BRAIN RESEARCH 1365 (2010) 3–17

available at www.sciencedirect.com

ScienceDirect

www.elsevier.com/locate/brainres

BRAIN RESEARCH

Research Report

Timing using temporal context

*Karthik H. Shankar**, *Marc W. Howard*

Department of Psychology, Syracuse University, USA

ARTICLE INFO

Article history:
Accepted 14 July 2010
Available online 21 July 2010

Keywords:
Timing
Weber's law
Trace conditioning
Episodic memory

ABSTRACT

We present a memory model that explicitly constructs and stores the temporal information about when a stimulus was encountered in the past. The temporal information is constructed from a set of temporal context vectors adapted from the temporal context model (TCM). These vectors are leaky integrators that could be constructed from a population of persistently firing cells. An array of temporal context vectors with different decay rates calculates the Laplace transform of real time events. Simple bands of feedforward excitatory and inhibitory connections from these temporal context vectors enable another population of cells, *timing cells*. These timing cells approximately reconstruct the entire temporal history of past events. The temporal representation of events farther in the past is less accurate than for more recent events. This history-reconstruction procedure, which we refer to as timing from inverse Laplace transform (TILT), displays a scalar property with respect to the accuracy of reconstruction. When incorporated into a simple associative memory framework, we show that TILT predicts well-timed peak responses and the Weber law property, like that observed in interval timing tasks and classical conditioning experiments.

1. Introduction

Timing the interval between two events is one of the basic cognitive capacities we all share. This has been rigorously studied in a wide variety of classical conditioning experiments on animals (Drew et al., 2005; Smith, 1968) and explicit interval timing experiments on humans and animals (Rakitin et al., 1998; Ivry and Hazeltine, 1995; Wearden, 1992; Roberts, 1981). One basic finding of these experiments is scalar variability in the underlying timing distributions. Suppose that subjects are trained to reproduce a time interval of a given duration, d_o. The reproduced duration d generally forms a smooth probability distribution peaked approximately at d_o. Moreover the data shows that the standard deviation of the response distribution is proportional to d_o. That is the ratio of the interval to be timed and the standard deviation of the distribution of responses is approximately constant, a manifestation of Weber's law (Gibbon, 1977). More specifically, the response distributions for different values of d_o overlap when they are scaled linearly. This is referred to as the scalar property. When the interval to be timed is short, the peak in the response distribution is narrow and the estimated duration is more accurate than when the interval to be timed is long. Superficially this appears fairly intuitive, but the underlying scalar property has very important implications for models of timing. Similar features are observed in classical conditioning experiments where animals are trained with a conditioned stimulus (CS) followed by an unconditioned stimulus (US) after a latency period. It is observed that the peak of the conditioned response (CR), which we can think of as a measure of the animal's anticipation of the US, approximately matches the reinforcement latency during training. In addition, the time

* Corresponding author.
E-mail address: karthik@memory.syr.edu (K.H. Shankar).

0006-8993/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.brainres.2010.07.045

distribution of the CR activity approximately exhibits the scalar property described above (Drew et al., 2005; Smith, 1968).

In order to model these and related tasks, we need an efficient timing mechanism. This timing mechanism then needs to be integrated with a memory mechanism in order to store and retrieve the timing information. It has been argued that in the 10 to 100 ms range, relevant to speech and motor processing, the dynamically evolving pattern of activity in a spatially distributed network of neurons, is intrinsically sufficient as a timing mechanism (Mauk and Buonomano, 2004), and hence it is unnecessary to postulate a specialized mechanism for timing. However, for longer time scales of the order of seconds to minutes, it seems necessary to have a specialized mechanism. There are many timing models developed over decades involving a variety of specialized mechanisms. They can be divided into two broad classes. See Mauk and Buonomano (2004), Gibbon et al. (1997), Miall (1996), Eagleman (2008), Ivry and Schlerf (2008) for reviews.

The more prominent class of models of timing relies on an internal clock-like mechanism. Models in this class use different mechanisms to construct a scalar representation of elapsed time. Some (Gibbon, 1977; Church, 1984; Gallistel and Gibbon, 2000) use a pacemaker whose pulses will be accumulated to represent perceived time, while others (Church and Broadbent, 1990; Treisman et al., 1990; Miall, 1990) use a population of neural oscillators of different frequencies. Still others (Matell and Meck, 2004; Buhusi and Meck, 2005) use a distributed idea of detecting the coincidental activity of different neural populations to represent the ticks of the internal clock.

The other class of models posits a distributed population of neural units, each of which responds to an external stimulus with a different latency. A straightforward approach is to use tapped delay lines (Moore and Choi, 1997) or chained connectivity between late spiking neurons (Tieu et al., 1999). In these models, the delays accumulated while traversing through each link of the chain add up, thereby making the different links of the chain respond with different latencies. A more sophisticated way to accomplish the same goal is to require the different members of the population to be intrinsically different and react to an external stimulus at different rates. The spectral timing model (Grossberg and Schmajuk, 1989; Grossberg and Merrill, 1992) and multi-time scale (MTS) theory (Staddon et al., 2002) both share this property. MTS, for instance, assumes a cascade of leaky integrators where the activity in each unit exponentially decays following a stimulus with a distinct decay rate.

In this paper, we construct a timing mechanism in the framework of the temporal context model (TCM), an associative memory model that has been extensively applied to problems in episodic recall (Howard and Kahana, 2002; Howard et al., 2005; Sederberg et al., 2008). This timing model falls into the second class of timing models—it has no explicit clock system like a pacemaker or synchronous oscillators. Instead, the model requires a population of persistently firing neurons with a range of decay rates, similar in many respects to the cascade of leaky integrators of MTS (Staddon et al., 2002). We show that this population of leaky integrators implements the Laplace transform of the stimulus sequence. Using this insight, we approximate the inversion of the Laplace transform, constructing a separate population of “timing cells”. We refer to this procedure as timing from inverse Laplace transform, TILT. The approximation of the inverse Laplace transform can be accomplished using bands of alternating feedforward excitation and inhibition from the leaky integrators. In effect, the leaky integrators implement the Laplace transform of the stimulus history and the timing cells approximately inverts this Laplace transform, thus generating an approximate reconstruction of the stimulus history. Each of the timing cells responds with peak activity at a different delay following a stimulus. The effect of this inversion is thus not unlike that generated by the spectral timing model or the delay line models. However, it turns out that the activity across the timing cells at any instant precisely shows the scalar property. When integrated into an analog of TCM’s learning and retrieval mechanisms, the model generates a prediction of the immediate future that reflects prior learning experiences. Rather than developing a detailed model of behavior, the focus of this paper will be on describing the qualitative features of the proposed timing mechanism. However, we do demonstrate that a simple behavioral model derived from this prediction qualitatively exhibits the Weber law property at the behavioral level.

We start with a brief description of the encoding and retrieval mechanisms of TCM. Following that, we construct the timing mechanism and discuss its neural representation. Finally, we integrate the timing mechanism with an analog of the learning and retrieval rules of TCM to qualitatively account for behavioral aspects of timing observed in classical conditioning experiments.

2. TCM

The initial goal of TCM was to account for the recency and contiguity effects observed in episodic recall tasks. The recency effect refers to the finding that, all other things being equal, memory is better for more recently experienced information. The contiguity effect refers to the finding that, all other things being equal, items experienced close together in time become associated such that when one comes to mind it tends to bring the other to mind as well (Kahana et al., 2008). The basic idea of TCM is that when a sequence of stimuli is presented successively, each stimulus is associated with a gradually-varying context state. Any two stimuli that have been experienced in close temporal proximity, though not directly associated, are indirectly linked as a consequence of associations to similar contexts.

The architecture of TCM can be formalized in terms of a two layer network comprised of a stimulus layer and a context layer with bidirectional connections between them as shown in Fig. 1. Each node in the stimulus layer denoted by $\mathbf{f}$ corresponds to a unique stimulus. The external input drives the activity in this layer. At any instant, only the specific node corresponding to the stimulus being perceived is active in this layer. Each of these nodes can be viewed as a distributed set of neurons and different nodes could potentially share some neurons. But we assume this overlap to be sparse; for