Increasing CS and US longevity increases the learnable trace interval

Xiangbao Wu & William B Levy
Department of Neurosurgery
University of Virginia Health System
P.O. Box 800420 Neurosurgery
Charlottesville, VA 22908-0420, USA

Abstract

It has been hypothesized that increasing CS longevity affects performance on trace conditioning. Using a hippocampal model, we find that increasing CS and US longevity increases the learnable trace interval. Our simulations show that, over a modest range, the maximal learnable trace interval is approximately a linear function of CS/US longevity.

Keywords: Recurrent networks; Sequence learning; Trace conditioning; Learnable trace interval; Hippocampus

1. Introduction

Trace conditioning, devised by Pavlov, can be a hippocampal-dependent task. In this task, a subject is given a stimulus (called the conditioned stimulus or CS). This stimulus is followed by an interval of no stimulus (called the trace interval). Finally, at the end of the trace interval comes the unconditioned stimulus (UCS or US) which produces the unconditioned response (UCR). Eventually, if the trace interval is not too long or too short, the subject learns to anticipate the UCS by generating a conditioned response (CR) at an appropriate time after the CS.

We have been able to produce models of hippocampal-dependent trace conditioning learning that map into real time of a training trial. This mapping is based on the measured off-rate time constant of the NMDA-receptor which controls the time-spanning synaptic modification rule of the model. Taking advantage of this mapping, we present a new result that is a quantitative, testable prediction. Specifically, we predict that longer trace intervals can be learned when the longevity of CS and US is increased.

2. The model

The model described here is an extension of our original hippocampal model of region CA3 [3, 7]. It still uses McCulloch-Pitts neurons that spike or do not spike on any one time-step. The input layer corresponds to a combination of the entorhinal cortex and dentate gyrus (Figure 1a). The CA3 model is a sparsely-interconnected feedback network of typically thousands of neurons where all direct, recurrent connections are excitatory. There is an interneuron mediating feedforward inhibition, and one mediating feedback inhibition. Inhibition is of the divisive form, but activity is only imperfectly controlled because of a delay in the feedback affect which activates these inhibitory neurons.

Region CA3 is modeled as a randomly connected network. Each excitatory neuron randomly connects to approximately $n \cdot c$ other neurons, where n is the number of neurons and c (0.1 here) is the connectivity ratio. Given that the output of neuron i at time t is $z_i(t)$, the net internal excitation of neuron j, $y_j(t)$, is given by

$$y_j(t) = \frac{\sum_{i=1}^{n} w_{ij} c_{ij} \phi(z_i(t-1))}{\sum_{i=1}^{n} w_{ij} c_{ij} \phi(z_i(t-1)) + K_R\left(\sum_{i=1}^{n} D_i(t-1) z_i(t-1)\right) + K_0 + K_I \sum_{i=0}^{n} x_i(t)}$$

where w_{ij} represents the weight value between neurons j and i, and c_{ij} is a 0/1 variable indicating a connection from neuron j to i. The term $\sum_{i=1}^{n} w_{ij} c_{ij} \phi(z_i(t-1))$ represents the excitatory synaptic conductance for the jth neuron. Parameters K_R and K_I are constants that scale the feedback and feedforward inhibitions, respectively. K_0 is a constant that controls the magnitude and stability of activity oscillations and can be considered the rest conductance in a shunting model [12]. Binary, {0, 1}, external input to neuron j at time t is given by $x_j(t)$. The neuron j fires (i.e. $z_j(t) = 1$) if either $x_j(t) = 1$ or if $y_j(t) \geq \theta$ where the threshold θ is fixed at one-half. Synaptic failures are included in this present model. The synaptic failure channel of the connection from neuron i to neuron j is represented by the function $\phi(z_j)$ [13], where $\phi_i(z_j = 0) = 0$ and a synaptic failure $\phi_i(z_j = 1) = 0$, occurs with probability f, and $\phi_i(z_j = 1) = 1$ with probability (1-f), i.e., the failure process is a Bernoulli random variable, which acts independently on each synapse at each time-step. Here the failure rate is 15%.

The model uses a biologically-inspired postsynaptic associative modification rule with time staggering between pre- and postsynaptic activity [2, 5, 6]. For more biological simulations, synaptic modification spans multiple time steps, approximating NMDA-dependent LTP and LTD [1, 11], i.e.

$$w_{ij}(t+1) = w_{ij}(t) + \mu z_j(t)\left(\left(\bar{z}_i(t-1)\right) - w_{ij}(t)\right),$$

where

$$\bar{z}_i(t) = \begin{cases} \bar{z}_i(t-1)\alpha & \text{if } \phi(z_i(t)) = 0 \\ 1 & \text{if } \phi(z_i(t)) = 1 \end{cases}$$

i is input and j is output, and α represents the decay time constant of the NMDA receptor. This decay of activity is exponential, as in all channels, and an e-fold decay in the NMDA receptor has been used in the lab (see section Methods and [4, 10]).

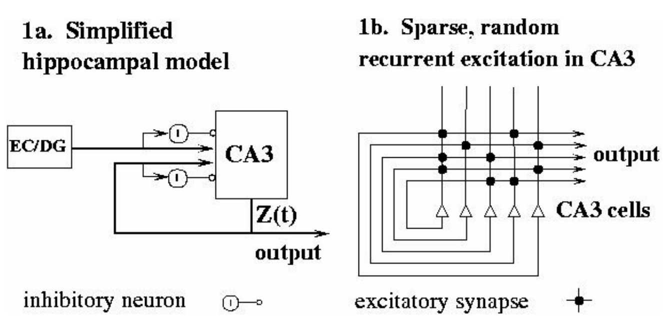

Figure 1 (a) The EC and the DG inputs are collapsed into a single powerful external input class, but (b) most cell firing is driven by recurrent excitation.

For better control of activity, a rule for modification of interneuron afferent synaptic strength is used [14].

$$D_i(t) = D_i(t-1) + \lambda z_i(t-1)\left[m(t)/n - A_D\right],$$

where D_i is the weight of excitatory connection from neuron *i* to the feedback interneuron, λ is the pyramidal-interneuron synaptic modification rate constant, A_D is a free parameter describing the desired percent of neurons activated at each time-step, and m(t) is the number of active neurons at time *t*.

3. Methods

The CS and UCS are represented in the model as input patterns presented at specific times to a network during training (e.g. see Figure 2). The trace interval is represented as no external activity. If, during testing, the network successfully anticipates the UCS at the appropriate time (e.g. about 140 ms prior to the onset of the UCS for a 500 ms trace interval, see [8]), then we say that the network successfully acquired trace conditioning. That is, the network must turn on at least 16 out of 40 neurons prior to US onset. However, turning on US neurons too early (e.g. 200 ms before the US onset) is also considered a failure to learn the trace interval.

In order for a model to accurately capture experimental observations, the behavioral time scale must be mapped onto the network. Time scale is derived from the e-fold NMDA-receptor off-rate time constant $\tau_A \approx 100$ ms (the amount of time required for the glutamate binding to the NMDA receptor to decay to $1/e$ of its previous value); in terms of the synaptic modification equation, we use, e.g., $\alpha = e^{-\Delta t/\tau_A} = e^{(-20/100)} = 0.8187$, which is used in the simulations for Figure 3. To model a CS pattern of duration *T* ms, the neurons represented the CS are externally activated for approximately $T/\Delta t$ time-steps. E.g. if each time-step is approximately 20 ms, a CS of 240 ms is activated for 12 time-steps.

4. Results

Figure 2 shows the development of neural codes in a simulation with a 100 ms CS, 500 ms trace interval, and 150 ms UCS [9]. Each panel is the cell firing of neurons 1-210 (211-2048 not shown) across the indicated training trial. Note that before training, Trial 0, activity is seemingly random except for the externally activated CS and UCS. At the end of training, Trial 200, there are many local context neuronal firings [7] and many of the US neurons fire before the US onset.

Figure 3 summarizes the data showing that increasing CS and US longevity enhances the longevity of the learnable trace interval. For example, when the CS is 100 ms and the US is 40 ms, the maximal learnable trace interval is 840 ms. However, when the CS is 240 ms and the US is 100 ms, the network can learn a trace interval of 1220 ms.

Figure 4 shows an example of the outcome of two simulations using the same network. Surprisingly, this simulation (like others) deteriorates in a rapid, nonlinear manner when the trace interval is made successively longer. As illustrated a 1220 ms trace interval is learnable but a 1240 ms interval is not. That is, the simulation on the left successfully predicts the US on trial 200 but the simulation on the right − with the 20 ms longer trace interval − does not produce an appropriate prediction; only 2 out of the 40 US neurons turn on prior to the US onset.

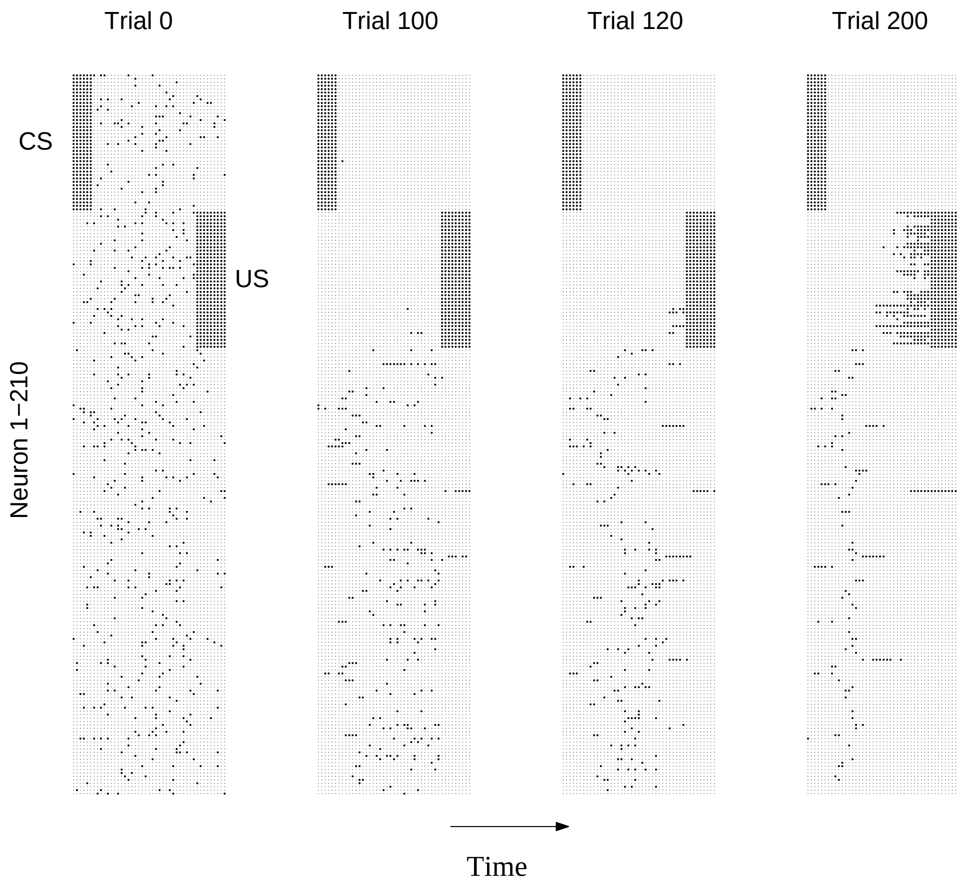

Figure 2. The development of neural codes that predict the UCS in a simulation with a 100 ms CS, 500 ms trace interval, and 150 ms UCS. Each plot is a cell firing pattern of neurons 1-210 (211-2048 not shown) across time steps in different learning trials. The CS is represented as neurons 1-40 at time steps 1-6; the trace interval is represented as no external input from time steps 7-36; and the US is represented as neurons 41-80 at time steps 37-45. Note that before training, Trial 0, activity during the trace interval is the result of random recurrent connections. At Trial 200, the US neurons fire earlier, i.e., before the US onset. A big dot represents a cell firing and a smaller dot represents a non-firing cell. Parameters were n = 2048, m_e = 40, K_R = 0.055, K_I = 0.018, K_0 = 0.596, μ = 0.01, α = 0.8465, λ = 0.5, the initial weights were set at 0.5, and the synaptic failure rate was 15%.

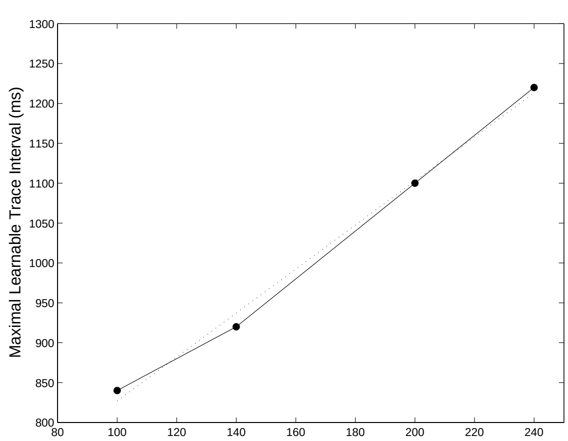

Figure 3. Increasing CS/US longevity increases learnable trace interval. The learnable trace interval is approximately a linear function of the CS/US longevity for the limited range presented here. The correlation coefficient of the linear fit is 0.99. The filled circles connected by the solid

line are the results of simulations and the dotted line is the linear fit. Parameters were n = 2048, me = 40, $K_R = 0.055$, $K_I = 0.050$, $K_0 = 0.001$, $\mu = 0.01$, $\alpha = 0.8187$, $\lambda = 0.5$, the initial weights were set at 0.45, and the synaptic failure rate was 15% which gives activity approximately 5%. Here the ratio of CS and US longevity is kept as close to 2.5 as possible. This ratio is similar to the ones used in Thompson's lab. That is, the CS and US values are {100 ms, 40 ms}, {140 ms, 60 ms}, {200 ms, 80 ms}, and {240 ms, 100 ms} respectively.

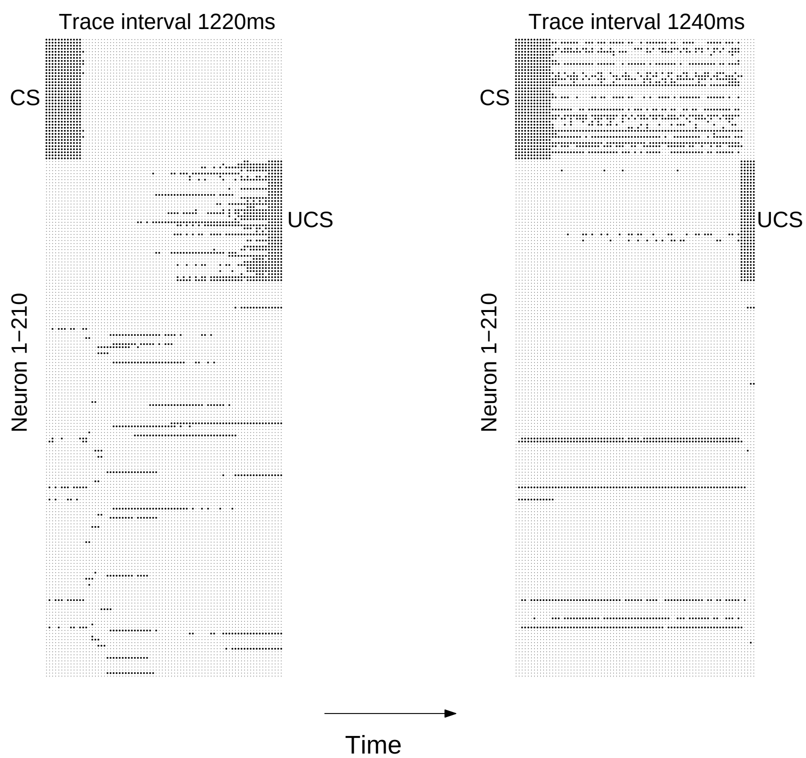

Figure 4. The network learns to predict a trace interval of 1220 ms but fails to learn a trace interval of 1240 ms after 200 training trials. Parameters were n = 2048, me = 40, $K_R = 0.055$, $K_I = 0.050$, $K_0 = 0.001$, $\mu = 0.01$, $\alpha = 0.8187$, $\lambda = 0.5$, the initial weights were set at 0.45, and the synaptic failure rate was 15% which gives an activity of approximately 5%.

5. Conclusion

The model solves the trace conditioning problem with some of the same foibles as have been quantified for eyeblink conditioning in rabbits and produces similar kinds of neural firing [9, 11].

The data presented here predicts that increasing CS/US longevity enhances the learnable trace interval. Thus a minimal biological model of hippocampal function can generate easily testable predictions of animal (and human) learning.

The sharp phase transition observed in the simulations at longer trace intervals (Fig. 4), however, may be sharper than what would be observed in behavioral experiments. One can reuse the same network in many simulations, but the same subject can only be used once in a trace conditioning experiment. Nevertheless, multiple simulations can be used to predict the performance of multiple subjects. Preliminary data indicate that the group phase transition will be noticeable but not as sharp as seen in multiple simulations of a single network.

Acknowledgments

This work was supported by NIH MH48161, MH63855 and RR15205 to WBL; and NSF NGSEIA-9974968, NPACI ASC-96-10920, and NGS ACI-0203960 to Marty Humphrey.

References

[1] D.A. August and W.B Levy, Temporal sequence compression by an integrate-and-fire model of hippocampal area CA3, J. Comp. Neurosci. 6 (1999) 71-90.
[2] W.B Levy, A computational approach to hippocampal function, in: R. D. Hawkins and G. H. Bower, Eds., Computational Models of Learning in Simple Neural Systems, (New York: Academic Press, 1989) 243-305.
[3] W.B Levy, A sequence predicting CA3 is a flexible associator that learns and uses context to solve hippocampal-like tasks, Hippocampus 6 (1996) 579-590.
[4] W.B Levy and P.B. Sederberg, A neural network model of hippocampally mediated trace conditioning, in: IEEE International Congress on Neural Networks, Vol. 1 (IEEE, 1997) 372-376.
[5] W.B Levy and O. Steward, Synapses as associative memory elements in the hippocampal formation, Brain Res. 175 (1979) 233-245.
[6] W.B Levy and O. Steward, Temporal contiguity requirements for long-term associative potentiation/depression in the hippocampus, Neuroscience 8 (1983) 791-797.
[7] W.B Levy, X.B. Wu and R.A. Baxter, Unification of hippocampal function via computational/encoding considerations, Intl. J. Neural Sys. 6(Supp.) (1995) 71-80.
[8] J.J. Kim, R.E. Clark, and R.F. Thompson, Hippocampectomy impairs the memory of recently, but not remotely, acquired trace eyeblink conditioned responses, Behavioral Neuroscience 109 (1995) 195-203.
[9] M.D. McEchron and J.F. Disterhoft, Sequence of single neuron changes in CA1 hippocampus of rabbits during acquisition of trace eyeblink conditioned responses, J. Neurophysiol. 78 (1997) 1030-1044.
[10] K.E. Mitman, P.A. Laurent, and W.B Levy, Defining time in a minimal hippocampal CA3 model by matching time-span of associative synaptic modification and input pattern duration, International Joint Conference on Neural Networks (IJCNN) 2003 Proceedings, 1631-1636.
[11] P. Rodriguez and W.B Levy, A model of hippocampal activity in trace conditioning: Where's the trace? Behav. Neurosci. 115 (2001) 1224-1238.
[12] A.C. Smith, X.B. Wu, and W.B Levy, Controlling activity fluctuations in large, sparsely connected random networks, Network 11 (2000) 63-81.
[13] D.W. Sullivan and W.B Levy, Quantal synaptic failures improve performance in a sequence learning model of hippocampal CA3, Neurocomputing 52-54 (2003a) 397-401.
[14] D.W. Sullivan and W.B. Levy, Synaptic modification of interneuron afferents in a hippocampal CA3 model prevents activity oscillations, International Joint Conference on Neural Networks (IJCNN) 2003 Proceedings, 2003b, 1625-1630.