non-reward trials in a peak procedure experiment as a function of time for three different time intervals (15, 30, and 60 seconds). This figure and following one are schematic views of the data in [5].

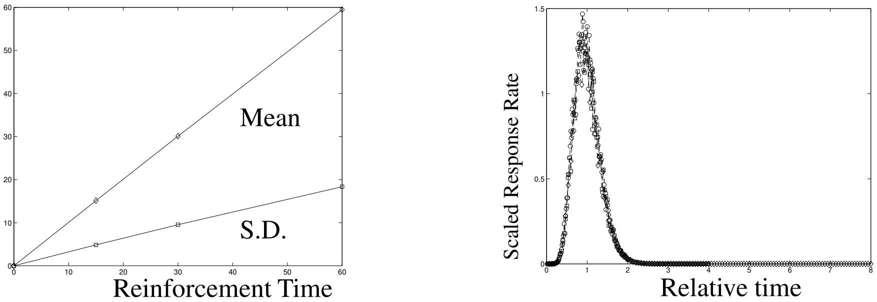

Figure 2: Two illustrations of the scalar property. The figure on the left shows that during non-reward trials the mean response time and its standard deviation are proportional to the time interval. The figure on the right illustrates that the data collapses when the response probability scaled with the time interval and plotted against relative time.

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 at least appears to perceive judging from its behaviour). The experiments show that $\tilde{T}$ varies for a given T. This variation can be expressed as a conditional probability, the probability of acting as though the time is $\tilde{T}$ given that the actual time is T, which is written $P(\tilde{T}|T)$. The fact that the data collapses implies this probability depends on T and $\tilde{T}$ in a special way,

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

Here P_{inv} is the function which describes the shape of the scaled curves. 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 clocks. But it has been difficult to think of processes which produce it. It is well known that any model based on the accumulation of independent errors, such as a clock with a variable pulse-rate, does not produce the scalar property. In such a model it would be the ratio of the *variance* to the mean response time which would be independent of the time interval (a consequence of the law of large numbers). In section 3, a simple stochastic process will be presented which gives rise to scalar timing.

In section 5, we show that reinforcement learning coupled to our model of scalar timing simulates experiments on peak procedure and another experiment called temporal bisection.

2. Previous Work

One way to generate the strong scalar property is to assume that there is a source of multiplicative noise. If one assumes that subjective time $\tilde{T}$ is related to actual time T through a noise source ξ

$$\tilde{T} = \xi T, \tag{2}$$

then equation (1) immediately follows, with P_{inv} equal to the noise distribution of ξ. A standard information processing model called Scalar Expectancy Theory [4] posits that multiplicative random variables intervene between the value of the count accumulated from a clock reinforcement time and retrieval time (see, for example [5] [6]). The clock is assumed to be an accurate time-keeper. The multiplicative noise is interpreted as motivational and attentional factors, and coding and decoding. It has been argued that it is possible to manipulate this noise source using drugs [9]. However, there does not appear to be any proposed mechanism for the noise, or why the accurate clock is not accessible. The scalar property is put in at the outset.

A connectionist model of scalar timing was proposed by Church and Broadbent [1]. This consists of a number of oscillators of differing frequencies, an associative memory to store time represented as the phases of the oscillators, and a comparator for decision making. The scalar property is built into the model through the assumption that the frequency of each oscillator is multiplied by a random variable generated from the same noise source. The model does simulate a number of interval procedures qualitatively [14], but does not provide quantitative predictions. A similar model used oscillators coupled to additive Gaussian noise in which the noise variance was proportional to the time interval [3]. This model described acquisition and particularly extinction between trials in rats, but again fails to explain the origin of the scalar property, as that is put into the model at the outset.

Another connectionist model was proposed by Grossberg and Merrill [7]. This model also consists of numerous clocks in the form of a number of neurons each of which is tuned to respond after a different time delay. This model also addresses the reinforcement learning problem. At reinforcement the response is associated with the clock nodes. Although the scalar property is not directly tested in this model, each clock node appears to have a tuning curve which has increasing width with increasing delay and the response appears to reproduce the scalar property qualitatively. This is a consequence of the dynamical equations used to described them. Grossberg argues that these equations describe plausible dynamics for real neurons. Thus, this model does seem to propose an explanation of the origin of the scalar property — it originates from the dynamics of neurons. However, the model does require many self-similar clock nodes, and contains a number of parameters which must be set to get an appropriate range of tuning curves of appropriate shapes.

A set of models which is very effective at modelling Pavlovian eye-blink experiments has been proposed by Desmond and Moore [10] based on the model of Sutton, Barto, and Desmond [11]. This model represents time using a tapped delay line; at each time-step, a different node in the delay line is activated. Time acts as one of the conditioned stimuli which using temporal difference (TD) reinforcement learning is associated with the response through the unconditioned stimulus. This models elucidates the reinforcement learning problem very clearly and simulates many experiments, but does not attempt to model the scalar property; time is represented accurately by the system.

3. A Simple Stochastic Network Which Produces the Scalar Property

We now consider a very simple connectionist model which reproduces the scalar property. It consists of a network of spiking neurons which are noisy and linear. 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. The connection strength is the same between all connected neurons and is taken to be 1. We assume that there is a characteristic time τ (10mS, say) required for a spike produced by a neuron to invoke a spike in a connected neuron. In principle, the spikes should be modelled as stochastic in time, e.g. via a Poisson process. However, we use strict, discrete units of size τ. There is no limit to the spiking rate of these neurons, so saturation effects are ignored.

The basic assumptions concerning the interactions between neurons are as follows,

Assumption 1 *There is a probability γ of a spike crossing a synapse,*

Assumption 2 *Spikes do not interact at synapses; they behave independently.*

Thus, these neurons act like noisy linear neurons. The expected number of spikes produced by a neuron is just γ times the number of pre-synaptic spikes. In other words, if $x_i(t)$ denotes the number of spikes produced by neuron i at time t, and $h_i(t)$ is the number of spikes feeding into that neuron,

$$h_i(t) = \sum_j C_{ji} x_j(t), \tag{3}$$

then $x_i(t+\tau)$ is drawn from a binomial distribution with mean $\gamma h_i(t)$ and variance $\gamma(1-\gamma)h_i(t)$.

Since the expected number of spikes produced by the neuron is less than the number exciting it, there is a loss of network activity when spikes cross synapses. However, the number of spikes also increases due to the fan-out of each neuron. Let C_i denote the fan-out of the ith node, $C_i = \sum_j C_{ij}$. To simplify the discussion, assume that all neurons have the same connectivity, $C_i = C$ for all i, (This is only necessary to hold on average and that is how the simulations in the next sections are done.) Under this assumption, the total activity of the network $n(t)$ can be expressed in terms of that in the previous time-step, $n(t-\tau)$, because $n(t)$ is drawn from a binomial distribution with mean $\gamma C n(t-\tau)$ and variance $\gamma(1-\gamma)Cn(t-\tau)$.

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. This is required to get the mean activity to be proportional to the time interval.

Assumption 3 *The amplification of neuron activity due to each neuron exciting multiple other neurons exactly balances the diminution of activity due to the finite probability a neuron exciting other neurons.*

In other words,

$$C\gamma = 1. \tag{4}$$

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

Assumption 4 *During a perceived time interval, the network receives constant or statistically stationary input.*

These assumptions modify the dynamics for those neurons connected to the inputs. Let $I(t)$ denote the number of spikes externally excited in the network at time t. The equation for the total neuron activity is,

$$n(t+\tau) = \sum_{\alpha=1}^{Cn(t)} \eta_\alpha + I(t). \tag{5}$$

where the η's are random variables — they are 1 with probability γ and 0 with probability $1-\gamma$ It is this equation which gives rise to the main features of scalar timing.

4. Discussion of the Model

This simple network is sufficient to give rise to scalar timing. Elsewhere [12] we show that for large time t, the network activation obeys the scalar property. Let $P(n|t)$ be the probability of network having total activity n at time t. The result is,

$$\lim_{t\to\infty} P(n|t) = \frac{1}{t} P_{\text{inv}}(n/t), \tag{6}$$

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 strong scalar property (equation 1).

For example, the coefficient of variation is

$$Q(t) = \frac{\sqrt{1-\gamma}}{\sqrt{2} m_I}\left[1 + \frac{\tau}{t}\left(1 + \frac{2\sigma_I^2}{1-\gamma}\right)\right] + O(\frac{\tau}{t})^2. \tag{7}$$

where m_I is the mean external excitation, and σ_I is the variance in the external excitation. This goes to a constant for large t. More generally, the scalar distribution, P_{inv} is a gamma-distribution with mean m_I/τ and variance $(1-\gamma)m_I/(2\tau^2)$. These two combinations of the three parameters of the model fully describe the system for large t.

5. Learning Behaviour

The above model represents a way of representing and measuring time. How would this be used to produce behaviour? Here a simple model is considered in which time is converted into a spatial code which is then associated with time intervals. Other methods of encoding time using this clock and more details will be presented elsewhere [12].

5.1 Learning Peak Procedure

In this model, the accumulator network feeds into a set of nodes which encode the activity of the accumulator. Each of these node is tuned to respond to a single value of accumulator activity. I.e. for any accumulator network output there is a single node which responds. These nodes for a spatial encoding of time.

For each possible response there is a response node. Each of the spatial code nodes feed into response nodes. The connection between the the ith spatial node and the jth response node is w_{ij}; it is these connections which learn to associate a response with an appropriate time delay. There is also a connection between the stimuli and the response nodes. It is the combined influence of the stimulus and the spatial nodes which cause the response node to fire. The association between stimulus and response is not modelled, but could follow a standard TD learning rule.

This model uses the TD learning equations of Desmond and Moore [2] as described in [11]. 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 association rather than a TD rule.

The model as 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 3 shows the average responses for non-reward trials, averaged over 100 runs. The figure also shows that the model produces the scalar property quite strongly.

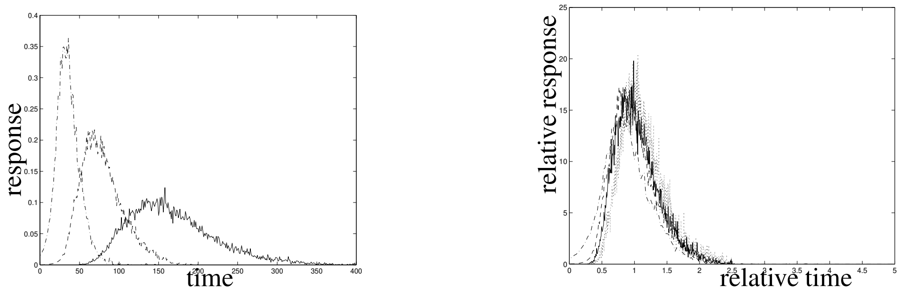

Figure 3: Left figure: Average response of the spatially encoded network for non-reward trials. The time intervals are 40τ (dot-dashed line), 80τ (dashed line), and 160τ (solid line). 100 non-reward trials have been averaged. Right Figure: The response probabilities scaled with the time interval and plotted against relative time. Same details as previous figure with addition of time interval of 320.

5.2 Learning Temporal Generalization

Another type of procedure which can be modelled is temporal bisection. In these experiments, a stimulus for a short duration S results in rewards for one action; the same stimulus for a longer duration L results in rewards for a different action. After an animal has been trained on the two intervals, intermediate intervals are used in order to see for what duration the animal produces either action with equal likelihood, the so-called bisection point. It has been observed that the bisection point is below the arithmetic mean and is closer to the geometric mean.

The model has two different response nodes, one trained on an interval of 40 time-steps and the other trained on longer intervals (60, 80, 100, 120, 160, and 320). The bisection point is interpreted as the point where the probabilities of producing each action are equal. In these simulations, there is nothing to stop the network from making both actions simultaneously or make no action. Thus, the two actions are independent, and the bisection point is essentially where curves in figure 3 cross. The results of this experiment are shown in figure 4. The bisection points are systematically below the geometric mean, although the difference is small. One way to estimate the bisection point is to take were the perceived time is the same. That is, the value of the network activity which is equally likely to be generated by time interval L as from time interval S. This is not the geometric mean, but is given by

$$\text{bisection point} = \frac{LS}{L-S}\log\frac{L}{S} \tag{8}$$

This is systematically less than the geometric mean, but the relative difference is less than 10% unless the long interval is more than 5 times greater the short one. The simulation results are consistent with this value.

6. 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. It is a simple, analytically tractable model. The parameters are: τ — the time for a spike on one neuron to excite spikes on connected neurons, m_I — the average number of spikes excited externally at each time step, and the variance of the branching process, which in this model is $1-\gamma$. A major weakness of this model is that it requires fine-tuning of a pair of parameters, equation (4), so that the expected number of spikes grows linearly during a measured interval.

Once a scalar clock is produced, simple reinforcement learning can be used to associate the clock signal with appropriate responses. An set of intermediate nodes which encodes time spatially was used here. Simple reinforcement learning between the intermediate nodes at reinforcement and an eligibility trace simulates peak procedure and bisection experiments, although the bisection point is below the geometric mean.

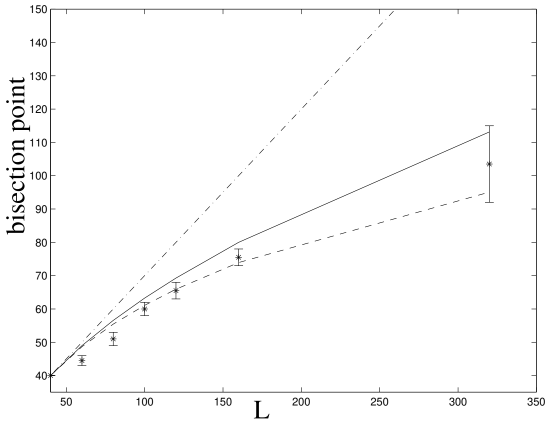

Figure 4: The results of bisection experiments. The short time interval is 40τ; the bisection point is shown for long intervals of 60τ, 80τ, 100τ, 120τ, 160τ, and 320τ. The dot-dashed line shows the arithmetic mean, the solid line shows the geometric mean, and the dashed line is the point where the clock output are as likely to have come from either interval, which is given by equation (8)

The next step is to draw connections to psychology and neurology. Because there are only a few parameters in this models, there will be relationships between measurable quantities. For example, it is not clear whether there is sufficient freedom to describe the range of pharmacological manipulations which have been carried out. Until we find what this relationships are in this model and test them experimentally, it will remain unclear how relevant this model is interval timing phenomena.

References

[1] R. Church and H. Broadbent. (1990). Alternative representations of time, number, and rate. *Cognition*, 37:55–81.

[2] John E. Desmond. (1990). Temporally adaptive responses in neural models: The stimulus trace. In Michael Gabriel and John Moore, editors, *Learning and Computational Neuroscience: Foundations of Adaptive Networks*, A Bradford Book, pages 421–456. The MIT Press.

[3] R. M. French and A. Ferrara. (1999). Modeling time perception in rats: Evidence for catastrophic interference in animal learning. In *Proceedings of the 21st Annual Conference of the Cognitive Science Conference*, pages 173–178. LEA.

[4] John Gibbon. (1977). Scalar expectancy theory and Weber's law in animal learning. *Psychology Review*, 84:279–325.

[5] John Gibbon. (1992) Ubiquity of scalar timing with a Poisson clock. *Journal of Mathematical Psychology*, 35:283–293. Theoretical Note.

[6] John Gibbon and Russell M. Church. (1990). Representation of time. *Cognition*, 37:23–54.

[7] Stephen Grossberg and John W. L. Merrill. (1992). A neural network model of adaptively timed reinforcement learning and hippocampal dynamics. *Cognitive Brain Research*, 1:3–38.

[8] S. C. Hinton and W. H. Meck. (1997). How time flies: Functional and neural mechansims of interval timing. In C. M. Bradshaw and E. Szadabi, editors, *Time and Behaviour: Psychological and Neurobehavioural Analyses*. Amsterdam: Elsevier Science.

[9] W. H. Meck. (1996). Neuropharmacology of timing and time perception. *Brain Research. Cognitive Brain Research*, 3:227–242.

[10] J. W. Moore, J. E. Desmond, and N. E. Berthier. (1989). Adaptively timed conditioned responses and the cerebellum: A neural network approach. *Biological Cybernetics*, 62:17–28.

[11] John W. Moore, Neil D. Berthier, and Diana E. J. Blazis. (1990). Classical eye-blink conditioning: Brain systems and implementation of a computational model. In Michael Gabriel and John Moore, editors, *Learning and Computational Neuroscience: Foundations of Adaptive Networks*, A Bradford Book, pages 359–387. The MIT Press.

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

[13] Richard S. Sutton and Andrew G. Barto. (1990). Time-derivative models of pavlovian reinforcement. In Michael Gabriel and John Moore, editors, *Learning and Computational Neuroscience: Foundations of Adaptive Networks*, A Bradford Book, pages 497–537. The MIT Press.

[14] J. H. Wearden and M. F. Doherty. (1995). Exploring and developing a connectionist model of animal timing: Peak procedure and fixed-interval simulations. *Journal of Experimental Psychology*, 21(2):99–115, 1995.