2022

Remembering the past to predict the future: a scale-invariant timeline for memory and anticipation

BOSTON UNIVERSITY

SCHOOL OF MEDICINE

Dissertation

REMEMBERING THE PAST TO PREDICT THE FUTURE:

A SCALE-INVARIANT TIMELINE FOR MEMORY AND ANTICIPATION

by

WEI ZHONG GOH

B.S., Brandeis University, 2015
M.S., Brandeis University, 2015

Submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

2022

Approved by

First Reader _____

Marc W. Howard, Ph.D.
Professor of Psychological and Brain Sciences
Professor of Physics

Second Reader _____

Mark A. Kramer, Ph.D.
Professor of Mathematics and Statistics

ACKNOWLEDGMENTS

I thank my advisor, Marc W. Howard, for his invaluable guidance; Graduate Program in Neuroscience program directors, Shelley Russek and Sandra Grasso, for their tireless support; my committee members, Mark Kramer, Benjamin Scott, Joe McGuire, and Vijay Mohan K Namboodiri; and my collaborators, laboratory colleagues, classmates, teachers, family and friends.

I pay tribute to the sacrifice of life and liberty of the laboratory experimental animals whose data was analyzed in this work.

REMEMBERING THE PAST TO PREDICT THE FUTURE:

A SCALE-INVARIANT TIMELINE FOR MEMORY AND ANTICIPATION

WEI ZHONG GOH

Boston University School of Medicine, 2022

Major Professor: Marc W. Howard, Professor of Psychological and Brain Sciences, Professor of Physics

ABSTRACT

To guide action, animals anticipate what events will occur, and when they will occur, based on experience. How animals anticipate future events is an unsettled question. Although reinforcement learning is often used to model anticipation, it is resource-intensive outside of the simplest scenarios. In this dissertation, I show evidence of memory that is persistent and carries timing information, and specify an algorithm for how animals might anticipate the identity and timing of future events.

This dissertation consists of two studies. In the first study, I found that identity and timing of remembered odors are jointly represented in the same cells in the dentate gyrus and lateral entorhinal cortex. Further, odor memories persist well after new odors emerge. The study analyzed results from an experiment conducted by Woods et al. (2020) on mice passively exposed to separate odors for a period of 20 s per exposure. The results are consistent with a memory framework known as timing using inverse Laplace transform (TILT).

In the second study, I constructed a computational algorithm based on the TILT memory framework to anticipate the identity and timing of future events. The algorithm generates

predictions based on memories of past events, and stored associations between cues and outcomes. The algorithm is resource-efficient even when the future depends on the indefinite past. The algorithm is scale-invariant and works well with chains of events.

Together, the studies support a novel computational mechanism which anticipates what events will occur, and when they will occur. The algorithm could be applied in machine learning in cases of long-range dependence on history. These studies predict that behavioral and neural responses of animals could depend on events well into the past.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 INTRODUCTION

1.1 Overview of This Section

This dissertation is about a mental framework that anticipates what events will happen, and when. By *anticipation*, I mean a feeling of expectation that an event is about to occur. This dissertation also includes a brief experimental study. The study's results are in line with features of that mental framework. In this section, first, I review behavioral and neural evidence for anticipation. Second, I point out pitfalls of an existing, widely used framework related to anticipation. Third, I lay out the structure of the rest of this dissertation.

1.2 Behavioral and Neural Evidence for Anticipation

The future guides our actions in the present. To be more precise, what we suppose will happen in the future, based on cues in the past, guides what we do in the present. While walking, we unconsciously anticipate obstacles as they near in order to keep our feet clear at the right moment (see, e.g., Potocanac and Duysens, 2017), even if the obstacles fall out of our field of view. While watching a talk or a movie, we anticipate less interesting moments to take bathroom breaks—perhaps an uninteresting question was asked. While making pasta, we check for doneness more and more often as the expected cook time nears, as we anticipate the time the pasta is done (see, e.g., Einstein et al., 1995). Sometimes our expectations are wrong, and we are worse off: we trip; we miss out on interesting content; we are left with mushy pasta. As the above examples show, not only do we anticipate which events would occur in the future, we also have a sense of when they would occur.

This dissertation is about how this sense of anticipation might arise. As a note on terminology, in order to be in line with the existing literature, I use *prediction* and *anticipation* interchangeably. I use both words to refer to an intuitive feeling of expectation about the future on the basis of past experience, examples of which were given in the previous paragraph. In other contexts, *prediction* might also carry the meaning of *forecast*, as in *meteorological prediction*. *Prediction* in this other sense is a forecast, made by humans or computers, after deliberation and calculation, perhaps on the basis of theory. I do not use *prediction* in this other sense.

Anticipating the time of future events allows animals to time reward-seeking and punishment-avoiding action. For instance, anticipation allows animals to return to a food source at the time it is expected to be replenished with food (Bateson, 2003), or to lick a spout at the appropriate time (Toda et al., 2017). In addition, the anticipated imminence of reward may help animals decide whether to stay in the current environment in hope of obtaining reward (Shintaki et al., 2021). Experiments have shown that, indeed, animals have a rich ability to anticipate future events. Animals demonstrate this at least both by recognizing the value of reward-predicting cues (e.g., Fiorillo et al., 2008) and expecting cued events around the time they are due, up to delay intervals of hours (Balsam et al., 2010). Behavioral results from experiments (e.g., Cardinal et al., 2001) suggest that mammals and some birds decide between immediate and future rewards based on the delay-discounted value of each alternative (Frost and McNaughton, 2017; Vanderveldt et al., 2016). These results suggest at least a basic ability to compute on the expected future. Recent psychology research has suggested that humans maintain a timeline of the future just like we do a timeline of the

past. In an experiment performed by Babcock et al. (2020), human participants exposed to a repeating sequence of letters were able to select letters that were about to appear within a time window 4–6 items (4–9 s) in the future. The future was decoupled from the past through random jumps in the sequence of letters. The idea that we maintain a timeline of the future is in line with the finding that humans conceive of time spatially along a mental timeline, in Western cultures usually with the past to the left and the future to the right (e.g., Weger and Pratt, 2008). Collectively, these experimental results corroborate our intuition that we anticipate both what and when future events will occur, as if they are laid out in an internal future timeline.

While neural evidence of encoding of such a future timeline remains elusive for now, there is extensive neural evidence that animals recognize the value of reward-predicting cues and predict reward at around the time that reward is due. This evidence comes in the form of phasic activation of midbrain dopaminergic cells in coding reward prediction error. Suppose there are two instantaneous events X (which is neutral, say, a bell) and Y (which is rewarding, say, the dispensing of a sugary treat), and X reliably precedes Y by 1 s. Midbrain dopamine cells respond to Y initially, but the response shifts to X with learning. The response to the reward-predicting cue X may be interpreted as an expectation that reward will occur at some point in the future. Crucially, on rare trials where Y fails to occur after X, the dopamine cell response dips below baseline shortly after the expected time of occurrence of Y. This suggests that the animal maintains a time-locked prediction for Y, which is available around the time when Y is due (Hollerman and Schultz, 1998; Morris et al., 2004; Daw et al., 2006). Thus, there is neural evidence that events are anticipated based on cues from the past.

1.3 The Dopamine Reward Prediction Error Hypothesis Is Unsuitable as a General Account of Anticipation

Researchers have attempted to explain the activity of dopamine cells using the framework of temporal difference (TD) reinforcement learning (RL). In the canonical account, the moment the cue occurs, the moment the reward occurs and all of the intervening moments each correspond to a separate state for the animal (Ludvig et al., 2012). The animal estimates an attribute for each state known as value V, which is the sum of discounted future reward expected starting from that state. In TD learning, when the animal enters a new state s' at time t, the value estimate $V_t(s)$ at time t for the previous state (or the previous few states) is incrementally updated (yielding $V_{t+1}(s)$). The target for this learning is the discounted value of the new state plus any reward obtained. The difference between the target and the old value estimate $V_t(s)$ is the TD error (Sutton and Barto, 2018), numerically equivalent to the reward prediction error (RPE). According to the dopamine RPE hypothesis, midbrain dopamine cell response encodes RPE (Glimcher, 2011). Overall, the dopamine RPE hypothesis has enjoyed wide experimental support (Glimcher, 2011).

However, the framework of TD RL is not completely satisfactory as a general account for anticipation, especially over time scales longer than tens of seconds. The TD RL framework crucially relies on the notion of an animal being in a state (or a set of states) at each point in time between the cue and the reward. However, in general, non-human animals can associate cues and rewards separated by anywhere between hundreds of milliseconds and tens of hours (Balsam et al., 2010). We can either suppose that the number of intervening states between the cue and the reward (and therefore the speed of progression through the

states) is independent or dependent on the delay period. Both options are problematic. For the number of states to be independent of the delay period, each state should last on the order of tens to hundreds of milliseconds for sufficient resolution to accommodate the shortest delay periods. To accommodate delay periods of tens of hours, we would need about a million such states, which can be resource-intensive to keep track of. To reduce the number of states, we can posit that the animal dwells longer and longer in each state as time since the cue elapses. However, in this case, implementing state evolution in a way that is neurobiologically realistic is nontrivial. To my knowledge, this idea has not been seriously pursued in the literature. Alternatively, we can suppose that the number of states is dependent on the delay period. This requires the animal to have foreknowledge of the delay period before acquiring the association, resulting in a catch-22 situation. Therefore, the framework of TD RL seems at odds with the wide variety of time scales across which animals associate cues and rewards. (See K Namboodiri, 2021, for additional objections to TD RL; for related discussion, see Sections 4.5.1 and 4.5.3.1.)

1.4 Structure of This Dissertation

To respond to the foregoing, the remainder of this dissertation proceeds as follows.

In Section 2, we formalize the prediction problem.

In Section 3, we look briefly at experimental results suggesting that brain regions of animals exposed to a train of odors carry information about past items. This observation reinforces the importance of a timeline of the past, which is a feature of a cognitive model known as timing using inverse Laplace transform (TILT), and not TD RL (Shankar and

Howard, 2012, 2013). Section 3 is a re-analysis of a previously published work (Woods et al., 2020), this time looking at the time evolution of cell activity within and across trials.

In Section 4, we build a cognitive architecture that allows an agent to anticipate future events. The algorithm we build is distinct from the framework of TD RL. Indeed, TD RL has been successful in accounting for the phasic activity of dopamine neurons. With sufficient modifications, the determined researcher may even be able to extend TD RL to a general framework for anticipation. However, instead of working under the paradigm of TD RL, we use the TILT paradigm. This paradigm is outlined in Section 4.1. Section 4 and Appendix A reproduce the main text and the appendix, respectively, of the author manuscript for Goh et al. (in press) in full.

TILT naturally operates across a wide variety of time scales, unlike TD RL, as explained in the previous section. Moreover, the framework we propose and TD RL differ in terms of the relative importance of value (as a function of states) and memory and prediction (as a function of time in the remembered past and expected future). In the framework we propose, a timeline of the past is used to generate a timeline of the future. Reward prediction error can be derived if desired (see Section 4.5.3.1), and there is hardly a notion of value of each state. On the other hand, to TD RL, the value of each state is central, and the reward prediction error functions crucially as the learning signal for the value of each state. Information about the past, if relevant, is encoded in the state, so a timeline of the past is unnecessary. A timeline of the future could, in theory, be generated if desired (see Kim et al., 2020), although this is unnecessary for things to work, and I am not aware of serious work to construct a timeline of the future from TD RL. Given the importance of timing in

anticipation in animals (see Section 1.2) and evidence for neural representation of a timeline of the past (Pastalkova et al., 2008; Salz et al., 2016; Shimbo et al., 2021; Taxidis et al., 2020; Tiganj et al., 2018; Mello et al., 2015), it may be preferable to adopt a framework in which these timelines are core.

Finally, in Section 5, we conclude with a brief discussion of future directions for the work described in Sections 3 and 4.

2 FORMALISM:

THE PROBLEM OF PREDICTION BASED ON CUES

In this section, we formalize the problem of prediction based on cues, which is the central problem considered in this dissertation. This will lay the groundwork for the proposed algorithm in Section 4.

2.1 Cues and Outcomes Are Formalized as Point Events

An agent observes various event occurrences in its environment. Each event occurrence, with index i, is completely described by its event type $\xi_i \in \mathcal{E}$ and occurrence time $\tau_i \in \mathbb{R}$. The events are instantaneous (i.e., they do not extend in time). The set of possible event types, $\mathcal{E}$, is fixed given the agent and its environment. We refer to the possible event types (i.e., the members of $\mathcal{E}$) with small capital letters ($\mathrm{X}, \mathrm{Y}, \mathrm{Z}$, etc.). These events encode, for example, the onset of a tone, light, odor, shock or reward.

2.2 Cue–Outcome Relationships Are Formalized as Motifs

In the environment, the events occur in series known as motifs. Motifs formalize cue–outcome relationships between events. Each motif, with index j, is completely described by its event occurrence type sequence $\left(\xi_1^j, \xi_2^j, \ldots, \xi_{L_j}^j\right) \in \mathcal{E}^{L_j}$ and joint probability density function $h\left(\tau_2^j, \tau_3^j, \ldots, \tau_{L_j}^j\right)$. The density h is followed by the time offsets, relative to the first event occurrence in the sequence, of each subsequent event occurrence. The time scale of the motif can be understood as any appropriate measure of central tendency of h.

When a motif j emerges at time τ, it means $\left\{\xi_\ell^j \text{ occurs at time } \tau + \tau_\ell^j\right\}_{\ell=1}^{L_j}$ ($\tau_1^j = 0$,

$\left(\tau_\ell^j\right)_{\ell=2}^{L_j}$ drawn according to h). (One may have non-occurrence of some a_ℓ^j by defining additional motifs.) Purely for convenience of simulation and interpretation, we may restrict $\tau_\ell^j \geq 0$ for all j, ℓ. The motif j emerges randomly as a Poisson point process with rate function $\lambda_j(t)$. In other words, $\lambda_j(t)\,dt$ is the probability that the motif j emerges between time t and $t+dt$, where dt is a short time interval. Where a motif is periodic, $\lambda_j(t)$ is a Dirac comb. Where a motif emerges uniformly at random, the rate function $\lambda_j(t) = \lambda_j$ is independent of time. The key challenge we face is that motifs of different time scales may be in play concurrently.

2.3 The Agent Encounters Events in Real Time

The event signal for events of type $E \in \mathcal{E}$ is given by $f_E(t) = \sum_i \delta(t - \tau_i) I(E = \xi_i)$, where $\delta(\cdot)$ denotes the Dirac delta function and $I(\cdot)$ denotes the indicator function. At regular intervals spaced Δt apart, the agent accesses $\int_{t-\Delta t}^{t} f_E(t')\,dt'$ for all $E \in \mathcal{E}$ (i.e., event counts occurring within the past Δt for each event type), where t is the time of access. The time resolution Δt is a property of the agent. For mammals, it is on the order of milliseconds to tens of milliseconds (see Muchnik et al., 1985). The agent has no direct access to past (and future) values of $f_E(t)$. In other words, the agent experiences the events in real time.

2.4 We Desire a Timeline of the Future

I will now write down a particular definition of *prediction* or *anticipation*. Given the above, the prediction problem is for the agent, at time t, to construct a prediction function $p_E(\delta; t)$ for all $0 \lesssim \delta < \delta_{\max}$, such that $p_E(\delta; t)$ approximates $f_E(t+\delta)$, for all $E \in \mathcal{E}$. Here, $\delta_{\max}$

refers to the time horizon of the agent, which is equal to the maximum time scale to which the agent is attuned. This could be on the order of days for animals (see Balsam et al., 2010). For each E, the function $p_E(\delta;t)$ encodes the agent's understanding of when E will happen in its future at time t. Hence, $p_E(\delta;t)$ evolves in time. We usually suppress the variable t, writing $p_E(\delta)$. Collectively, the function $\mathbf{p}(\delta) = (p_E(\delta))_{E\in\mathcal{E}}$ represents a timeline of the future, which includes information about what will happen, and when. Our desire for the availability of the prediction for all δ at time t is important, especially when considering the framework of TD RL.

2.5 For Feasibility, Our Timeline of the Future Averages Over Interdependence Between Events

At this stage, I keep the discussion general as to how, mathematically, $p_E(\delta)$ should approximate $f_E(t+\delta)$. However, I note that at time t, $p_E(\delta)$ represents the prediction marginalized on that event E, based on the information available up to time t.

Evidently, the agent is powerless to accurately predict events belonging to motifs that emerge between t and $t+\delta$, beyond baseline rates of emergence.

In addition, in principle, an agent could have richer information about the interdependence between event occurrences in the future that $\mathbf{p}(\delta)$ cannot encode due to event marginalization. For instance, suppose X cues Y at (a delay of) 1 s, X cues Z at 1 s (both with some small variance), and both motifs emerge uniformly at random at the same rate. When X occurs, say at t_1, in principle, an agent could expect that either Y or Z, but probably not both, occurs about 1 s later. However, information about the exclusion between Y and Z

cannot be in $\mathbf{p}(\delta;t_1)$, due to event marginalization. Instead, $\mathbf{p}(\delta;t_1)$ would just contain an expectation that Y (magnitude 0.5) would occur 1 s later, and, separately, the same for Z. The function $\mathbf{p}(\delta;t_1)$ has no space for the mutual relationship between Y and Z. It can and should be the case that when either Y or Z indeed occurs (at t_2), that the prediction $\mathbf{p}(\delta;t_2)$ for the other event type be muted, but that would only occur at t_2 and not t_1.

Note that encoding the relationship between Y and Z in a prediction when X occurs, would not be simply a matter of using a bivariate joint probability density function for the occurrence times of Y and Z. This is because unlimited numbers of Y and Z can occur in the imminent future (unlike in a motif, whose event type occurrence sequence has finite length). Thus, an analog of $\mathbf{p}(\delta)$ that encodes joint relationships in generality would be intractably high-dimensional. To be feasible, we keep $\mathbf{p}(\delta)$ as an average over all possible futures. If desired, the agent can separately store relationship information among events. Indeed, the algorithm we propose does so.

3 EXPERIMENTAL STUDY: HISTORY ENCODING IN THE DENTATE GYRUS AND LATERAL ENTORHINAL CORTEX

3.1 Introduction

Animals learn from experience, gaining the ability to avoid threats, hunt prey and seek mates. This learning requires episodic memory, which is memory for specific events at specific times and places, or temporal and spatial contexts. As discussed in Section 1, the binding of time to item is a hallmark of episodic memory, but the cellular origins of our abilities to bind time to item are not well understood. Encoding of episodic memory depends crucially on the hippocampus (Hainmueller and Bartos, 2020), which houses the dentate gyrus (DG) and adjoins the lateral entorhinal cortex (LEC). The LEC, together with the medial entorhinal cortex, provides most of the cortical input to the hippocampus. The LEC receives information from cortical sensory areas relating to objects and events (Eichenbaum et al., 2012), and projects via the perforant path to granule cells in the DG and pyramidal cells in the CA3 and CA1 subfields of the hippocampus; DG granule cells also project to CA3 via mossy fibres (Hainmueller and Bartos, 2020). The LEC is thought to bind sensory stimuli to context, while the DG is thought to enhance discrimination of memories of contexts, locations and stimuli (Pilkiw et al., 2017; Hainmueller and Bartos, 2020; Woods et al., 2020). Accordingly, inactivating either the DG or the LEC adversely affects the memory performance of animals (Tanninen et al., 2013; Hainmueller and Bartos, 2020).

This study explores the encoding of time and item in the DG and LEC in a task-free paradigm, from two perspectives. Firstly, we study whether information about item and time is carried in separate cell subpopulations, or in the same cells. Secondly, we study whether information about the item associated with the previous trial is retained in the DG and LEC after a new trial begins, just like for novelty, past choice and past reward in other brain regions (Ben-Yakov et al., 2014; Larkin et al., 2014; Fried et al., 1997; Bernacchia et al., 2011; Morcos and Harvey, 2016).

To expand on the first objective, the episodic memory for an event consists of the event identity ("item") and time. Broadly, we can distinguish between two ways of representation of time and item. For the first alternative, a set of cells carry information about event time irrespective of item identity ('time-only cells'), and a separate set of cells carry information about the corresponding item identity irrespective of event time ('item-only cells'). In this case, time–item conjunctive representation is absent. The role of time-only cells may be fulfilled by time cells that are not stimulus-specific (in our case, responding to any odor or tone); time cells are cells that activate around a specific time delay relative to the occurrence of a specific salient stimulus (Eichenbaum, 2014). The role of item-only cells may be fulfilled by cells that fire persistently during a delay period in response to specific items (e.g., Masse et al., 2020; Major and Tank, 2004; Kamiński et al., 2017; Suzuki et al., 1997; Xu et al., 2014). The activity of these stimulus-specific persistently firing cells have been thought to underpin working memory (Masse et al., 2020).

For the second alternative, individual cells carry information about both time and item. In this case, time–item conjunctive representation is present. Stimulus-specific time cells,

each of which responds only to a subset of stimuli, are one example of cells with such representation (Pastalkova et al., 2008; Salz et al., 2016; Shimbo et al., 2021; Taxidis et al., 2020; Tiganj et al., 2018; Mello et al., 2015). Temporal context cells are another example. The activity of temporal context cells steps up in response to a stimulus of interest, then decays exponentially with a range of time constants (e.g., Tsao et al., 2018; Bright et al., 2020). Conjunctive representation is closely related to the concept of mixed selectivity, which has been previously demonstrated for discrete task-relevant variables (Rigotti et al., 2013). A cell displays mixed selectivity when its activity is a function, possibly nonlinear, of more than one task-relevant variables. Cells that conjunctively represent item and time, in other words, display mixed selectivity for item and time within trial. Time–item conjunctive encoding has previously been shown in the monkey hippocampus and lateral prefrontal cortex (Cruzado et al., 2020; Tiganj et al., 2018). Here, we investigate potential conjunctive representation of item and time, which is a continuously changing continuous variable, within a task-free paradigm.

This study is a reanalysis of data from Woods et al. (2020). The Woods et al. study focuses on the different roles of the DG and LEC in supporting odor-based associative learning. Woods et al. found that DG odor representation requires LEC synaptic input, and that discriminability of DG representation supports behavior. Woods et al. also found that over the course of learning, DG responds more to the conditioned odor and less to the non-conditioned odor, more so than the LEC. This suggests a crucial role for the DG in discriminating among similar stimuli.

To form a basis for comparison, Woods et al. collected baseline DG and LEC data during

passive odor exposure in naive mice, before behavioral intervention. Our study focuses on this baseline data. In the bulk of their analysis involving this baseline data, Woods et al. averaged over the cell activity over the 4 s odor exposure phase. Our study looks at odor representation within the 4 s odor exposure phase, as well as the 16 s period after odor exposure, at the level of 0.27 s time bins, which is the time resolution of the calcium imaging performed. Whereas Woods et al. relates cell activity with the simultaneously presented odors, our study relates cell activity with odors presented on previous trials. Therefore, our study focuses on the rich time evolution of the DG and LEC odor representation within and across trials.

3.2 Results

3.2.1 DG and LEC Contain Single Cells That Are Both Item- and Time-Specific

Mice were passively exposed to a series of odors in randomized complete blocks, while calcium imaging was performed on the DG or LEC. Figure 3.1 shows rasters for single cells in the DG and LEC that are reliably active at specific times within trial, for specific items. The cells resemble stimulus-specific time cells, peaking in activity around a single time within trial, although cell reliability is mixed.

3.2.2 DG and LEC Encode Time Within Trial

To explicitly demonstrate that DG and LEC encode time within trial, for each animal, we trained a linear discriminant analysis (LDA) decoder to decode time within trial. Specifically, we divided the 20 s trial duration evenly into 4 s quintiles, and trained the decoder to decode

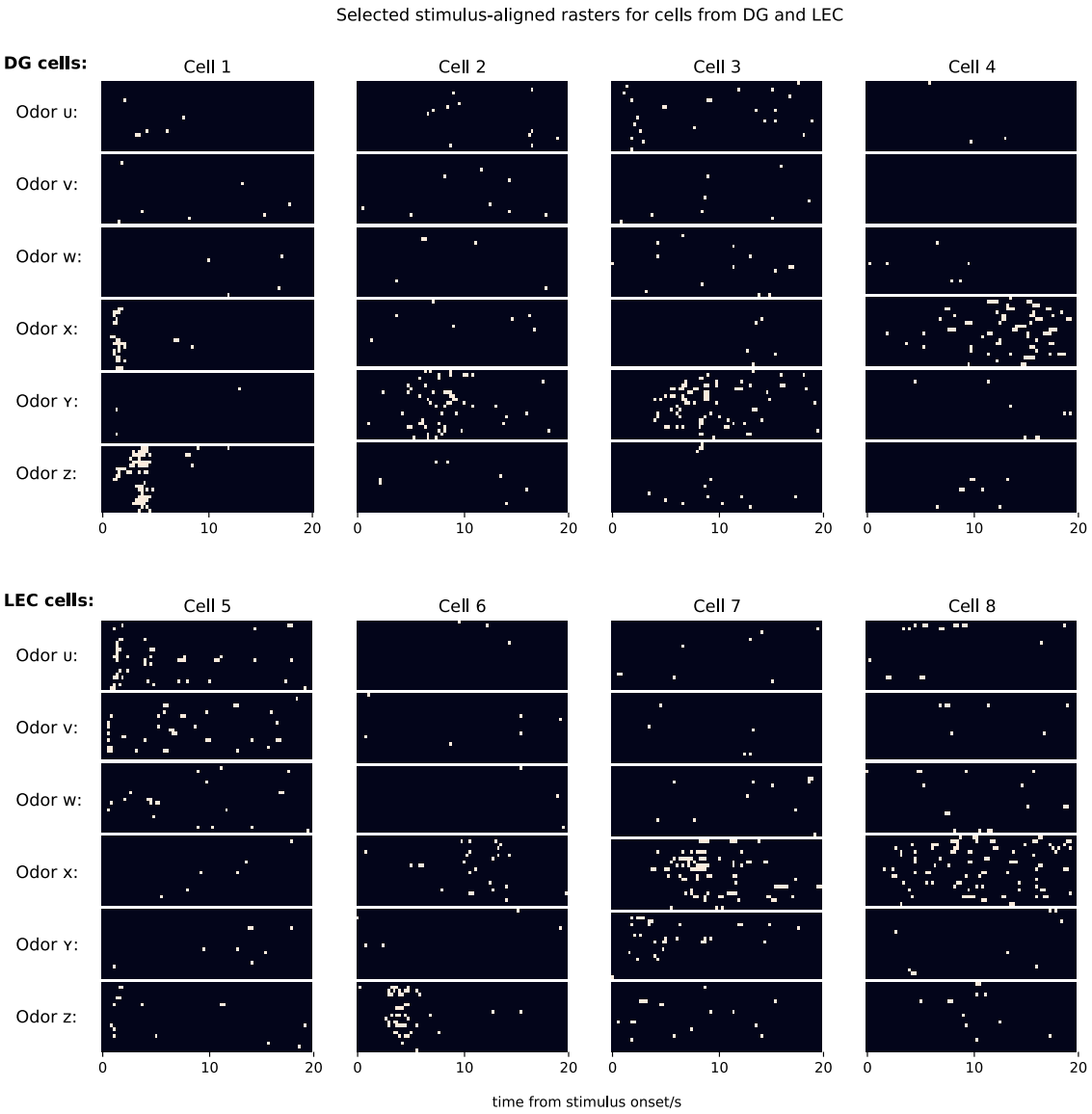

Figure 3.1: **DG and LEC contain single cells that are both item- and time-specific.** Calcium imaging rasters from eight regions of interest, each from a different animal, from our data are shown. The cells are active for specific items at specific times. For instance, Cell 3 appears active specifically in Odor Y trials at about 8 s relative to odor onset, while Cell 4 appears most active specifically in Odor X trials at about 12–18 s relative to odor onset.

Calcium images were taken while mice were passively exposed to a series of odors, one odor per 20 s trial. Each row is associated with one 20 s trial, and each column is associated with one time bin, each lasting 0.27 s. Each white dot indicates detected calcium events, used as a proxy for neural activity, during the time bin. The trials are grouped by item. There are 20 trials in total for each item, and 6 possible items.

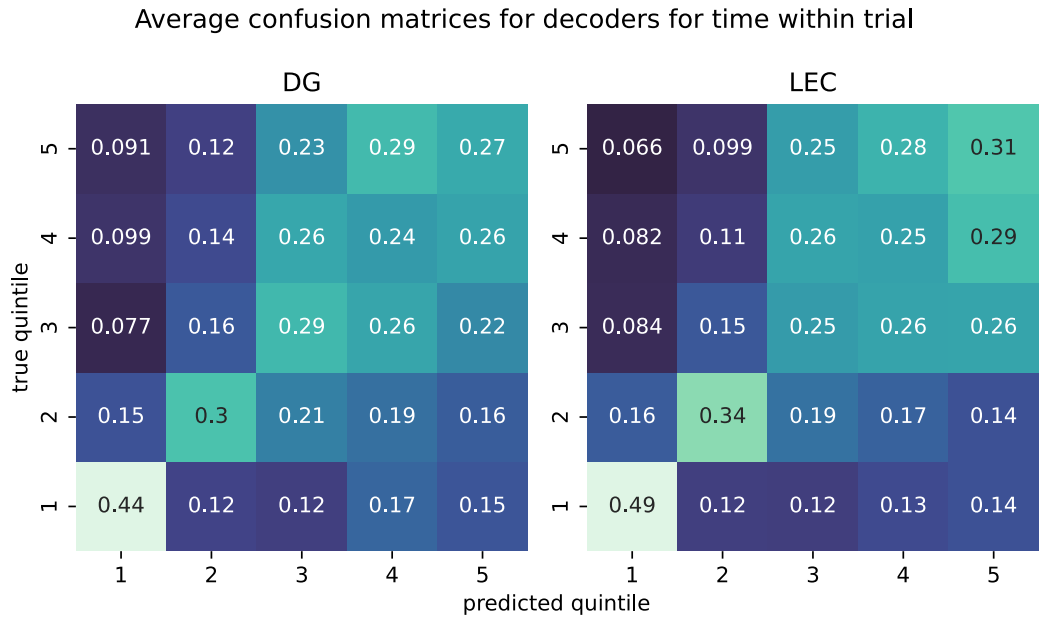

Figure 3.2: **DG and LEC contain information about time.** Confusion matrices, averaged across 8 animals each, for linear decoders for time within trial in animals with DG (left) and LEC (right) imaged respectively. The duration of each trial was evenly divided into five quintiles. The first quintile coincides with odor presentation. Values are fractions of data predicted to be associated with the quintile denoted on the horizontal axis, using linear discriminant analysis decoders trained on time-averaged population vectors over the quintile denoted on the vertical axis. Elements that are more luminant represent larger values. The vertical axis is the true quintile associated with the population vectors, and the horizontal axis is the quintile predicted by the decoder.

the quintile associated with the input based on the population vectors averaged over the quintile. To test for significance, we permuted the labels of the quintiles. The decoders performed significantly above chance (Bonferroni corrected $p < 0.05$) for 7 out of the 8 DG animals, and 7 out of the 8 LEC animals. The confusion matrices are shown in Fig. 3.2.

3.2.3 DG and LEC Each Conjunctively Encode Item and Time

To investigate whether information about item and time is carried in separate cell subpopulations, or in the same cells, we used a cross-temporal classification approach. We train

decoders for item based on time-averaged population vectors from a time window within the trial, and test using data from a different time window. The data are consistent with the presence of time–item conjunctive encoding (Fig. 3.3; see also King and Dehaene, 2014).

The cross-temporal classifier heat maps are consistent with time–item conjunctive encoding because accuracy peaks along the diagonal and decreases as one moves away from the diagonal (Fig. 3.4a). This is consistent with a population representation of item that is changing smoothly with time within each trial. We performed linear regression based on the patterns we expected from a decaying item signal and smoothly changing item representation (Fig. 3.4b). As a group, the coefficients are significant for both DG and LEC animals, suggesting a decaying item signal and smoothly changing item representation based on the accuracy pattern (Fig. 3.4c).

3.2.4 DG and LEC Encode Previous Item Identity

In addition to showing time–item conjunctive encoding within a 20 s trial, we are able to decode the previous stimulus based on the inferred calcium events time-averaged between 21 s and 40 s after odor onset. To investigate this, we build a supermouse by merging the data across 8 animals whose DG was imaged, and did the same for another 8 animals whose LEC was imaged. We compared the performance of decoders trained on trial-averaged population vectors. Evaluation was based on accuracy in decoding item identity at various trial offsets. Table 3.1 summarizes the results from the supermice analysis.

The +1 decoders were evaluated on their ability to decode the next trial. These were unable to decode above chance, the randomized complete block design notwithstanding.

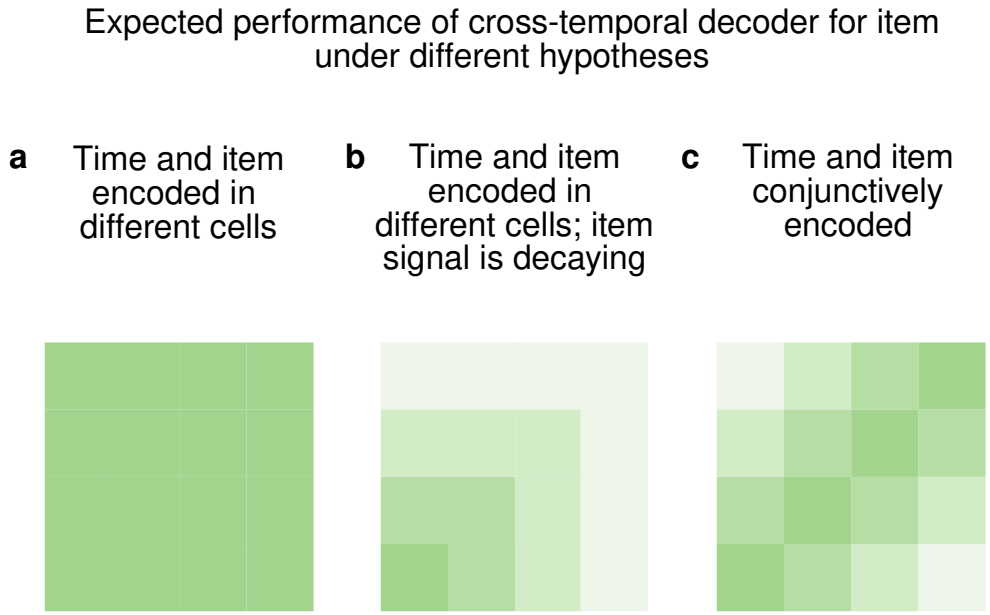

Figure 3.3: **A cross-temporal decoder for item reveals the presence of time-item conjunctive encoding.** The figure depicts the expected performance of a cross-temporal decoder for item under three different hypotheses. Higher decoder performance is associated with similarity of item encoding between training and testing time windows. Higher decoder performance is depicted as higher saturation. The horizontal axis is training time window and the vertical axis is testing time window. Random error is not shown. **(a)** If time and item are encoded in different cells, and the item signal is maintained throughout the trial, the performance of the item decoder should be unaffected by training and testing time. **(b)** If time and item are encoded in different cells, and the item signal strictly decays as the trial unfolds, the performance of the item decoder would decrease with the training or testing time, whichever is the later. **(c)** If time and item are conjunctively encoded, and the item representation is smoothly changing with time and maintained throughout the trial, the performance of the item decoder would smoothly decrease with distance from the diagonal, that is, the distance between the training and testing time window.

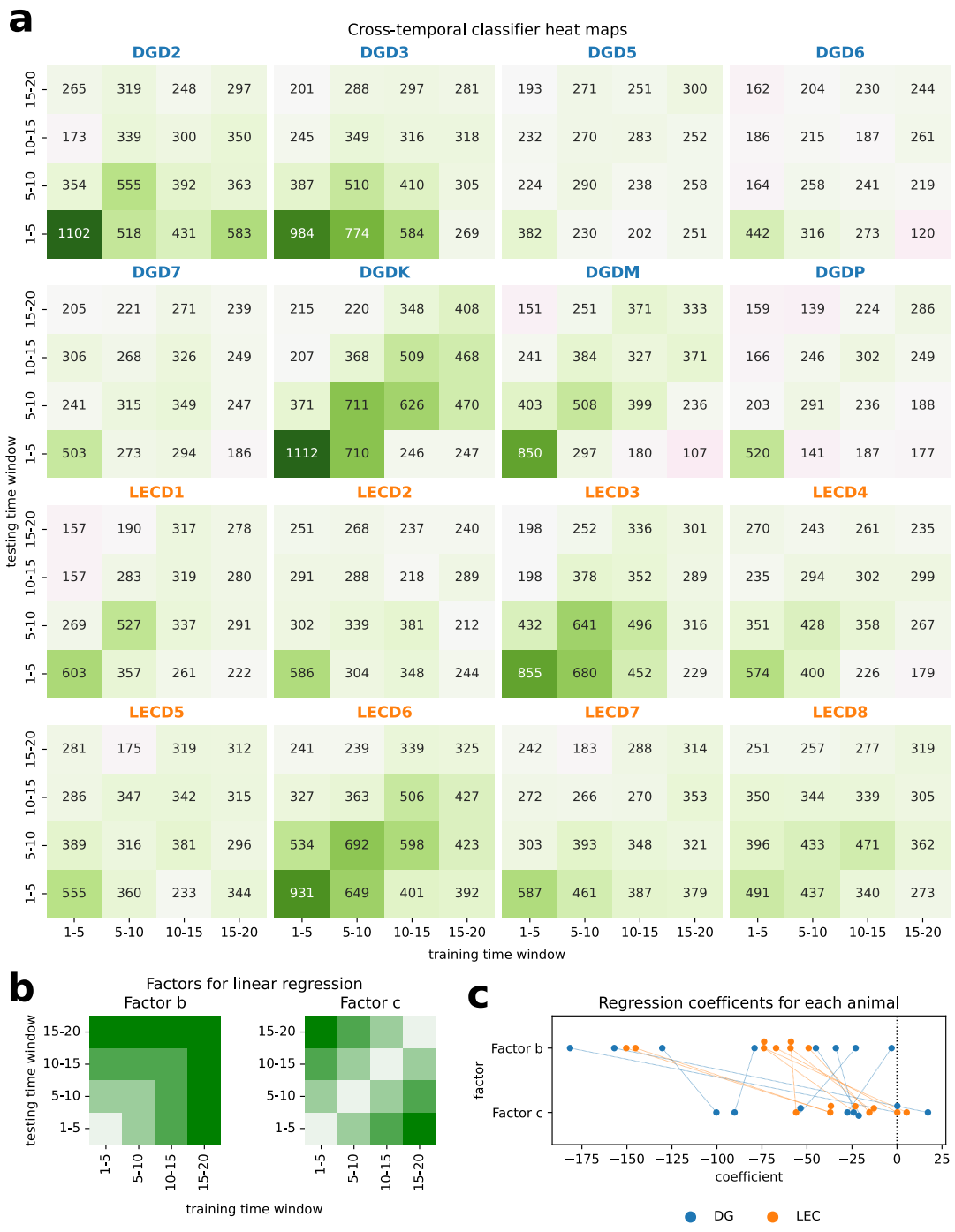

Figure 3.4: (Continued on the following page.)

Figure 3.4: **DG and LEC each conjunctively encode item and time. (a)** Cross-temporal classifier heat maps for 16 animals. The numbers are scores based on 10-fold cross-validation repeated 10 times, for a full score of 1190. Since there are 6 possible items, chance is 198. Note that each cell is derived from 100 classifiers. The horizontal axis is the training time window and the vertical axis is the testing time window. Classifiers are trained on population vectors time-averaged over entire time windows. The heat maps are consistent with time–item conjunctive encoding because accuracy peaks along the diagonal and decreases as one moves away from the diagonal. **(b)** To quantify the patterns in the accuracy values, for each animal, the 16 accuracy scores are fit as the response variable of a linear model with the two factors shown. A negative coefficient for Factor b suggests the influence of a decaying item signal (Fig. 3.3b). A negative coefficient for Factor c suggests the influence of a smoothly changing item representation (Fig. 3.3c). **(c)** The regression coefficients for Factors b and c tend to be negative, for both the group of eight DG and the group of eight LEC animals, suggesting that the cross-temporal classifier accuracies exhibit patterns corresponding to a decaying item signal and a smoothly changing item representation. (One-sample Wilcoxon signed-rank tests, DG–Factor b, $p < 0.05$, LEC–Factor b, $p < 0.05$, DG–Factor c, $p < 0.05$, LEC–Factor c, $p < 0.05$.)

Item identity offset	DG accuracy	DG corrected p-value	LEC accuracy	LEC corrected p-value
−4 prev	16.4%	1.0	17.3%	1.0
−3 prev	17.1%	1.0	20.7%	1.0
−2 prev	25.0%	.31	25.0%	.31
−1 prev	32.7%	$< .001$	30.1%	.005
0 current	100%	$< .001$	89.5%	$< .001$
+1 next	20.4%	1.0	17.7%	1.0
+2 next	24.1%	.58	20.5%	1.0

Table 3.1: DG and LEC supermice encode previous item identity. For accuracy, chance is 16.7%.

The −1 decoders were evaluated on their ability to decode the previous trial. These decoders were able to decode the previous stimulus identity above chance, based on the inferred calcium events time-averaged over 21–40 s after item onset. This suggests that encoding of the previous item in the DG and LEC is maintained beyond 20 s. The p-values are based on binomial tests with success probability 1/6 under the null hypothesis, with Bonferroni correction for the family of 14 comparisons.

3.3 Discussion

We have observed time–item conjunctive encoding that may emerge within minutes of first-time odor exposure in the DG and LEC. This is consistent with prior observations of time cells that do the same within minutes of first-time odor exposure in hippocampal CA1 (Taxidis et al., 2020). Similar findings apply for CA1 place cells, some of which form within minutes of exposure to a novel environment (Frank et al., 2004; Dong et al., 2021). The possible rapid emergence of time–item conjunctive encoding suggests the possibility of pre-wired time cell assemblies that are activated upon exposure to stimuli (Taxidis et al., 2020), and may constrain hypotheses regarding the cellular and dendritic mechanism of time cell formation (Frank et al., 2004; Lee et al., 2015). We note that to our knowledge, this is the first report of observations of such time–item conjunctive encoding in the DG.

What neural circuits may underlie the generation of time–item conjunctive encoding observed here in the DG and LEC? Within the brain, odors are first encoded at the olfactory bulb (OB), which is reciprocally connected with the piriform cortex and the LEC (Iravani et al., 2021; Gretenkord et al., 2019; Wouterlood and Nederlof, 1983; Liu, 2020; Imamura

et al., 2020); the piriform cortex also sends odor information to the LEC, which in turn transmits odor information to the DG (Woods et al., 2020). Like what we have shown in the DG and LEC, the OB and piriform cortex both carry slowly changing item representation over the scale of seconds (Bolding et al., 2020; Zhang et al., 2019). Recurrent networks within pyramidal cells of the piriform cortex have been credited with the emergence of slowly changing item representation in the same area in mice (Bolding et al., 2020). It is possible that these networks could drive the slowly changing item representation in the DG and LEC that we observed. Alternatively, under the TILT hypothesis, the slowly changing item representation may result from activity of an ensemble of time cells **f**, which derives from activity of an ensemble of leaky integrators known as temporal context cells **F** (Shankar and Howard, 2012, 2013; Bright et al., 2020). Temporal context cells have been identified in the monkey entorhinal cortex (Bright et al., 2020). In principle, the temporal context cells that may be responsible for the slowly changing item representation in the aforementioned regions, may be located in the OB, piriform cortex or elsewhere in the olfactory cortex. Further studies on these regions would clarify this possibility. Involvement of the olfactory system in the generation of time cell sequences would explain the observation that odors bias time perception (Yue et al., 2016; Millot et al., 2016; Zhou et al., 2017).

In addition, we have observed that item information in the DG and LEC persists across trials. It is known that hippocampal responses to stimuli are modulated by whether the item has ever been previously seen (e.g., Ben-Yakov et al., 2014; Larkin et al., 2014), up to 10 h prior (Fried et al., 1997), indicating that neural responses reveal information about item history. It is also known from experiments involving visual tasks that neural activity in

various cortical regions is modulated by the reward and choice history of recent trials, up to 30 s prior, even after new trials begin (Bernacchia et al., 2011; Morcos and Harvey, 2016). In this study, we found that even in the absence of tasks and rewards, item identity associated with the previous trial can be directly decoded from the neural responses averaged over a given trial, even though this response occurs after a new item emerges. This indicates modulation of neural responses due to items presented specifically within a narrow window of time (21–40 s) in the past. This neural response may play a role in working memory (Leszczynski, 2011) or efficient coding (Brenner et al., 2000). Moreover, the response may arise from time cells with time constants on the range of 21–40 s, which have been reported in other brain regions (Mello et al., 2015; Bolkan et al., 2017).

These results are at odds with the proposition that the TD RL framework, traditionally associated with midbrain dopamine neurons, could be at work generally in the brain. Key to the TD RL framework is the notion of states, which encode all there is to know about the environment that will influence the future (this is the "Markov assumption"). In most applications of the TD RL framework to neuroscience, the state is denoted by the time elapsed since the (presumably most recent) cue. Under the TD RL framework, information about the previous items is superfluous. In the results we have seen in this section, information about at least the previous item is present in the DG and LEC.

The presence of such information does not falsify the TD RL framework necessarily. For instance, the observation that the previous item remains encoded, may be accommodated by expanding the definition of the RL state to include the previous item. However, our findings open the door to alternative frameworks that take advantage of information about previous

items. One such framework is the TILT framework, which maintains, for each possible event, an array of leaky integrators with a range of time constants. From this array of leaky integrators, fuzzy information about the timing of past events can be recovered through the inverse Laplace transform (hence the name *timing using inverse Laplace transform*). Information about past events is present as stimulus-specific time cells, which collectively encode a timeline of the past. The validity of the TILT framework implies the presence of time-item conjunctive representation of memory. The results of this section affirms the presence of time-item conjunctive memory in the DG and LEC and are thus consistent with the tenets of the TILT framework.

In the results we have presented, we have analyzed the data at a population level. To extend these analyses, one may perform single-cell analyses to detect stimulus-specific time cells or temporal context cells directly. In performing these single-cell analyses, one may use the raw calcium traces, instead of the deconvolved traces, as we have. The CNMF-E procedure estimates the decay time constant separately for each cell, then performs the deconvolution using a kernel sized according to the estimated time constant (Zhou et al., 2018). If the intrinsic activity of the cells decay with a range of time constants, the deconvolution algorithm would obscure this. Using the raw calcium traces preserves the time constants of each cell. The distribution of time constants for the DG and LEC can be used to differentially characterize the roles of each region.

Under the TILT framework, information about cue–outcome relationships can be used together with information about past events to generate a timeline of the future. In the next section, we build out a specific mechanism to do this.

3.4 Method

3.4.1 Animal Behavior, Imaging and Analysis

The experimental design is described by Woods et al. (2020). Briefly, surgeries were performed to inject a virus carrying a genetically encoded calcium indicator, GCaMP6f, and implant a gradient-index lens into the brain regions of interest in adult male C57BL/6J mice initially at 8–12 weeks of age. After recovery of at least two weeks, two-photon calcium imaging experiments were performed. Mice were fitted with a nose cone with a tubing insert to deliver odors and a gentle vacuum to evacuate odors. An imaging session consisted of 20 back-to-back blocks of 6 trials. Each trial consisted of a 4 s odor presentation period and a 16 s delay period. Within each block, the six possible odors were presented once each, in random order. The total session length was about 40 min. Inferred calcium events, serving as a correlate of neural activity, were obtained through the constrained non-negative matrix factorization for microendoscopic data (CNMF-E) algorithm (Zhou et al., 2018). Each imaging time bin lasted for 0.27 s. Throughout this analysis, data for the last trial was not considered, due to incomplete imaging data for the 16 s delay after the presentation of the final odor. The numbers of regions of interest for the animals in our analysis are stated in Table 3.2.

3.4.2 Time Encoding

Analyses were done in the Python package scikit-learn. Time within trial was evenly divided into five quintiles (0–4 s, 4–8 s, etc.). Within each trial, the calcium events were

DG animal code	Number of regions of interest	LEC animal code	Number of regions of interest
DGD2	352	LECD1	172
DGD3	78	LECD2	57
DGD5	102	LECD3	83
DGD6	138	LECD4	164
DGD7	105	LECD5	87
DGK	374	LECD6	70
DGM	259	LECD7	34
DGP	381	LECD8	70
Total	1789	Total	737

Table 3.2: Numbers of regions of interest for the animals in our analysis. Codes refer to animals.

time-averaged over each quintile. For the decoder, the average calcium activity of each putative neuron was used to predict the quintile number associated with that activity. Linear discriminant analysis (LDA) with Ledoit–Wolf automatic shrinkage was used for the decoder. To obtain predictions, five-fold stratified cross-validation was used without shuffling. To test for significance, a permutation test with 4,999 permutations was conducted. The negative mean absolute error of predictions of quintile number was used as the test statistic. To sample the test statistic under the null hypothesis, the labels for quintile number were shuffled. The p-value, before correction, is given by $(C+1)/(n+1)$, where C is the number of permutations whose test statistic was equal to or larger than the observed value without label shuffling, and n is the total number of permutations. Bonferroni correction was performed for the family of 16 p-values.

3.4.3 Time–Item Conjunctive Encoding

We addressed the presence of time–item conjunctive encoding by using a cross-temporal decoder that used cell activation (inferred calcium events) to determine the item identity. We segmented each trial into several time bins, and time-averaged the inferred calcium events within each time window. Using LDA, we trained decoders to decode the odor associated with the trial, based on population activity (averaged over segment i centered at time t_i), and tested the decoder's performance on population activity (averaged over segment j centered at time t_j) on a held-out set of trials. This analysis yielded decoding accuracies as a function of i and j.

LDA separates population activity along a hyperplane that is orthogonal to a dimension that maximizes the distance between the means of each class while minimizing variance within each class. Analytically, LDA is derived with the assumption of multivariate normality and identical covariances for every class, which do not apply for our data. Nonetheless, LDA provides good results in practice. Importantly, due to its simplicity, LDA is resistant to overtraining. Since the number of features (cells) for most subjects exceeded the number of trials, the covariance matrix estimate Σ was not invertible, which is a prerequisite of LDA. To address this, we used the estimate

$$\Sigma' = (1-\lambda)\Sigma + \lambda I,$$

where I is the identity matrix and λ is a regularization parameter given by Ledoit and Wolf (2004).

The LDA classifiers could only combine the information from different features (cells) linearly, since the boundaries between classes are linear. The analysis would produce different results depending on the presence of time–item conjunctive representation. Under the model without conjunctive representation, a set of cells encode time and another set of cells encode item identity. In an LDA cross-temporal decoder for item, the time-only cell dimensions would be noise whereas the item-only cell dimensions would be consistently discriminative irrespective of i and j. The variance of the noise may certainly depend on i and j, but the mean is always zero because the cells do not represent items. Therefore, under the model without conjunctive representation, the boundaries between classes would be constant with time within trial, and there would be no performance penalty if the training time window i and the testing time window j are different. On the other hand, under the model with conjunctive representation, individual cells encode both time and item. In other words, the population code for each item would vary (smoothly) with time within trial. In an LDA cross-temporal decoder for item, performance would be highest when the training time window coincides with the testing time window, that is, when $i = j$, along the diagonal of the accuracy matrix. Performance would degrade (smoothly) as $|i - j|$ increases, as we move away from the diagonal, since the class boundaries are changing (smoothly) with time within trial.

Analysis using the same method (reviewed in King and Dehaene, 2014) has been performed so far on monkey hippocampus and lateral prefrontal cortex, showing time–item conjunctive encoding in both regions (Cruzado et al., 2020; Tiganj et al., 2018). This method has also been applied elsewhere (e.g., Stokes et al., 2013; Astrand et al., 2015; see also

Crowe et al., 2010).

3.4.4 Linear Regression for Time–Item Conjunctive Encoding

To test for significance, for each animal, we used the following linear model for the accuracy as a function of training and test time:

$$y^{\alpha}\left(w_{\text{train}}, w_{\text{test}}\right) = \beta_0^{\alpha} + \beta_1^{\alpha} x_1\left(w_{\text{train}}, w_{\text{test}}\right) + \beta_2^{\alpha} x_2\left(w_{\text{train}}, w_{\text{test}}\right) + \text{error},$$

where α indexes the animals in the experiment; y^{α} are the cross-temporal decoder accuracies; $w = 1, 2, 3, 4$ corresponds to the 1.1–5.6 s, 5.6–10.4 s, 10.4–15.0 s and 15.0–19.8 s time windows respectively; $x_1\left(w_{\text{train}}, w_{\text{test}}\right) = \max\left(w_{\text{train}}, w_{\text{test}}\right)$ corresponds to the factor associated with decay in item information with time; $x_2\left(w_{\text{train}}, w_{\text{test}}\right) = \left|w_{\text{train}} - w_{\text{test}}\right|$ corresponds to the factor associated with slow changes in item representation; and $\beta_0^{\alpha}, \beta_1^{\alpha}, \beta_2^{\alpha}$ are constant parameters. The time windows were derived from the calcium imaging frame rate and are hence not round numbers. We fit the model for each animal using ordinary least squares, based on the 16 observations corresponding to the elements of the cross-temporal decoder accuracy matrix (Fig. 3.4a). The factors x_1 and x_2 are depicted in Fig. 3.4b. The factor $-x_1$ decreases with the training or testing time window, whichever is later, since the later time window would be the limiting factor for the accuracy if the item signal were to decay with time. The factor $-x_2$ decreases with time between the training and testing time window, corresponding with degraded accuracy as the training and testing time window becomes far apart, if the item representation were smoothly changing with time. We evaluated whether the location of the set $\left\{\beta_1^{\alpha}\right\}$ for the DG animals was significantly different from zero using

a one-sample Wilcoxon signed-rank test, which is the non-parametric counterpart to the one-sample t-test. The location of the set $\{\beta_1^\alpha\}$ being negative, supports the hypothesis of decay in item information with time. We did the same for the set $\{\beta_2^\alpha\}$. The location of the set $\{\beta_2^\alpha\}$ being negative, supports the hypothesis of slow changes in item representation. We repeated the Wilcoxon tests for the LEC animals.

3.4.5 Previous Item Identity

We built two supermice (one for LEC and one for DG) to probe previous item encoding beyond 20 s. For the supermice, we combined the cells across different subjects. Since the item series were different for each subject, we merged the trials on item identity (at the relevant offset: current, previous, and so on). We preserved the relative order of trials. For example, we merged the earliest Item 1 trial of Subject 1 with the earliest Item 1 trial of Subject 2; these two trials, in general, were of different trial numbers (out of 114). We then trained an LDA decoder with Ledoit–Wolf automatic shrinkage to decode the item at the relevant offset, based on inferred calcium events averaged over the 1.1–19.8 s time window of the trial. Decoder accuracy was evaluated with 10-fold cross-validation. Binomial tests were performed on the accuracy scores with success probability $1/6$ under the null hypothesis. Bonferroni correction was performed for the family of 14 accuracy scores. The data for the final block of six trials was not included, due to incomplete imaging data.

We were mindful of the autocorrelation in the series of items. Because the item series was blocked, items were unlikely to repeat back-to-back. Thus, guessing that the current stimulus was any other item than the previous stimulus, would have spuriously improved

performance. That is, decoding of the current stimulus may mediate performance when evaluated on the previous item identity. There were five possible previous stimuli, but since the number of trials was small, some stimuli pairs were more likely than others by chance, enhancing this effect in some subjects.

Our method of analysis mitigated the spurious effects due to autocorrelation in the series of items. Because the item series were different for each subject, the items other than the one at the merged position would be different for different subjects. This mitigated the problems associated with some item pairs being more prevalent than others by chance. We also trained decoders for offsets +1 and +2 to decode item identity of the next trial and the trial after. Due to causality, we expect chance performance. We expect the autocorrelation enhancement to be similar for offsets -1 and $+1$, so this served to inform whether the autocorrelation was sufficient to lead to above-chance supermouse decoder performance in the absence of signal.

4 THEORETICAL STUDY:

PREDICTING THE FUTURE WITH A SCALE-INVARIANT TEMPORAL MEMORY FOR THE PAST

Wei Zhong Goh[1], Varun Ursekar[2], Marc W. Howard[2,3]

[1]Graduate Program in Neuroscience,

[2]Department of Physics,

[3]Department of Psychological and Brain Sciences, Center for Systems Neuroscience, 610 Commonwealth Avenue, Boston University.

Goh, W. Z., Ursekar, V., and Howard, M. W. (In press). Predicting the future with a scale-invariant temporal memory for the past. *Neural Computation*.

Abstract

In recent years it has become clear that the brain maintains a temporal memory of recent events stretching far into the past. This paper presents a neurally-inspired algorithm to use a scale-invariant temporal representation of the past to predict a scale-invariant future. The result is a scale-invariant estimate of future events as a function of the time at which they are expected to occur. The algorithm is time-local, with credit assigned to the present event by observing how it affects the prediction of the future. To illustrate the potential utility of this approach, we test the model on simultaneous renewal processes with different time scales. The algorithm scales well on these problems despite the fact that the number of states needed to describe them as a Markov process grows exponentially.

4.1 Using Memory to Predict the Future

Reinforcement learning (RL) models that are designed for Markov processes (e.g., Watkins and Dayan, 1992; Sutton, 1988) have been extraordinarily successful in accounting for reward systems in the brain (e.g., Schultz et al., 1997; Waelti et al., 2001) and led to remarkable achievements in artificial intelligence (e.g., Mnih et al., 2015; Silver et al., 2018). Despite the success of RL, its affinity for Markov statistics may be a serious limitation. The real world contains many distinct causes that predict their effects at a range of time scales, presenting a challenge for learners optimized for Markov statistics. Of course, random processes with memory can be turned into Markov processes at the cost of defining additional states. However, the cost in terms of memory, and time to learn transition probabilities among an exponentially growing number of states, may be excessively costly in some settings.

It has been proposed that a primary function of the mammalian brain is to predict future events to enable adaptive behavior (Clark, 2013; Friston, 2010). Evidence from neuroscience has made clear that the brain contains robust memory for the identity and time of recent events extending well into the past. For instance, sequentially activated time cells in the hippocampus, prefrontal cortex, and striatum (e.g., MacDonald et al., 2011; Tiganj et al., 2018; Mello et al., 2015) maintain information about the time at which recent events were experienced over at least tens of seconds, and perhaps much longer. Experimental presentation of distinct stimuli triggers different sequences of time cells (e.g., Tiganj et al., 2018; Taxidis et al., 2020; Cruzado et al., 2020) so that these populations maintain information about what happened when. In addition to sequentially activated time

cells, neurons in the entorhinal cortex (Tsao et al., 2018; Bright et al., 2020) and other cortical regions (Bernacchia et al., 2011; Murray et al., 2017) carry temporal information *via* populations of neurons that respond with a spectrum of characteristic time scales, in some cases up to at least minutes (Tsao et al., 2018). This paper, inspired by work arguing that conditioning results from an attempt to learn temporal contingencies between stimuli (Balsam and Gallistel, 2009; Gallistel et al., 2019), presents a formal model that learns to predict the future given a temporal record of the past. This proposed mechanism is computable given a temporal history that can be translated in time and proposes a solution for how to estimate the future from a past that includes information about many past events.

This paper proceeds as follows. In the rest of this section, we review a model for retaining a record of past events, and associations between event pairs. In Section 4.2, we present the model for predicting the future given a temporal record of the past. In Section 4.3, we discuss its computational complexity, time scale invariance and several other properties. In Section 4.4, we present a numerical demonstration of the efficacy of this algorithm. Finally, in Section 4.5, we compare this algorithm to traditional RL algorithms, and point out its connections to neuroscience.

4.1.1 A Formal Model for Temporal Record of the Past

We start with an agent which is capable of observing and remembering several types of events, such as the onset of a 440 Hz tone or the appearance of an image of an apple. In this section, we will describe a model for its capabilities. We will see that the agent maintains a fuzzy timeline of past events, which it uses to make pairwise associations between events.

Neurobiological justification for this model is outlined in Section A.1 of the Appendix.

4.1.1.1 Events in Continuous Time

We assume that the world provides a series of discrete events that occur in continuous time. For simplicity, without loss of generality, suppose there are three types of events, which we call X, Y and Z respectively. Whenever we need to avoid confusion, we will use *event type* to refer to type of event, and use *event episode* to refer to an individual occurrence of an event. We encode the occurrence of the event type X as a signal $f^{\mathrm{X}}(t)$, which is the sum of Dirac delta functions centered at the occurrence times of episodes of X (Fig. 1a). (We will discuss quantities in relation to X; such statements hold analogously for Y and Z.) We call t, the argument for the signal $f^{\mathrm{X}}(t)$, *real time* or *external time*, to emphasize that this time axis is a feature of the world instead of being constructed by the observer. We denote the collection of all three signals as $\mathbf{f}(t)$, and analogously for the quantities to follow. At every instant in (external) time t_0, the agent has direct access to $\mathbf{f}(t_0)$ (which is zero precisely unless the event of interest occurs at t_0), but not $\mathbf{f}$ at any other time value. Signals are shown in Fig. 4.1a in the case where X, Y and Z occur at times 0, 1 and 2 respectively.

4.1.1.2 Temporal Memory

At every instant in time t_0, the agent's memory for X, denoted $\tilde{f}^{\mathrm{X}}(\overset{*}{\tau};t_0)$, is a fuzzy representation of the signal up to the present, $f^{\mathrm{X}}(t_0-\overset{*}{\tau})$. From the agent's perspective, the internal past time, $\overset{*}{\tau}>0$, indexes how long ago events in memory might have occurred. The degree of fuzziness of the memory varies inversely with a sharpness parameter k, which is typically

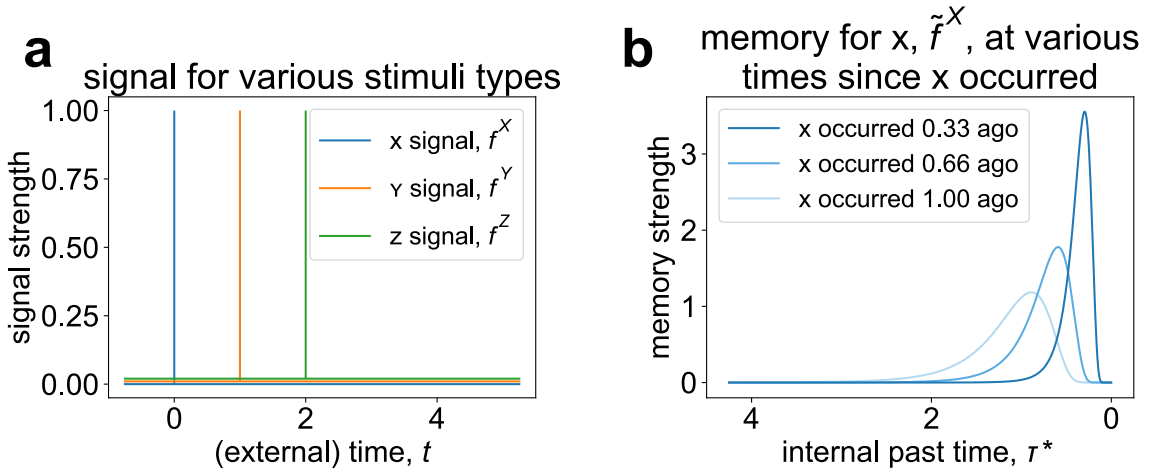

Figure 4.1: **Memory is a fuzzy representation of the signal up to the present.** **(a)** Signal as a function of external time, for three event types, X, Y and Z. This is the scenario considered in Fig. 4.2 through 4.6. **(b)** Memory for a recent event as a function of internal past time, at varying (external) times since the event occurred. As a function of internal past time, peaks in the memory are present at approximately the time interval since the event.

a small even integer; throughout this paper, it is fixed at 8.

Consider an event that exactly happens at a particular time, τ_0. At time $\tau_0 + t$, the memory element for that event is given by $\tilde{f}(\overset{*}{\tau}; \tau_0 + t) = \Phi_k(t/\overset{*}{\tau})/\overset{*}{\tau}$, where the fuzziness, $\Phi_k(\cdot)$, is given by the dimensionless equation

$$\Phi_k(x) = u(x)\kappa_0 x^k e^{-kx}, \tag{4.1}$$

$\kappa_0 = k^{k+1}/k!$ is a normalizing constant and u is the unit step function. Memories for a recent event are shown in Fig. 4.1b for various values of t. For an arbitrary signal $\mathbf{f}$, the associated memory up to time t is

$$\tilde{\mathbf{f}}(\overset{*}{\tau}; t) = \frac{1}{\overset{*}{\tau}} \int_{-\infty}^{t} \mathbf{f}(\tau)\, \Phi_k\left(\frac{t-\tau}{\overset{*}{\tau}}\right) d\tau. \tag{4.2}$$

(The origin for Eqs. 4.1 and 4.2 is outlined in Sec. A.1 in the Appendix.) In other words, the memory for an event type is the sum of the memory elements associated with each episode of that event type. On its face, Eq. 4.2 appears to assume that the agent has access to the infinite past of $\mathbf{f}(t)$. However, previous work has shown that $\tilde{\mathbf{f}}(\overset{*}{\tau};t)$ can be efficiently and time-locally constructed from a set of leaky integrators with a spectrum of time constants (see Section A.1 in the Appendix;Shankar and Howard, 2013). Using this approach, the number of leaky integrators necessary to remember the past to some bound T goes up like $\log T$. Previous papers (e.g., Shankar and Howard, 2013) using this formalism have made explicit use of this property in choosing to sample values of $\overset{*}{\tau}$ on a logarithmic scale and logarithmically compress the $\overset{*}{\tau}$ axis in integrals. In this paper, we do not logarithmically compress the $\overset{*}{\tau}$ axis in integrals. However, one may adopt an alternative interpretation, consistent with this paper, as follows. Within this alternative interpretation, the expressions for integrals over $\overset{*}{\tau}$ used below are logarithmically compressed (i.e., a factor of $1/\overset{*}{\tau}$ is added to the integrands). At the same time, the prefactor of $1/\overset{*}{\tau}$ is removed from Eq. 4.2. Neurally, this would amount to a statement that the peak firing rate of time cells triggered by a delta function is constant as a function of $\overset{*}{\tau}$.

The signal $\mathbf{f}$ up to any given external time t_0 fixes the event occurrence history. However, due to the agent's fuzzy memory, the agent is only able to form a fuzzy subjective belief distribution about the event occurrence history leading up to the present. We may interpret the memory for X as the agent's subjective estimate of the instantaneous rate of occurrence

of X at time $t - \overset{*}{\tau}$. In other words, we have, for an infinitesimal time element $d\overset{*}{\tau}$,

$$\tilde{f}^{\mathrm{X}}(\overset{*}{\tau};t)\, d\overset{*}{\tau} \approx P\left(\mathrm{X} \ @\ t - \overset{*}{\tau}\ (d\overset{*}{\tau})\right), \tag{4.3}$$

where $P(\cdot)$, the probability of an event, is used in the subjective Bayesian sense to describe the agent's belief, and "X @ $t - \overset{*}{\tau}\ (d\overset{*}{\tau})$" stands for "an episode of event X occurred within the infinitesimal time interval between $t - \overset{*}{\tau}$ and $t - \overset{*}{\tau} + d\overset{*}{\tau}$." Since $\tilde{\mathbf{f}}$ allows the agent access to the identity of and approximate time at which past events might have happened, we describe $\tilde{\mathbf{f}}(\overset{*}{\tau})$ to be a timeline of the past.

At each instant in time t, the agent is also able to compute the state of the memory a time interval δ into the future, assuming that no events of interest occur during that interval. For an arbitrary signal $\mathbf{f}$, this quantity is given by

$$\tilde{\mathbf{f}}_{\delta}(\overset{*}{\tau};t) = \frac{1}{\overset{*}{\tau}} \int_{-\infty}^{t} \mathbf{f}(\tau)\ \Phi_k\left(\frac{t+\delta-\tau}{\overset{*}{\tau}}\right) d\tau. \tag{4.4}$$

Translation can be efficiently implemented based on the set of leaky integrators. Prior work has shown that this can be done in a neurobiologically reasonable way see (Sec.~A.2 in the Appendix; Shankar et al., 2016).

4.1.1.3 Estimating Pairwise Time-Lagged Statistics

Many models of memory make use of associations between the temporal context describing the recent past and the currently available stimulus. The agent described here builds pairwise associations from X (the cue) to Y (the outcome) as the average state of memory for X

whenever Y occurs, and analogously for other event pairs:

$$\Delta M_{\mathrm{X}}{}^{\mathrm{Y}}(\overset{*}{\tau}) \propto \tilde{f}^{\mathrm{X}}(\overset{*}{\tau};t) f^{\mathrm{Y}}(t). \tag{4.5}$$

We denote the collection of pairwise associations between event pairs as $\mathbf{M}(\overset{*}{\tau})$, which may be thought of as an $n \times n$ matrix at every $\overset{*}{\tau}$, where n is the number of possible events. We denote the collection of pairwise associations with X as the cue as $\mathbf{M}_{\mathrm{X}}(\overset{*}{\tau})$, which may be thought of as a vector with n elements, one for each possible outcome, at every $\overset{*}{\tau}$.

Note that as a neural network, Eq. 4.5 simply requires Hebbian learning. At the end of learning, we normalize $\mathbf{M}_{\mathrm{X}}$ by the number of episodes of the cue X, $\int f_{\mathrm{X}}(t)\,dt$.

For example, suppose that X always precedes Y by a time interval τ_{XY}. Then, by the end of learning, we would have the pairwise association

$$M_{\mathrm{X}}{}^{\mathrm{Y}}(\overset{*}{\tau}) = \Phi_k(\tau_{\mathrm{XY}}/\overset{*}{\tau})/\overset{*}{\tau} \tag{4.6}$$

Fig. 4.2 shows the pairwise associations between two pairs of events, occurring 1 and 2 time units apart respectively.

We may view $M_{\mathrm{X}}^{\mathrm{Y}}$ from two complementary perspectives. Firstly, given the occurrence of Y in the present, the agent could use $M_{\mathrm{X}}^{\mathrm{Y}}(\overset{*}{\tau})$ as a subjective estimate, based on an average over occurrences of Y, of the instantaneous rate of occurrence of X at time $\overset{*}{\tau}$ in the past, that is,

$$M_{\mathrm{X}}^{\mathrm{Y}}(\overset{*}{\tau})\, d\overset{*}{\tau} \approx P\left(\mathrm{X} \text{ @ } t_{\mathrm{Y}} - \overset{*}{\tau}\,(d\overset{*}{\tau})\right). \tag{4.7}$$

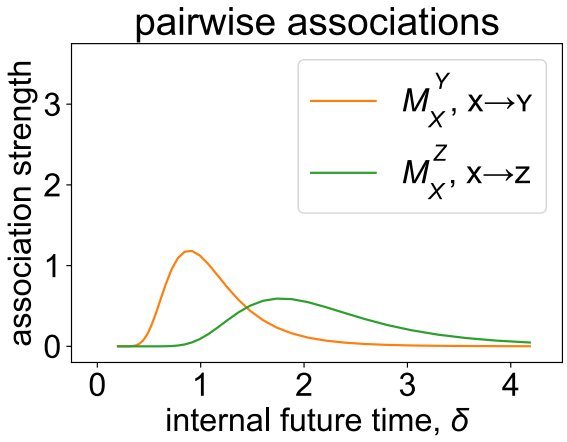

Figure 4.2: **Pairwise associations fuzzily represent the rate of finding two events occurring a certain time interval apart.** Other event types are ignored in computing the association between two event types. The associations shown here are based on the signals in Fig. 4.1a. As a function of internal time, the associations peak at around the time interval separating the event pairs.

Secondly, given the occurrence of X in the present, the agent may use $M_{\mathrm{X}}^{\mathrm{Y}}(\delta)$ as a subjective estimate of the instantaneous rate of occurrence of Y at time δ in the future, that is,

$$M_{\mathrm{X}}^{\mathrm{Y}}(\delta)\, d\delta \approx P\left(\mathrm{Y} \text{ @ } t_{\mathrm{X}} + \delta \ (d\delta)\right). \tag{4.8}$$

We use $\overset{*}{\tau}$ or δ as the time argument for $\mathbf{M}$ according to the interpretation that applies.

A limitation of directly selecting an element of the pairwise association $\mathbf{M}$ to predict the future is that the prediction can only be based on a single cue in the present (i.e., the cue corresponding to that element). To overcome this, the agent constructs the pairwise prediction $\mathbf{m}$, which integrates cue–outcome pairwise information (encoded by $\mathbf{M}$) from multiple simultaneous cues in the present to predict the future. We estimate the rate of future

occurrences of Y based on the events in the present (time t) as

$$m^{\mathrm{Y}}(\delta;t)\,d\delta \approx P\left(\mathrm{Y}\ @\ t+\delta\ (d\delta)\right), \tag{4.9}$$

where

$$\mathbf{m}(\delta;t) = \kappa_1 e^{[\boldsymbol{\mathcal{M}}\mathbf{f}_\delta](t)} \tag{4.10}$$

is the collection of rates of outcomes (a vector with n elements, one for each possible outcome, at every δ), and κ_1 is a normalization constant whose form is given in Sec. A.3. The exponential applies element-wise. The operator $\boldsymbol{\mathcal{M}}$ is defined by

$$[\boldsymbol{\mathcal{M}}\mathbf{f}_\delta](t) = \frac{1}{|\mathcal{E}_t|}\sum_{\alpha\in\mathcal{E}_t}\int f_\delta^\alpha(\overset{*}{\tau};t)\log\mathbf{M}_\alpha(\overset{*}{\tau})\,d\overset{*}{\tau}, \tag{4.11}$$

where $\mathcal{E}_t$ is the set of events occurring at time t, $|\mathcal{E}_t|$ is the number of events occurring at time t, $\mathbf{M}_\alpha$ is the collection of pairwise associations with α as the cue (a vector with n elements, one for each possible outcome in correspondence with $\mathbf{m}$, at every δ), and

$$f_\delta^\alpha(\overset{*}{\tau};t) = \Phi_k(\delta/\overset{*}{\tau})/\overset{*}{\tau} \tag{4.12}$$

denotes the future state of the memory element associated with the currently occurring episode of α. (Contrast this with $\tilde{f}_\delta^\alpha$ (see Eq. 4.4), which denotes the future memory state induced by all past occurrences of α.) The logarithm in Eq. 4.11 applies element-wise. The operator $\boldsymbol{\mathcal{M}}$ may be thought of as operating on the pre-computed future memory state of the current events (Eq. 4.4) to generate a prediction for the future. In general, Eq. 4.8 and

Eq. 4.9 provide similar estimates for the future. The normalization constant κ_1 is such that precisely when X is the only cue for Y and the time delay between them is fixed, Eq. 4.8 and Eq. 4.9 provide exactly the same estimate for Y. Mathematical details can be found in Sec. A.3.

Intuitively, the integral that is the second term on the right hand side of Eq. 4.11 takes the product of $\log \mathbf{M}_{\mathrm{X}}$ with the future state of the memory element f_{δ}^{X}. Let us consider Y to be the outcome of interest, and consider the case where the time interval between X and Y, τ_{XY}, is constant. In this case, the integral (and thus, m^{Y}) attains a maximum as a function of δ when δ coincides with the peak of $\log M_{\mathrm{X}}^{\mathrm{Y}}(\overset{*}{\tau})$, which is approximately τ_{XY}. Fittingly, this is behavior we expect of m^{Y}.

Both $M_{\mathrm{X}}^{\mathrm{Y}}$ and the corresponding Eq. 4.11 integral are smooth functions that peak around the delay interval between the two events. This may prompt the question of why the integral is used instead of $M_{\mathrm{X}}^{\mathrm{Y}}$ in its place (e.g., Tiganj et al., 2019). The strength of the present formulation is that Eqs. 4.10–4.12 closely parallel Eqs. 4.14–4.15 in the next section, for which the integral is necessary. Using equations of similar form is more neurobiologically realistic, because it suggests that analogous neural architecture supports the computation for both pairwise prediction **m** and prediction **p**, to be introduced later.

4.2 Predicting the Future With a Scale-Invariant Past

It would be straightforward to build a prediction for the future based on a single event (e.g., the most recent event) using the pairwise associations **M**. The challenge is to build a prediction that is based on multiple events in the recent past. One difficulty arises when

associations overlap. For example, we associate the sound of rain (X) with a chance of hearing thunder (Z). We also associate the sight of wet ground (Y) with a chance of hearing thunder (Z). Having heard the sound of rain, the prediction for thunder should not be increased by the sight of wet ground when we step outdoors. This example illustrates one of the pitfalls of simply adding the predictions suggested by the pairwise associations.

To address double-counting, in addition to pairwise associations, we construct credit associations between event pairs, which is the key for this algorithm to generating a timeline for the future. In Sec. 4.2.1, we explain how the agent constructs a timeline of the future by integrating over a timeline of the past, weighted by credit associations between cues and outcomes. In Sec. 4.2.2, we show how the agent learns credit associations between cues and outcomes based on comparing predictions prior to the cue with predictions due to the cue.

4.2.1 Generating Predictions From Credit Associations

In addition to the pairwise associations $\mathbf{M}$, we build the credit associations $\mathbf{C}$ between each pair of events (a cue and an outcome) as a function of internal time δ since the cue. The credit associations $\mathbf{C}(\delta)$ may be thought of as an $n \times n$ matrix at every δ. We denote the collection of credit associations with X as the cue as $\mathbf{C}_{\mathrm{X}}(\delta)$, which may be thought of as a vector with n elements, one for each possible outcome, at every δ. We interpret $C_{\mathrm{X}}^{\mathrm{Y}}(\delta)$ as logarithm of the factor by which an agent adjusts its subjective estimate of the instantaneous rate of occurrence of Y at time δ in the future, having just observed X. Denoting $\mathbf{p}^{-}(\delta)$ as

the agent's prior estimate (just before observing X), we have

$$\left[p^{-}(\delta)\right]^{\mathrm{Y}} \exp C_{\mathrm{X}}^{\mathrm{Y}}(\delta)\, d\delta \approx P\left(\mathrm{Y} \text{ @ } t_{\mathrm{X}} + \delta \ (d\delta)\right). \tag{4.13}$$

Eq. 4.13 relates to the observation of one cue at one time (the present). For cues in the past, the further in the past they occur, the more imminent outcomes should seem. For example, if X has credit for Y peaking at $\delta = 5$ and X occurred three time units ago, Y should be expected in two time units. Accounting for multiple cues over the past, we find that at time t, the agent's internal timeline for a time δ into the future, is

$$\mathbf{p}(\delta;t) = \Lambda \odot e^{[\mathcal{C}\tilde{\mathbf{f}}_{\delta}](t)}, \tag{4.14}$$

where $\mathbf{p}$ stands for *prediction* and is a vector over event types, Λ consists of the long-term average of each event type, $\odot$ denotes the element-wise product, and the exponential applies element-wise. The operator $\mathcal{C}$ is defined by

$$[\mathcal{C}\tilde{\mathbf{f}}_{\delta}](t) = \sum_{E} \int \mathbf{C}_{E}(\overset{*}{\tau}) \tilde{f}_{\delta}^{E}(\overset{*}{\tau};t)\, d\overset{*}{\tau}, \tag{4.15}$$

where the index of summation E indexes the possible cue types, and $\mathbf{C}_{E}(\overset{*}{\tau})$ represents the collection of credit associations with E as the cue (a vector with n elements, one for every possible outcome, at every $\overset{*}{\tau}$). Intuitively, the integral sums products of $\mathbf{C}_{E}$ with the projected memory $\tilde{f}_{\delta}^{E}$. Let us consider Y to be the outcome of interest and X the only cue, and consider the case where the time interval between X and Y, τ_{XY}, is constant. In this case,

the integral (and thus, p^{Y}) attains a maximum as a function of δ at the value of δ at which the peaks of $C_{\mathrm{X}}^{\mathrm{Y}}$ and $\tilde{f}_{\delta}^{\mathrm{X}}$ coincide. The credit $C_{\mathrm{X}}^{\mathrm{Y}}$ peaks around τ_{XY}, the time delay between X and Y. The projected memory $\tilde{f}_{\delta}^{E}$ peaks around the time $\tau_{\mathrm{X}}+\delta$, where τ_{X} is the time that has elapsed since X. Therefore, the integral (and thus, p^{Y}) peaks around $\delta=\tau_{\mathrm{XY}}-\tau_{\mathrm{X}}$, which is the time remaining to Y, the outcome of interest. In other words, the agent's expectation for Y would be the highest at a time when Y is, in fact, due. Mathematical details can be found in Sec. A.4.

We interpret $p^{\mathrm{Y}}(\delta;t)$ as the agent's subjective estimate, made at time t, of the instantaneous rate of occurrence of Y at time δ in the future, that is,

$$p^{\mathrm{Y}}(\delta;t)\,d\delta \approx P\left(\mathrm{Y}\ @\ t+\delta\ (d\delta)\right). \tag{4.16}$$

Unlike Eq. 4.8 and Eq. 4.13, this estimate takes into account all of the events that have occurred in the recent past. A schematic distinguishing the utility of the pairwise associations $\mathbf{M}$ and the credit associations $\mathbf{C}$ in making predictions is shown in Fig. 4.3. Just as we consider $\tilde{\mathbf{f}}(\overset{*}{\tau})$ a timeline of the past, we consider $\mathbf{p}(\delta)$ a timeline of the future. Note that $p^{\mathrm{Y}}(\delta=0;t)$ would correspond to the agent's internal model for, in the language of point process theory, the conditional intensity function of Y (see Rasmussen, 2018).

As an illustration, consider again the scenario where events X, Y and Z always occur consecutively, once on each trial, at relative times 0, 1 and 2 respectively, with a very long gap between trials. Once X occurs, the proposed algorithm (explained in the following sections) generates predictions for Y and Z that become more and more imminent as time elapses (Fig. 4.4). As a function of δ, the predictions peak at approximately the time when

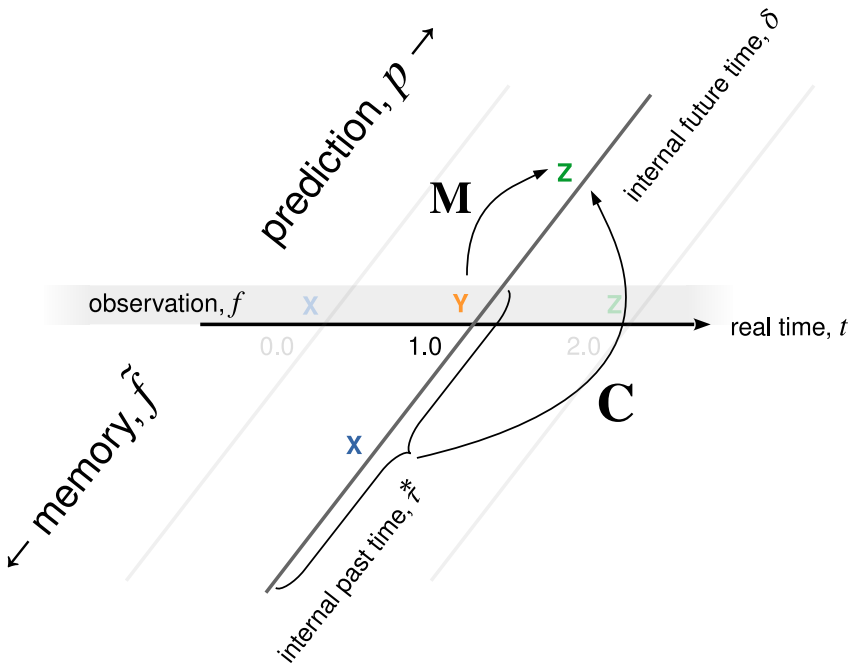

Figure 4.3: **Predictions can be made using credit associations C based on memories of the multiple events in the recent past.** The horizontal axis shows events occurring in real time. The event signal for this scenario is shown in Fig. 4.1a. Associated with each point in real time is an agent's internal time axis, shown here diagonally at $t = 1.0$, which indexes memories of the past (bottom half) and predictions for the future (top half). The agent may make a prediction for the future with **M** (Eq. 4.8) based on the currently observed event (here, Y). As a better alternative, the agent may make a prediction for the future with **C** (Eq. 4.14) based on multiple events in the present and the recent past.

the events are, in fact, due.

4.2.2 Computing Credit Associations

Loosely speaking, we assign credit for an outcome to an event according to how much the event's occurrence would revise the prediction for that future outcome. In our example, wet ground would be assigned little to no predictive value, since the chance of thunder has already been predicted by the sound of rain. During training, we update the credit assigned to an event when that event occurs. In this section, we will describe the update that happens when X occurs with no loss of generality.

Formally, as we have stated, $\exp C_{\mathrm{X}}^{\mathrm{Y}}(\delta)$ is the factor by which we should adjust the prediction for Y at time δ in the future, having just observed X. Therefore, to compute $\exp C_{\mathrm{X}}^{\mathrm{Y}}(\delta)$, whenever X is observed, we will first compute the prediction for Y before and due to the observation of X, and analogously for other possible outcomes.

4.2.2.1 Prediction Prior to Event Observation

Prior to event observation at time t, the prediction associated with internal future time δ is given simply by

$$\mathbf{p}^{-}(\delta;t) = \lim_{t' \to t^{-}} \mathbf{p}(\delta;t'). \tag{4.17}$$

This prediction arises from the memory of cues in the past, and specifically excludes the effects of what occurs at time t.

Consider the scenario in Fig. 4.1, where X, Y and Z occur consistently at trial times 0, 1 and 2 respectively. When X occurs, $(p^{-})^{\mathrm{Y}} = \Lambda_{\mathrm{Y}}$, the long-term average rate of Y, for all δ.

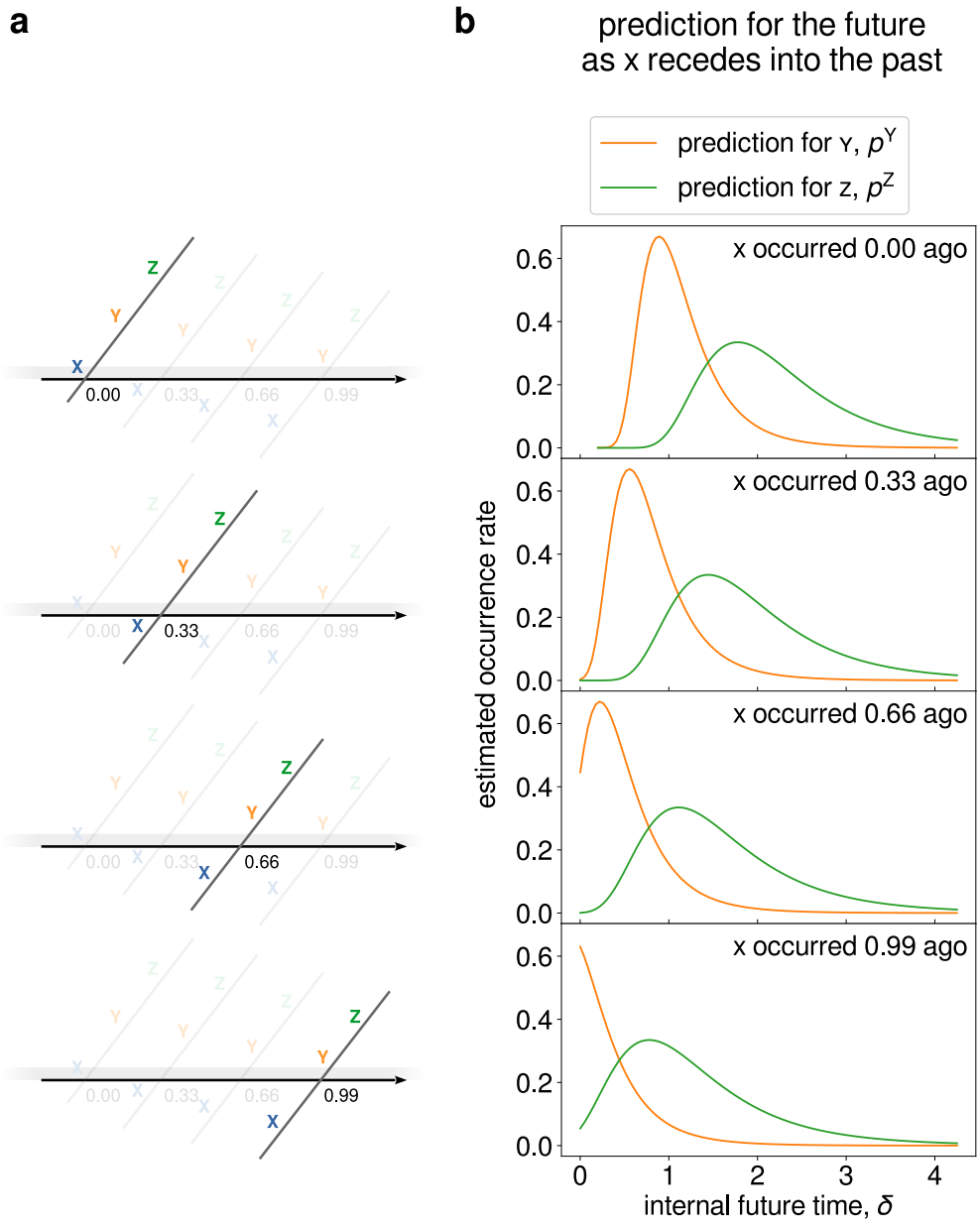

Figure 4.4: (Continued on the following page.)

Figure 4.4: **Predictions for events peak at about the right time and become more imminent with time.** The events X, Y and Z occur on each trial at times 0, 1 and 2 respectively, as with previous figures. **(a)** A schematic of the state of memory and prediction as a function of time. The axes have the same interpretation as in Fig. 4.3. At real time 0.00, X is observed, leading to a prediction for Y and Z, depicted along the diagonal internal time axis. As real time passes, the memory of X recedes into the past, and the prediction for Y and Z become more imminent, depicted by the events' downward movement along the internal time axis. **(b)** Prediction for Y and Z generated by simulation using the proposed algorithm, as the memory for X recedes into the past, depicted at four time points. The peak times for the prediction for Y and Z correspond roughly to when the events are in fact due, and move towards zero as time passes. For example, in the topmost plot, right after X occurs, Y and Z are to occur in 1 and 2 time units respectively. Indeed, the generated predictions for Y and Z peak at approximately $\delta = 1$ and 2 respectively.

This is because $\mathbf{p}^-$ is computed based on memory of events occurring before X, of which there are none (Fig. 4.5c). In contrast, when Y occurs, $(p^-)^{\mathrm{Z}}$ shows a peak at $\delta = 1$, based on memory of events occurring before Y (i.e., X), and the credit association between X and Z (Fig. 4.5f).

4.2.2.2 Prediction Due to Event Observation

For the prediction due to the observed event X itself, we use the pairwise prediction in accordance with Eq. 4.9,

$$\mathbf{p}^+(\delta;t) = \mathbf{m}(\delta;t). \tag{4.18}$$

For the scenario in Fig. 4.1, when X occurs, $(p^+)^{\mathrm{Y}} = m^{\mathrm{Y}} = M_{\mathrm{X}}^{\mathrm{Y}}$ (Fig. 4.5c), and when Y occurs, $(p^+)^{\mathrm{Z}} = m^{\mathrm{Z}} = M_{\mathrm{Y}}^{\mathrm{Z}}$ (Fig. 4.5f). Both of these have the same form, peaking sharply at $\delta = 1$, since the time intervals between X and Y and between Y and Z are fixed and equal.

4.2.2.3 Updating C

When X is observed at time t, we update $\mathbf{C}_{\mathrm{X}}$ in the following manner:

$$\Delta \exp \mathbf{C}_{\mathrm{X}}(\delta) \propto \frac{\mathbf{p}^{+}(\delta;t)}{\mathbf{p}^{-}(\delta;t)} - \exp \mathbf{C}_{\mathrm{X}}(\delta). \tag{4.19}$$

The division and exponentiation are performed element-wise for each possible outcome. This update depends on the previous state of $\mathbf{C}$ (through Eq. 4.17 and Eq. 4.14). During training, as events occur, we update respective components of $\mathbf{C}$, which in turn enhances the agent's predictions of the future as training proceeds. This update rule squares with the intuition that events be assigned credit in accordance with their association with outcomes that are not previously predicted. As training proceeds, $\exp \mathbf{C}_{\mathrm{X}}(\delta)$ approaches $\mathbf{p}^{+}(\delta;t)/\mathbf{p}^{-}(\delta;t)$ in expectation, up to the variability of event occurrence history in recent episodes of X during training. Since we assume stationary statistics, a small learning rate (i.e., constant of proportionality in Eq. 4.19) should be used to minimize the effects of such variability.

For the scenario in Fig. 4.1, as noted, the observation of X generates a prior prediction for Z that is present when Y occurs. Thus, via Eq. 4.19, Y receives less credit for Z than X for Y at each δ (Fig. 4.5), even though the X–Y and Y–Z pairwise associations are the same (Fig. 4.6). As a practical matter, since the learning of $\mathbf{C}$ depends on the accurate learning of $\mathbf{M}$, for best results, $\mathbf{C}$ should be learned only after $\mathbf{M}$ stabilizes during training.

4.2.3 Summary

The agent's memory $\tilde{\mathbf{f}}$ encodes a timeline of past events (Eq. 4.2). Using Hebbian association, the agent makes pairwise associations $\mathbf{M}$ between each pair of event types as a function of internal time (Eq. 4.5). This lets the agent form a pairwise prediction $\mathbf{m}$ for the future whenever events occur, but only based on the pairwise correlations associated with those events as cues. To predict future events based on past events, the agent learns credit associations $\mathbf{C}$ between each pair of event types as a function of internal time. The agent uses $\mathbf{C}$ and $\tilde{\mathbf{f}}$ to generate a timeline of future events (Eq. 4.14). While the agent learns, each time an event occurs, we step $\exp C_\alpha^\beta$ (where α is the event that occurred) towards the ratio of the prediction for β due to α (based on $\mathbf{M}$), to the prediction for β prior to α (based on $\mathbf{C}$) (Eq. 4.19). This design curbs double-counting of correlations for an outcome associated with multiple cues at different points in the past. Through learning, we expect the agent to produce better and better predictions for events in its future.

4.3 Properties of the Prediction Algorithm

The algorithm described above has interesting computational properties. We will discuss how it scales with the number of event types that can be distinguished and the time scales over which prospection is implemented. It can be shown that the model is optimal for pairwise predictions modulo the uncertainty that comes from finite temporal resolution of memory. Moreover, the model is invariant to rescaling of time, which may be useful in applications where the relevant time scale is not known *a priori*.

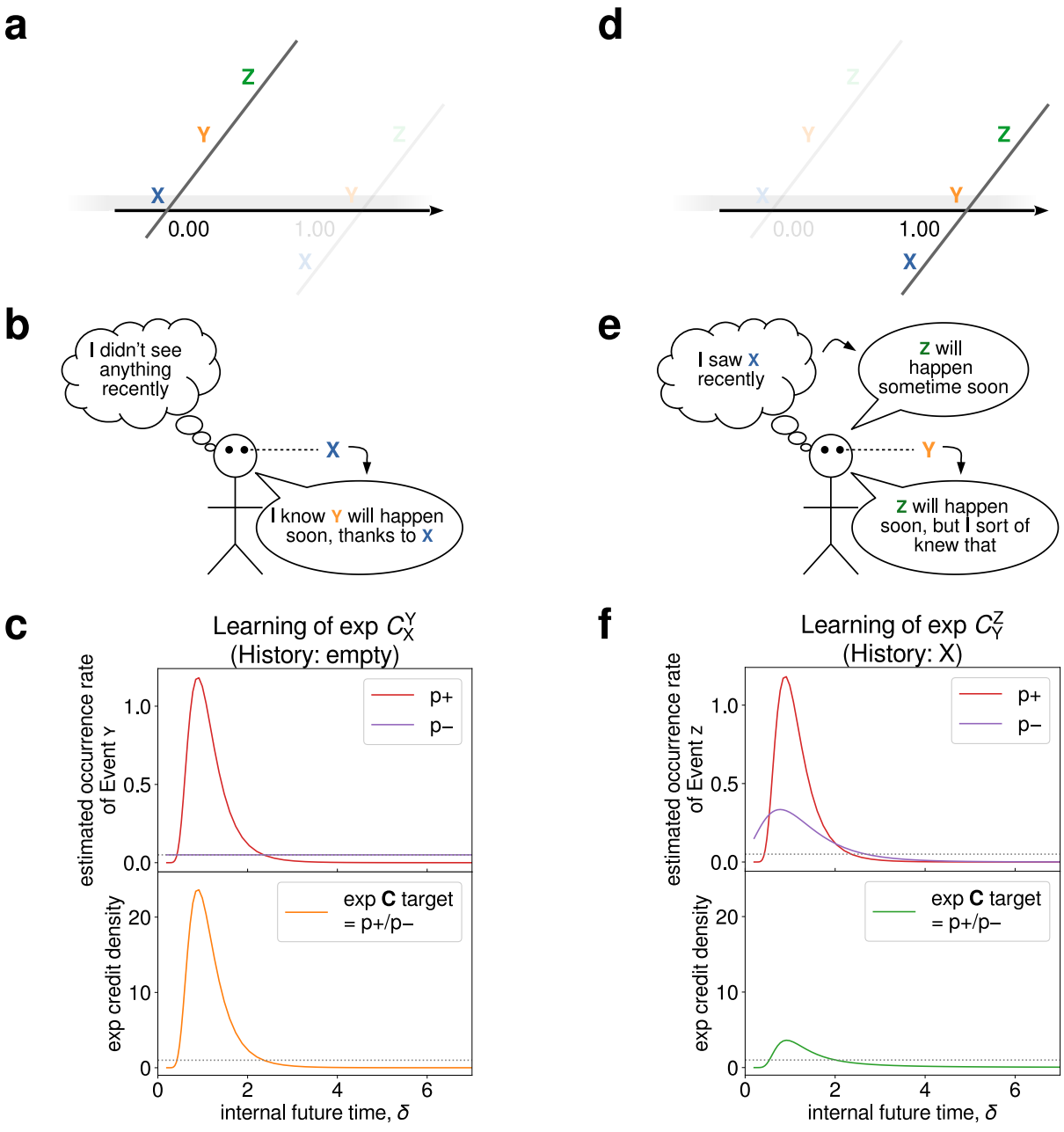

Figure 4.5: (Continued on the following page.)

Figure 4.5: **Observed events receive less credit for future events which have already been predicted based on past events.** As with all previous figures, the events X, Y and Z occur on each trial at times 0, 1 and 2 respectively. **(a)** A schematic of the state of memory and prediction as a function of time, as in Fig. 4.4(a). At real time 0.00, X is observed and the memory is empty. **(b)** An illustration of an agent's inferences at the time X occurs. No memory of past events exists to suggest a prediction, whereas the currently observed event X suggests that Y occurs soon. **(c)** Plots of $\mathbf{p}^+$, $\mathbf{p}^-$ (top) and $e^{C_{\mathrm{X}}^{\mathrm{Y}}}$ (bottom) as a function of internal future time, δ, at the time X occurs, for the prediction of Y. The quantity $\mathbf{p}^+$ (red) is the pairwise association between X and Y, while $\mathbf{p}^-$ (purple) is flat as a function of δ as there is no memory of events. The quantity $e^{C_{\mathrm{X}}^{\mathrm{Y}}} = \mathbf{p}^+/\mathbf{p}^-$ (orange). **(d)** Same as (a), but at real time 1.00. Y is observed and X is in memory. **(e)** An illustration of an agent's inferences at the time Y occurs. The agent remembers X, prompting a prior prediction of Z. The currently observed event Y suggests the same, but the agent does not gain much information from Y, and hence assigns Y less credit. **(f)** Same as (c), but at the time Y occurs, for the prediction of Z. The quantity $\mathbf{p}^+$ is the pairwise association between Y and Z, which is the same as that between X and Y. However, $\mathbf{p}^-$ reflects the prior prediction for Z based on the memory for X. (This is p^{C} from the bottommost plot in Fig. 4.4b.) Thus, $e^{C_{\mathrm{X}}^{\mathrm{Y}}}$ is diminished.

4.3.1 Scaling Properties

As with traditional associative models, the computational time and space required for this algorithm vary quadratically with the number of event types considered. In typical RL models, each state s must be defined to include all of the information that could affect the transition to the next state, in order to fit into the Markov structure. If transitions depend on the indefinite past, the number of possible states would become unwieldy. In contrast, the event types used here are economically defined to be those events that occupy a single point in time (X, Y, etc.), which are much smaller in number.

In addition, this algorithm runs in time and space polynomial in the number of $\overset{*}{\tau}$ time points considered in $\tilde{\mathbf{f}}(\overset{*}{\tau})$ and δ time points considered in $\mathbf{p}(\delta)$. For example, in Eq. 4.15, for each δ, the numerical integral is computed in time linear in the number of $\overset{*}{\tau}$, the variable of integration, corresponding to how far in the past memories are considered. The full

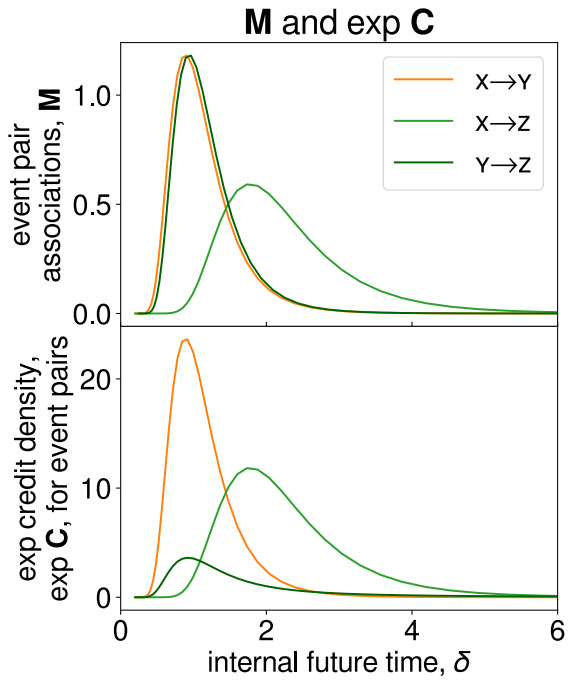

Figure 4.6: **For event pairs, credit density can differ despite having the same pairwise associations.** A summary of the event pair associations (top) and credit densities (bottom) for all nontrivial event pairs for the scenario in all previous figures, where the events X, Y and Z occur on each trial at times 0, 1 and 2 respectively. The pairwise associations $M_{\mathrm{X}}^{\mathrm{Y}}$ and $M_{\mathrm{Y}}^{\mathrm{Z}}$, overlapping perfectly, are slightly displaced for clarity. However, $C_{\mathrm{X}}^{\mathrm{Y}}$ is greater than $C_{\mathrm{Y}}^{\mathrm{Z}}$ due to the memory of X allowing a prior prediction for Z when Y occurs, as shown in Fig. 4.5(f).

prediction, over all δ that the agent considers, is computed in time linear in the number of δ. Translation to different values of δ can either be implemented serially, consistent with neural considerations (Shankar et al., 2016), or be parallelized *in silico*. The quick performance comes at the cost of the ability to directly handle some forms of joint statistics among cues. We discuss this shortcoming in Sec. 4.5.2.

The longest time scale over which predictions are based and are made increases exponentially with computational demands. Although the integral form in Eq. 4.2 would seem to require memory for the entire history up to the present, $\tilde{\mathbf{f}}$ can be generated from leaky integrators with a number of time constants (Shankar and Howard, 2013). The scale invariance of Φ_k allows us to choose the distribution of time constants as a geometric series, resulting in a logarithmic relationship between the number of integrators and the longest time scale that can be represented.

4.3.2 Equivalence of Fuzzy Memory and Input Temporal Uncertainty

Even when the time interval between events is fixed, fuzzy memory (finite k) leads to temporal fuzziness in both the pairwise association $\mathbf{M}$ and prediction $\mathbf{p}(\delta)$. At every instant in time, this induced fuzziness is equivalent, in its effect on the prediction, to fuzziness due to intrinsic temporal uncertainty in the signal f faced by an agent with perfect memory (infinite k).

As an example, consider an agent with fuzzy memory encountering X, followed by Y after a fixed time interval τ. Precisely at the time X occurs, the agent's prediction for Y is

given by

$$p^{\mathrm{Y}}(\delta;t_{\mathrm{X}}) = \frac{\kappa_1^{-1}}{\delta}\Phi_k\left(\frac{\tau}{\delta}\right), \tag{4.20}$$

where κ_1 is as given in Sec. A.3. Another agent with perfect memory encountering X, followed by Y after a random time interval τ, whose probability density function is given by $q_\tau(t) = \Phi_k(t/\delta)/\delta$, makes an optimal prediction following X equivalent to Eq. 4.20. The derivation of Eq. 4.20 is given in Sec. A.5.1.

Although the fuzzy memory agent's prediction for Y some time after encountering X is different from Eq. 4.20, this equivalence property still holds: at every instant in time, there exists a perfect memory agent, with observations subject to some density function of τ, with an equivalent optimal prediction.

4.3.3 Time Scale Invariance

The prediction algorithm inherits the time scale invariance of the temporal record of the past. If the input signals are time-dilated, the resulting predictions would be time-dilated, rescaled in magnitude and otherwise unchanged (Fig. 4.7). Therefore, the prediction algorithm, with an appropriate range of $\overset{*}{\tau}$ and δ, supports chains of events that happen over any time scale.

Formally, for any constant λ, the estimated probability of event occurrence within a small duration $d\delta$, $p(\delta;t)\,d\delta$, is invariant under the transformation

$$t \to \lambda t$$

$$\overset{*}{\tau} \to \lambda \overset{*}{\tau}$$

$$\delta \to \lambda \delta.$$

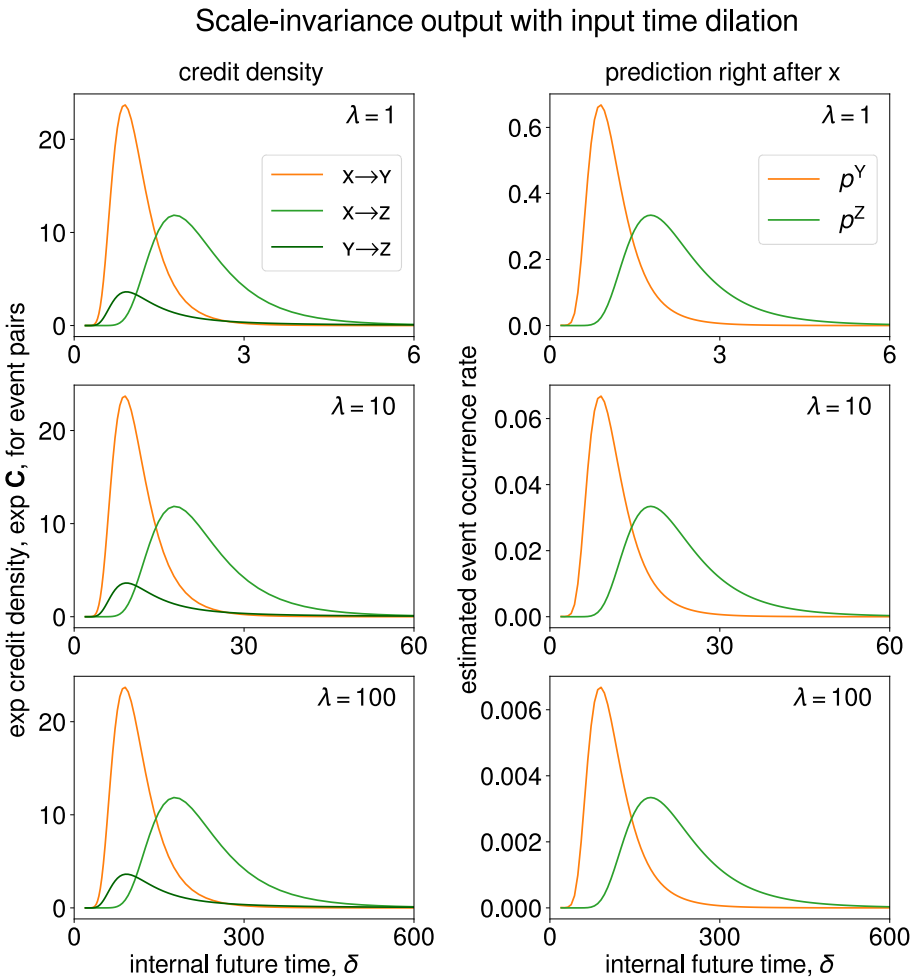

Figure 4.7: **Predictions are time scale–invariant.** Top: Credit density and, as an example, the prediction after X occurs for Y and Z, as a function of internal future time, δ, for the scenario in all previous figures. Middle, bottom: When the scenario is time-dilated, shown here by 10 and 100 times, the model output is unchanged as a function of dilated internal time. In the case of predictions, the magnitude rescales to preserve the area under the curve. This suggests that the proposed algorithm supports chains of events that occur over any time scale.

This means that within the limits of a computational implementation, i.e., far from the smallest and largest values of $\overset{*}{\tau}$ and δ (which grow exponentially with the resources committed to representing time), the model provides the same relative temporal resolution.

One may wonder whether, as an alternative to computing $\mathbf{C}$, one can generate a future timeline $\mathbf{p}(\delta)$ and directly update it using $\mathbf{p}^+$ and $\mathbf{p}^-$ whenever an event occurs. A difficulty with this approach is that a time scale would have to be chosen for the evolution of $\mathbf{p}(\delta)$ between events, violating the time scale invariance property that we desire.

4.3.4 With Fuzzy Memory, Credit Is Assigned Based on Temporal Proximity

Consider the scenario where X occurs, then Y, then Z, always with the same time delays. In the limit of perfect memory, Y would receive no credit for Z. This is because the occurrence of X would allow the time of occurrence of Z to be predicted perfectly at all times. The occurrence of Y would not improve the (already perfect) prediction. When memory is fuzzy, the X–Z pairwise association would have a larger temporal uncertainty than the Y–Z pairwise association, since Y and Z are closer in time than X and Z (Eq. 4.5). Therefore, the occurrence of Y would improve the prediction for Z. The closer Y occurs to Z, the more Y sharpens the prediction for Z, and the more credit is assigned to Y for Z. Fig. 4.8 illustrates this effect, and supporting equations are worked out in Sec. A.5.2.

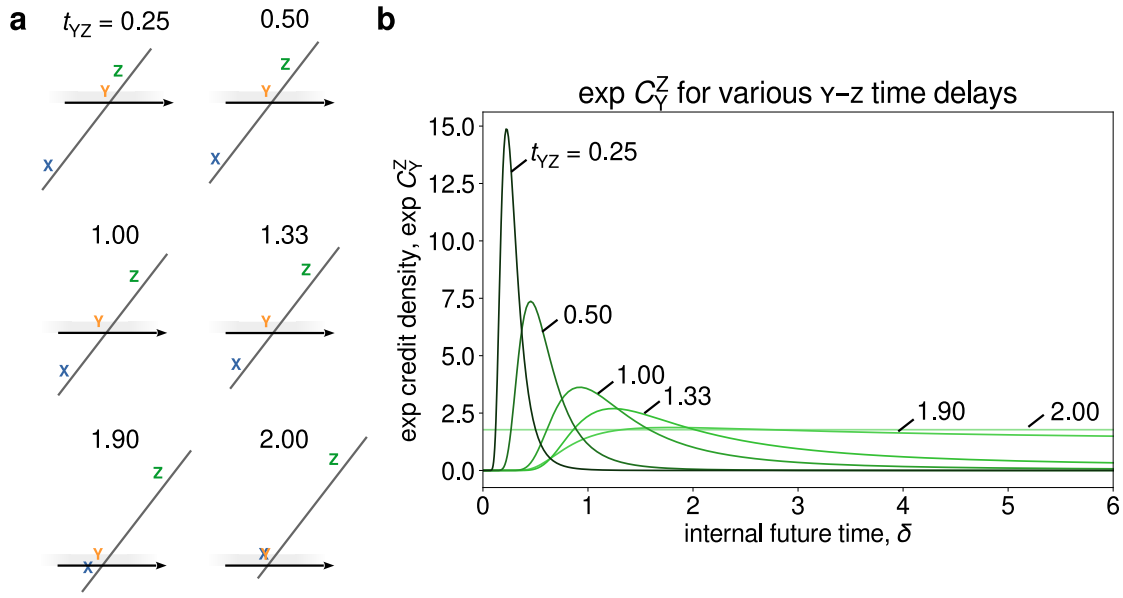

Figure 4.8: **Temporal proximity promotes credit assignment.** Events X, Y and Z occur at times 0, $2-t_{YZ}$ and 2 respectively. **(a)** A schematic of the state of memory and prediction at the time Y occurs, for six values of t_{YZ}. The axes have the same interpretation as in Fig. 4.3. **(b)** Credit assigned to Y for Z is shown here for the six values of t_{YZ}, as a function of internal future time, δ. In other words, each line represents different amounts of temporal proximity between Y and Z, while the interval between X and Z remains fixed. For $t_{YZ} = 1.9$, Y is much closer in time to X than to Z. In this case, the credit is almost flat, as the prediction for Z due to X is still fresh. The case $t_{YZ} = 1$ is the scenario in Fig. 4.1 through 4.6. For lower and lower values of t_{YZ}, credit density is more and more sharply peaked. The prediction for Z due to X has flattened out, allowing the effect of the pairwise association between X and Z to dominate. The analytic form of the lines plotted are worked out in Sec. A.5.2.

4.4 Demonstration: Event Streams With Memory and Multiple Characteristic Time Scales

We have seen that the algorithm described here is able to predict the future based on a temporally extended record of the past containing multiple possible cues. In addition, this prediction does not require selection of a preferred time scale, allowing for generalization across an exponentially large range of times. As a consequence of these two properties, this approach is well-suited to applications where the relevant time scale is not known *a priori* or to situations where there are multiple processes at different characteristic time scales that must be simultaneously learned. To illustrate these properties, we demonstrate learning of the algorithm on a time series of discrete events generated from multiple Markov renewal processes (MRP).

In principle, the algorithm we describe is capable of handling multiple cues with additive effects (but see Sec. 4.5.2) stretching into the indefinite past. However, for simplicity, we generate a scenario such that each event has exactly one cue. This cue is mostly found at most 15 time steps before the event. For comparison, most consecutive events have an intervening time of between 0.1 and 15 time units. Crucially, the cue is not usually the immediately preceding event, but one of the several preceding events. Thus, one cannot merely predict the future based on the most recent event. To add realism, we introduce a small amount of variability in the event type of the outcome, as well as a small amount of Gaussian variability in the time of the outcome.

The way we generated a scenario with such properties is to superpose several MRPs, each with three base event types, U, V, W. MRPs have the property that the type of each event

is the sole determiner of the probability distribution of the type and time of the next event. In other words, each event has a single cue. Superposing MRPs destroys the guarantee that the cue immediately precedes its outcome. We generated the scenario using two approaches, mainly differing in the way event types are determined in the superposed process. For the first approach, event types in the superposed process are determined according to the base type of the event and the MRP of origin. For example, for a superposition of 7 MRPs, there would be $3 \times 7 = 21$ event types ($1\text{U}, 1\text{V}, 1\text{W}, \ldots, 7\text{W}$). An example of such a scenario with two MRPs superposed is shown in Fig. 4.9a. Figure 4.9b shows the corresponding mean transition times for each type of transition. The drawback of this approach is that as the number of MRPs increases, the number of event types increases, making the prediction task inherently harder. For the second approach, event types in the superposed process are determined only according to the base types of the events, even if they originate from different MRPs. This way, for the prediction of the type of an event, there are always two wrong answers and one correct answer, for a fair comparison regardless of the number of MRPs superposed.

The algorithm we describe can be used to predict both the time and type of likely events in the future. However, for simplicity, we evaluate the algorithm on its average accuracy of predicting the type of the next event, given the time to the next event, whenever an event occurs. We generate this prediction *via* $\text{argmax}_i \; p^i(\delta = t_{n+1} - t_n; t = t_n)$, where t_n is the time of the nth event. We call this the $\mathcal{C}$-based prediction. As a comparison, we generate an $\mathcal{M}$-based prediction *via* $\text{argmax}_i \; m^i(\delta = t_{n+1} - t_n; t = t_n)$, where j is the type of the event at t_n, and evaluate its average accuracy. Notice that the $\mathcal{M}$-based prediction only invokes

pairwise associations with event j as the cue, whereas the $\mathcal{C}$-based prediction invokes credit associations with current and past events as cues. Finally, we compare these to a baseline of always predicting the most frequent event type. Our method is described in detail in Appendix A.6.

The average accuracies of the prediction methods are shown in Fig. 4.9c and 4.9d, as a function of the number of MRPs superposed, for the first and second approach of scenario generation respectively. The $\mathcal{C}$- and $\mathcal{M}$-based predictions generally outperform the baseline model. Across both figures, the results are qualitatively similar. The accuracies of $\mathcal{C}$- and $\mathcal{M}$-based predictions are comparable for a single MRP. This is expected since for an MRP, the cue and its outcome are neighbors. Whenever an event occurs, the $\mathcal{M}$-based predictor uses the pairwise associations between that event and its possible outcomes to predict the type of the next event. However, as more and more MRPs are superposed, the $\mathcal{C}$-based algorithm outperforms.

What drives the difference in performance between the $\mathcal{C}$- and $\mathcal{M}$-based algorithms? Although the $\mathcal{C}$-based algorithm uses the credit associations $\mathbf{C}$ while the $\mathcal{M}$-based algorithm uses the pairwise associations $\mathbf{M}$, this difference is immaterial in this case. Since each event only has one cue, $\exp \mathbf{C}^{\alpha}$ is proportional to $\mathbf{M}^{\alpha}$. Rather, the $\mathcal{M}$-based algorithm suffers when successive events originate from separate MRPs, and the pairwise association between the respective event types would not be predictive. The $\mathcal{M}$-based algorithm makes predictions only based on events in the present. In contrast, the $\mathcal{C}$-based algorithm makes predictions based on events in the present and in the past, where the correct cue would be included in such situations.

This demonstration provides a proof of concept that the algorithm provides reasonable predictions for cues at time scales spanning one order of magnitude. We accomplished this without selecting any single operating scale. The demonstration gives a flavor for the advantages of the algorithm we describe over Markov models. A classic approach based on n-th order Markov models would entail discretizing time at some lowest-level scale (but see Kurth-Nelson and Redish, 2009; Ludvig et al., 2008), and sizing the memory buffer to encompass most of the longest transitions. For simplicity, we have constructed a relatively tame scenario for this demonstration, in which most event relationships only span about 1–15 time units, and events are sparse. In reality, the wider the range of time scales, the harder it is for standard algorithms operating at the lowest-level time scale, which fumble at time scales significantly different from their operating scale (Mozer, 1992). In scenarios where events have long-range temporal dependencies, Markov models would be significantly limited by the exponential growth in the number of states (and thus, computational demands) with the size of the memory buffer. The algorithm we describe does not face these limitations (see Sec. 4.3.1).

4.5 Discussion

We have proposed an algorithm that generates a scale-invariant timeline of the future. This algorithm is time-local in the sense that predictions at time t are derived from $\tilde{\mathbf{f}}(\overset{*}{\tau}, t)$, which represents events that are, in fact, non-local in real time. Moreover, the translation mechanism enables event rates at future time points to be estimated. In addition to associative memory, as developed by model-free RL algorithms, this capability would let an agent

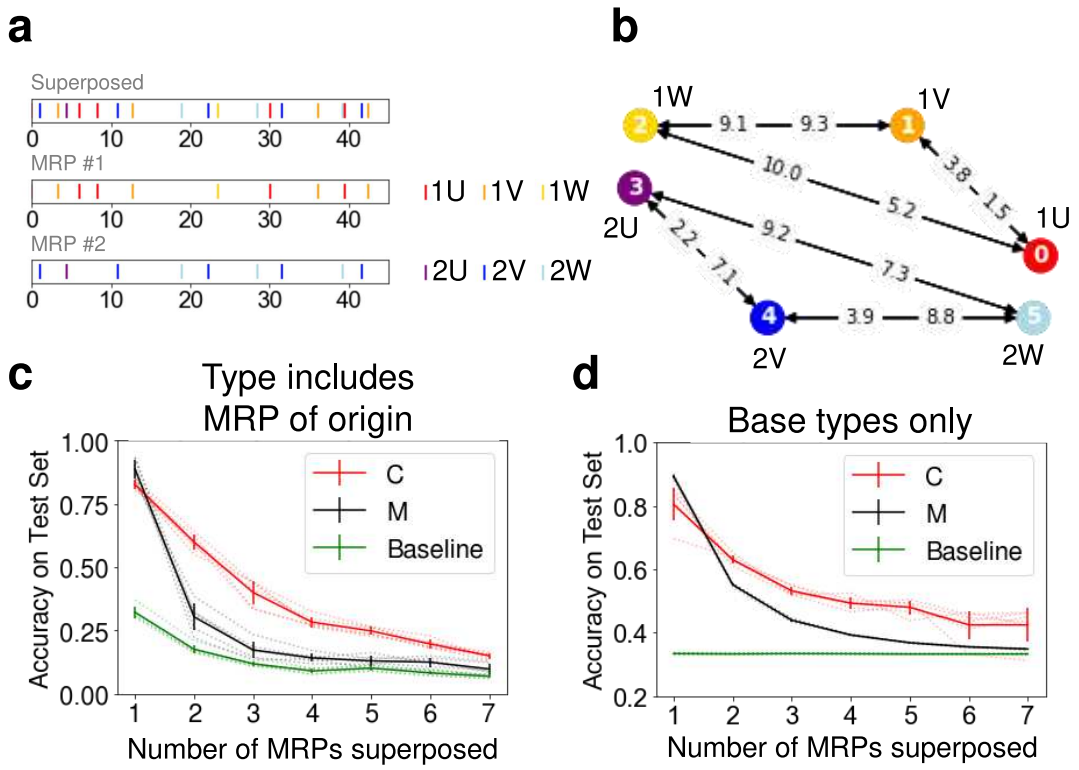

Figure 4.9: **The algorithm provides good predictions for cues at multiple time scales.** **(a)** The top panel shows the first few events in a superposed process. The bottom two panels show the corresponding events from the two component MRPs, which are composed of events $\{1\text{U}, 1\text{V}, 1\text{W}\}$ and $\{2\text{U}, 2\text{V}, 2\text{W}\}$, respectively. Note that successive events in the superposed process (e.g., the first two events in the topmost panel) may be from different MRPs, and thus the earlier event not predictive of the later event. **(b)** Graph depicting mean transition times between event types within each component MRP. Weights are associated with the arrowhead closest to them. The variances of the normally distributed transition times are not shown here. Note, for example, how the $1\text{V} \rightarrow 1\text{U}$ transition takes place at the scale of about 1.5 time units, while the $1\text{U} \rightarrow 1\text{W}$ transition takes place at a different scale of about 10 time units. The two MRPs depicted here are two of the component MRPs in the simulation used to generate (c). **(c)** We superpose MRPs such that event types from different MRPs are deemed different event types in the superposed process. **(d)** We superpose MRPs such that event types from different MRPs are identified by base types (U, V, W) irrespective of their MRP of origin, resulting in exactly three event types in the superposed process. For (c) and (d), each point represents an average accuracy computed by repeating the training and testing procedures 6 times for each choice of number of MRPs superposed. Regardless of method of superposing MRPs, the algorithm (labeled $\mathcal{C}$) performs well above chance, showing that it provides good predictions for cues at multiple time scales. See text for a comparison of the $\mathcal{C}$- and $\mathcal{M}$-based predictions.

construct the estimate over possible futures (McGuire and Kable, 2013).

4.5.1 Theoretical Properties of the Current Model

This model has properties that are quite different from traditional RL paradigms. First, this algorithm naturally runs in continuous time, which suits applications dealing with natural processes unfolding in time. This feature contrasts with basic RL algorithms, which only allow agents to move among discrete states in discrete time. In principle, this proposed algorithm can be extended such that position in higher-dimensional spaces replaces or augments time, allowing agents to navigate real and abstract spaces. Translation can be along an angle or perhaps even along a trajectory, instead of being confined to a given axis (Eq. 4.4).

Second, the scale invariance of the model is useful in applications where the time scale of event relationships is not known in advance. In principle, the model is indifferent towards the absolute time intervals between events. Instead, within a given scenario, it is only concerned about time intervals relative to other time intervals. In comparison, in traditional RL systems, a time scale for history dependency, if any, is set by the size of the history that the designer defines as part of the state s. Moreover, in many aspects of the world that we might be interested in, such as in natural language (Altmann et al., 2012), network traffic (Cohen et al., 2000) and financial markets (Cont, 2005), event dependencies exist simultaneously across a wide range of scales. This model is potentially suited for such applications, since it incorporates past events across a range of time scales, and an increase in computing resources provides an exponential increase in the length of history considered.

Third, in the context of RL, this model may be incorporated into algorithms to allow agents to naturally form a prediction of its own trajectory as a function of future time. This can be done by considering the agent's arrival at some or all states as events. In addition, by combining the predictions for future states s as a function of future time, $p^s(\delta)$, with a reward function over future states, the agent can generate the predicted future reward as a function of future time, $r(\delta)$. By learning and comparing weighted integrals of $r(\delta)$ for several alternative policies, the agent can choose flexibly among these policies according to task demands. For instance, if the agent knows it only has 10 time units to complete the task, it can choose the policy with the highest $\int_0^{10} r(\delta)\, d\delta$. The model's ability to form a prediction as a function of future time stands in contrast to RL paradigms, which tend to flatten the dimension of future time. For example, a naive RL agent assigns values to states according to the expected *sum* of future reward starting from that state; a successor representation agent (Dayan, 1993) learns the expected future state occupancy, *summed over future time*, starting from each state (but see Tano et al., 2020; Momennejad and Howard, 2018).

Finally, we note that this model provides information usually associated with model-based RL, but with very different computational properties. Like model-based RL, this model provides an explicit prediction as a function of future time δ. However, a constraint of model-based RL is that the time to compute an event δ in the future goes up linearly with δ. In the present model, because the calculation of the prediction at a particular value of δ does not depend on the prediction at previous values of δ, one could in principle compute all values of δ in parallel. Moreover, this means that it is possible to sample the δ axis in

whatever way is convenient. Integrals over δ give hyperbolic discounting if the δ axis is sampled evenly as a function of $\log \delta$. See also Shankar et al. (2016) for considerations related to physically instantiating translation across a population of neurons.

4.5.2 Theoretical Limitations of the Current Model

We highlight two limitations relating to applying this algorithm toward machine learning. First, the algorithm, as currently described, is not directly sensitive to joint statistics of two or more cues. For example, the model would be unable to capture the conditional structure "Z occurs exactly if either X or Y occurs, but not both". As a consequence, the algorithm is also unable to deal appropriately with number of events. For instance, the algorithm has no basis to differentiate "X precedes Y by 10 s" from "half of the time, X precedes two closely-spaced occurrences of Y by 10 s and the other half of the time, Y does not occur". We can mitigate this issue by perceiving events depending on context. For example, the agent can perceive the Y after an X as the event XY, enabling sensitivity to joint statistics of at most two cues. In terms of computational complexity, naively implementing this would introduce a quadratic factor in the number of base events. However, we can reduce the resource complexity by finding a compressed representation of the event history while preserving information about future events: that is, dealing with the information bottleneck problem (Tishby et al., 2000). Since existing deep neural network algorithms efficiently extract joint statistics, it would be natural to pursue research that seeks to merge this approach with deep network algorithms.

Second, this algorithm is limited in prescribing how to achieve optimal policies in the context of RL. Our focus has been on how to predict future events, and not how to learn the

best policy. In many contexts, it is natural to define events such that events occur depending on actions of the agent (e.g., in a spatial navigation task where events occur based on the agent's trajectory). In these cases, in effect, we presume that the agent follows an existing policy π, and the model deduces event associations and makes predictions with respect to π. The agent can certainly flexibly choose among several alternative policies, say, between π and π', by comparing predictions from the start state and selecting the more rewarding alternative. However, unlike basic RL algorithms, we do not prescribe a method for learning a policy that scales in complexity with the number of states, such as a policy to navigate a grid. In the context of grid navigation, we have, in effect, avoided assigning values to coordinates on the grid, since this contradicts our design principle of allowing history to influence events (rewards). More research would be needed if one wished to pursue policy learning within the framework we describe.

4.5.3 Neuroscience Considerations

This subsection discusses two potential points of contact between the formal model presented in this paper and computational and systems neuroscience.

4.5.3.1 Reward Prediction Error and Dopamine

The success of RL algorithms in accounting for the firing of dopaminergic neurons in the basal forebrain (Schultz et al., 1997) is arguably the greatest achievement in computational cognitive neuroscience. The basic empirical story is well-known. Dopaminergic neurons respond to unpredicted rewarding outcomes. However, with learning, as the reward becomes

predicted by a neutral stimulus, the cells no longer fire to the predicted reward but instead fire to the neutral stimulus that predicts the future rewarding outcome (see Schultz, 2006, for a review of the early literature). While there are undoubtedly many details that would need to be worked out, at least the rough outline of the classical empirical story about dopamine can be mapped onto this framework.

Let us suppose that expected future value is computed at each moment by integrating over future time δ, taking the projection from the vector of predicted events $\mathbf{p}(\delta)$ onto some vector $\mathbf{A}$ that describes the reward value of each possible event type in $\mathbf{p}$:

$$V(t) = \int_0^\infty \mathbf{A} \cdot \mathbf{p}(\delta; t) g(\delta)\, d\delta, \tag{4.21}$$

where $g(\delta)$ denotes the factor arising from compression of the δ axis. As discussed above, it is reasonable to sample δ on a logarithmic scale to implement hyperbolic discounting, in which case $g(\delta) = 1/\delta$. Let us suppose further that reward prediction error E is computed as the difference between $r(t)$, the actual reward observed at time step t, and the change in V at time step t:

$$E(t) = r(t) + [V(t) - V(t - \Delta t)], \tag{4.22}$$

where we have chosen a discrete time interval Δt to acknowledge that the computation of value may take a substantial amount of time in the brain. For instance, Shankar et al. (2016) proposed that integrals over δ could be completed within a theta oscillation, suggesting Δt could be as long as a few hundred milliseconds. Now, consider slowly learning an association between an inherently neutral event X and a rewarding event Y, separated by

a fixed delay τ. Initially, Y is unpredicted. When X is presented, there is no change in V. Similarly, V is zero both before and after Y is presented. Because Y is rewarding, the reward prediction error is positive around the time of presentation of Y. After learning, immediately after X is presented, $\mathbf{p}(\delta)$ includes the prediction for Y, a time $\delta \simeq \tau$ in the future. This means that $V(t)$ changes abruptly around the time that X is presented, resulting in a positive reward prediction error. Now, after learning, consider a time τ after presentation of X. If the rewarding stimulus is omitted, negative reward prediction error is observed as the peak in $\mathbf{p}(\delta)$ corresponding to Y becomes increasingly truncated. However, if Y is presented at the time it is expected, then the positive reward from Y is balanced by a rapid decreasing $V(t)$. Note that because Y does not predict itself at a short lag, observation of Y abruptly decreases the prediction of itself.

This approach aligns well with the classic understanding of reward prediction error with one very important exception. Rather than estimating expected future reward *via* temporal difference learning, predictions for an extended future are available at each moment. Unlike temporal difference learning algorithms, there is no sense in which value moves gradually along intermediate time points between X and Y. This model thus has no difficulty accounting for the finding that value seems to rapidly "jump" between events (Pan et al., 2005).

4.5.3.2 Translation and Theta Oscillations

The algorithm described here relies on the ability to translate $\tilde{\mathbf{f}}$ towards the past. Shankar et al. (2016) suggested that hippocampal theta (4–12 Hz) oscillations could provide a mechanism for translation of temporal representations. The basic conjecture of that model for translation

is that different values of δ map onto different phases of theta oscillations. If the timeline δ maps onto different phases of the theta oscillation, this places a lower limit on the order of 100 ms on the timelines indexed by $\overset{*}{\tau}$ and δ. Theoretical and neurobiological considerations led Shankar et al. (2016) to the conclusion that δ ought to accelerate exponentially with the theta cycle, resulting in a logarithmic sampling of the δ axis.

This conjecture made sense of several neurophysiological findings, including the gradual ramping of firing in striatal neurons accompanied by phase precession with respect to theta recorded in the hippocampus, a brain structure that is relatively distant from the striatum (van der Meer and Redish, 2011). The fact that spikes in the striatum are organized by hippocampal theta suggests that theta oscillations reflect a computation that is extended over a significant part of the brain. The learning rule presented here, Eq. 4.19, describes changes in the strength of connections in $\mathbf{C}$ by noting the difference between $\mathbf{p}^+$ and $\mathbf{p}^-$ at each value of δ. This suggests convergent connections between axons communicating $\mathbf{M}$ and $\mathbf{C}$ arriving at target neurons representing predicted future outcomes. Perhaps the coordination implied by the involvement of theta oscillations in prediction could lead to a difference in the timing of spikes communicated *via* $\mathbf{M}$ and $\mathbf{C}$. Coupled with spike-timing-dependent plasticity, perhaps this could lead to the learning rule in Eq. 4.19.

Acknowledgments

This work was supported by NIBIB R01EB022864 and NSF IIS 1631460. The authors gratefully acknowledge inspiring conversations with Randy Gallistel and work in early stages of this project by Kostya Tiurev.

Code Availability

The code that supports the demonstration in Sec. 4.4 can be found at `https://predicting.gitlab.io`.

5 GENERAL DISCUSSION

5.1 Summary

In Section 3, we looked briefly at some results that suggested that DG and LEC item representation is time-conjunctive and persistent. Mice were exposed to a series of odors in a task-free condition as calcium images were taken of DG and LEC. Heat maps were made of accuracies of cross-temporal odor classifiers. These heat maps peak along the diagonal and decrease in value away from the diagonal. This suggests population encoding of odor that evolves smoothly with time within each trial. Pooled analysis suggests that the identity of the previous odor can be decoded accurately, suggesting that DG and LEC item representation is persistent. In the context of this dissertation, these results motivate the possibility of using the TILT framework to build out an algorithm for anticipation.

In Section 4, we constructed a computational algorithm to generate a fuzzy timeline of future events (prediction), based on a fuzzy timeline of past events (memory) and pairwise associations between event types. The algorithm stores credit associations between event types, which are the factors by which one should increase our expectation for outcomes given cues. The algorithm constructs a timeline of the future by operating on the timeline of the past and the credit associations in a neurally plausible manner. When a cue occurs, the credit associations are updated towards the pairwise associations divided by the prediction prior to the cue. Dividing out the prior prediction curbs double-counting in the case of chains of events. Through experience, the algorithm refines credit associations and forms better predictions for the future.

5.2 Applications to Neuroscience

For agents operating using the TILT-based algorithm we have proposed, two properties hold. Firstly, predictions may depend on the indefinite past. Secondly, time-bound predictions for future events are available in advance, not only just before the events are about to occur. The framework of RL has become widely used in accounting for results in neuroscience experiments involving anticipation and reward prediction error. However, the standard RL framework has difficulty with these two properties. Regarding the first property, dependence on the indefinite past is associated with an exponential increase in computational complexity with the number of past time steps considered (K Namboodiri, 2021; Section 4.3.1). Suppose the environment is either 0 or 1 at each time step. To take into account the effect of the past n time steps, the RL state space would have to be 2^n in size, whose value functions require 2^n resources to maintain. The second property has been discussed in Section 1.4.

Neuroscience experiments have begun to unravel the time scales of the history dependence and the future horizon of animal prediction. On the issue of history dependence, Nakahara et al. (2004) has demonstrated that the activity of monkey midbrain dopamine neurons can depend on the reward history of the past 6 trials, lasting about 40 s in total. In the experiment, monkeys were presented with one of the four cardinal directions in randomized complete blocks, with exactly one direction cueing reward. The block design means that exactly one reward appears within each block. Therefore, the conditional expectation of reward increases with the length of the streak of unrewarded trials. In particular, no more than six consecutive trials would be unrewarded. The activity of dopamine neurons on cue presentation was consistent with the monkeys exploiting the information on the number of

unrewarded trials to inform their prior expectation of cue-associated reward. This suggests that monkey dopamine neurons are able to integrate information over about the past 40 s.

This finding does not necessarily imply that the monkey has used information as laid out along a timeline of the past. This is because the bulk of the history dependence is contained in the number of trials since the previous reward. A scalar counter can be used to store this information. Indeed, for the dopamine response, Nakahara et al. used a simple RL model whose states are imbued with information on the number of trials since reward. However, the RL model suggested that if all previous six trials were unrewarded, the dopamine response would be zero, but the experimental dopamine response was positive. Hence, the researchers introduced a counting error for the number of trials since reward, to simulate imperfect counting and improve the fit to the data.

The TILT-based model we have introduced can be alternatively used to model the dopamine response in the experiment. Each cardinal direction would inhibit itself partially, and be associated with the other directions, over the next three or so trials. Therefore, each unrewarded trial would increase the expectation for the rewarded cue. In this context, the TILT-based model has three conceptual advantages over the researchers' model. Firstly, fuzziness is in-built instead of being added as an afterthought. Secondly, the TILT-based model averts the question of how the states were obtained in the case of RL. Nakahara et al. overcame the difficulty of exponential resources for the states by using a summary statistic, which is the number of trials since reward. This would raise the question of how the monkey brain derived this summary statistic in the first place. Thirdly, although the history dependence is capped at the past 7 trials in this experiment, the TILT-based model unlocks

the possibility of dependence on the indefinite past. Experiments can be performed that allow for unbounded history dependence, such as where the rate of reward is a self-correcting process (see Example 2.4 in Rasmussen, 2018). These experiments would push the known boundaries of history dependence of single cell responses in animals, and discriminate between RL-based models and the model we propose.

5.3 Limitations and Further Study

Thus far, we have dispensed entirely with the problem of action selection. However, action selection is crucial for animals, and the key goal in RL. Therefore, the algorithm would be improved with a treatment of action selection. The problem of action selection is nontrivial. Actions and predictions influence each other, since each action would change an agent's future. The first step in an optimal series of actions may temporarily lower the agent's expected future reward. For example, suppose an agent wishes to maximize happiness by starting a business. Owning a successful business would make the agent happy, but losing capital would make the agent sad. Most new businesses fail. Suppose this is because most entrepreneurs do not follow through with the actions required to sustain a business (e.g., planning, networking, staying vigilant against competitors). Thus, the optimal series of actions may be to start a business *and then* following a complicated series of actions to sustain it, but the first action (i.e., starting the business) lowers reward in expectation. Thus, a greedy approach of taking the action that increases expected reward, or brings the expected future toward what the agent desires, may not always be desirable. On the other hand, a brute force search through the space of combinations of actions would be exponential in the

amount of action steps that the agent is able to make. Therefore, efficient algorithms are needed to make the problem of action selection tractable.

The beginnings of a strategy to speed up action selection may be related to the distinction between prospective and retrospective associations. In the algorithm we propose, the credit associations are prospective, in the sense that the associations are updated when cues arise, and are normalized relative to the cue occurrence. However, neural evidence suggests that animals also build retrospective cognitive maps, which are maintained from the perspective of rewarded outcomes, and normalized relative to the rewarded outcomes. Since rewarded outcomes tend to be rarer than cues, such retrospective cognitive maps may be more efficient. A combination of prospective and retrospective maps may speed learning and serial action selection (K Namboodiri and Stuber, 2021). Therefore, retrospective updating of credit associations, when reward is encountered, as well as building retrospective representations to aid action selection, may ease the problem of action selection in the algorithm we propose.

5.4 Conclusion

Collectively, the results presented in this dissertation suggest looking beyond the TD RL framework when thinking about what the brain of an animal might be doing in a conditioning experiment. I have presented an algorithm that does not require the use of states, unlike in RL. The algorithm is efficient and runs in real time. It operates natively in continuous time and over a wide range of time scales. In addition, it generates predictions natively as a function of future time, enabling flexible decision-making.

A THEORETICAL STUDY: APPENDIX

A.1 A Formal Model for Temporal Record of the Past

Let multiple types of discrete events occur in continuous time. For each event type, we denote the signal by $f(t)$, where each event is represented by a Dirac delta function at the instant it occurs. For each event type, an array of leaky integrators, F, with a range of decay rates s, receive the signal as input:

$$\frac{\partial}{\partial t}F(s;t) = -sF(s;t) + f(t). \tag{A.1}$$

The array of leaky integrators $F(s;t)$ encodes the real Laplace transform of the signal up to time t, where s is the Laplace domain variable. For each event type, an array of time cells $\tilde{f}(\overset{*}{\tau})$ approximately inverts the Laplace transform (see Post, 1930). This yields an estimate of the signal up to time t, at time offsets $\overset{*}{\tau}$ prior:

$$\tilde{f}(\overset{*}{\tau};t) = \tilde{f}(k/s;t) = \frac{(-1)^k}{k!}s^{k+1}\frac{\partial^k}{\partial s^k}F(s;t) = \mathbf{L}_k^{-1}F(s;t). \tag{A.2}$$

The constant k is a sharpness parameter. As $k \to \infty$, the estimate $\tilde{f}(\overset{*}{\tau})$ becomes precise, at the cost of infinite resources to implement the model. As stated in Eq. 4.2, for an arbitrary signal $\mathbf{f}$,

$$\tilde{\mathbf{f}}(\overset{*}{\tau};t) = \frac{1}{\overset{*}{\tau}}\int_{-\infty}^{t}\mathbf{f}(\tau)\Phi_k\left(\frac{t-\tau}{\overset{*}{\tau}}\right)d\tau. \tag{A.3}$$

In other words, for a given $\overset{*}{\tau}$, $\tilde{\mathbf{f}}(\overset{*}{\tau};t)$ is proportional to a causal convolution of the signal $\mathbf{f}$ with a kernel Φ_k that describes the smearing.

A.2 Time-Translation to Estimate the Future State of the Past

The future state of the memory (Eq. 4.4) can be readily computed through translation in the Laplace domain:

$$\tilde{\mathbf{f}}_\delta(\overset{*}{\tau};t) \equiv \mathbf{L}_k^{-1}\mathbf{R}^\delta\mathbf{F}(s;t) \equiv \mathbf{L}_k^{-1}\left\{e^{-s\delta}\mathbf{F}(s;t)\right\}. \tag{A.4}$$

Building a translation operator out of realistic neurons and synapses is a non-trivial, but tractable problem. It has been proposed that the brain implements translation to various amounts δ by mapping δ on to different phases of theta oscillations (Shankar et al., 2016). Previous work has long argued that theta oscillations, a prominent 4–12 Hz oscillation in the local field potential, have long been believed to be crucial in the neurobiology of memory (Buzsáki, 2002; Hasselmo et al., 2002; Kahana et al., 2001). Requiring scale invariance, and also consideration of the problem from the perspective of the individual neurons requires the sweep through δ to accelerate exponentially through the theta cycle.

A.3 Pairwise Association and Pairwise Prediction

The agent makes pairwise associations $\mathbf{M}$ between each pair of event types using Hebbian learning. As the agent experiences the world, the pairwise prediction $\mathbf{m}$ allows the agent to generate predictions for the future based on pairwise associations with the currently occurring events as cues. The pairwise prediction is a building block for the learning of the credit associations $\exp\mathbf{C}$, from which the prediction $\mathbf{p}$ is derived. This section consists

of two subsections. The first subsection motivates the form of the pairwise prediction **m** (Eqs. 4.10–4.12; in particular, the form of the integral in Eq. 4.11). The second subsection highlights and proves the numerical coincidence between **M** and **m** in a simple case, from which the normalization for **m** derives.

A.3.1 Equation for Pairwise Prediction

In this subsection, we motivate the equations for the pairwise prediction. We do this by showing that when memory is perfect, **m** reduces to the geometric mean of the elements of **M** associated with the events at time t, as desired.

Hebbian learning can be used to make pairwise associations **M** between events (Eq. 4.5). The parameters of **M** are the possible cue, the possible outcome and the internal time. The pairwise prediction **m** uses the pairwise associations **M** to make a prediction about future events based on possibly multiple currently occurring events. In other words, **m** serves to integrate cue–outcome pairwise information from multiple simultaneous cues in the present. **M**, in the definition for **m**, plays the role that $\exp \mathbf{C}$ does in the definition for **p**. In the algorithm, **m** serves the function of $\mathbf{p}^+$ (Eq. 4.18).

The pairwise prediction **m** is computed, when a set of events $\mathcal{E}_t$ occur at time t, as follows:

$$m^\beta(\delta;t) = \kappa_1 \exp\left(\frac{1}{|\mathcal{E}_t|}\sum_{\alpha\in\mathcal{E}_t}\int f_\delta^\alpha(\overset{*}{\tau};t)\log M_\alpha^\beta(\overset{*}{\tau})\,d\overset{*}{\tau}\right), \qquad \delta>0, \tag{A.5}$$

where the constant $\kappa_1 = \left[ke^{-\psi(k)}\right]^{k+1}$, $\psi(k)$ is the digamma function, and

$$f_\delta^\alpha(\overset{*}{\tau};t) = \Phi_k\left(\delta/\overset{*}{\tau}\right)/\overset{*}{\tau}$$

denotes the future state of the memory element associated with the currently occurring episode of α. (For $k=2$, $\kappa_1 \sim 2.3$; for $k=8$, $\kappa_1 \sim 1.8$; as $k \to \infty$, $\kappa_1 \to \sqrt{e} \sim 1.65$.) The notation $|\mathcal{E}_t|$ denotes the number of elements in $\mathcal{E}_t$, that is, the number of events co-occurring at time t. The function $f_\delta^\alpha(\overset{*}{\tau};t)$ is a Gaussian-like function that peaks around $\overset{*}{\tau} = \delta$, and reflects the fact that at a time δ in the future, α would have occurred $\overset{*}{\tau}$ in the past, and the memory $\tilde{f}^\alpha$ would reflect this.

We can motivate the form of the pairwise prediction $\mathbf{m}$ as follows. Let us imagine that memory were perfect such that events were localized in time exactly, so $\phi_\delta^\alpha(\overset{*}{\tau}) = \delta(\overset{*}{\tau} - \delta)I(\alpha \in \mathcal{E}_t)$, where ϕ_δ^α is the analog of f_δ^α when memory is perfect (i.e., $k \to \infty$), $\delta(\cdot)$ is the Dirac delta function and $I(\cdot)$ is the indicator function. We would then have

$$\begin{aligned}\int \phi_\delta^\alpha(\overset{*}{\tau};t)\log M_\alpha^\beta(\overset{*}{\tau})\,d\overset{*}{\tau} &= I(\alpha \in \mathcal{E}_t)\int \delta(\overset{*}{\tau} - \delta)\log M_\alpha^\beta(\overset{*}{\tau})\,d\overset{*}{\tau} \\ &= I(\alpha \in \mathcal{E}_t)\log M_\alpha^\beta(\delta).\end{aligned}$$

The integral on the right hand side is a convolution of $\log M_\alpha^\beta$ with the delta function, which

returns the former unchanged. In the case where exactly X occurs at time t,

$$\begin{aligned} m^{\beta}(\delta;t) &\propto \exp \int f_{\delta}^{\alpha}(\overset{*}{\tau};t) \log M_{\alpha}^{\beta}(\overset{*}{\tau})\, d\overset{*}{\tau} \\ &= \exp \log M_{\mathrm{X}}^{\beta}(\delta) \\ &= M_{\mathrm{X}}^{\beta}(\delta), \end{aligned} \tag{A.6}$$

so $\mathbf{m}$ would be proportional to the appropriate elements of $\mathbf{M}$. In the case where exactly X and Y occur at time t,

$$\begin{aligned} m^{\beta}(\delta;t) &\propto \exp \left[\frac{1}{|\mathcal{E}_t|} \sum_{\alpha \in \mathcal{E}_t} \int f_{\delta}^{\alpha}(\overset{*}{\tau};t) \log M_{\alpha}^{\beta}(\overset{*}{\tau})\, d\overset{*}{\tau} \right] \\ &= \exp \left\{ \frac{1}{2} \left[\log M_{\mathrm{X}}^{\beta}(\delta) + \log M_{\mathrm{Y}}^{\beta}(\delta) \right] \right\} \\ &= \sqrt{M_{\mathrm{X}}^{\beta}(\delta) M_{\mathrm{Y}}^{\beta}(\delta)}, \end{aligned}$$

so $\mathbf{m}$ would be proportional to the geometric mean of the elements of $\mathbf{M}$ associated with the events at time t.

In the model with fuzzy memory, f_{δ} approximates ϕ_{δ}, so the above relationships hold only approximately. In other words, in general, according to Eqs. 4.10–4.12, $m^{\beta}(\delta;t)$ is not always exactly $M_{\mathrm{X}}^{\beta}(\delta)$ (when X occurs at time t) or $\sqrt{M_{\mathrm{X}}^{\beta}(\delta) M_{\mathrm{Y}}^{\beta}(\delta)}$ (when X and Y occur at time t), and so on. Instead of using Eqs. 4.10–4.12, which involves an integral, one could have defined $m^{\beta}(\delta;t)$ directly as the geometric mean of the relevant elements of $M^{\beta}(\delta)$. However, we do not do so. We use Eqs. 4.10–4.12 because they closely parallel Eqs. 4.14–4.15 for the prediction, for which the integral is necessary (see Sec. A.4). Using

equations of similar form is more neurobiologically realistic, because it suggests that analogous neural architecture supports the computation for both pairwise prediction **m** and prediction **p**. Integrals in time are straightforward to implement with neural networks.

In summary, we have shown that when memory is perfect, **m** reduces to the geometric mean of the elements of **M** associated with the events at time t, as desired. The strength of the equations underlying **m** is that they closely parallel those for **p**.

A.3.2 Normalization for Pairwise Prediction

In the bulk of the previous subsection, we imagined that memory were perfect such that events were localized in time exactly. In the actual formulation, memory is fuzzy, and this is reflected in the form of $\mathbf{f}_\delta$. Therefore, the above proportionality relationships do not hold exactly, in general. However, Eq. A.6 holds at the time of occurrence of X in the case where X and Y occur at a constant time interval τ, and no event co-occurs with X, even in the case of fuzzy memory. In this subsection, we will prove this result (in Lemma 3), which suggests the value that the normalization constant κ_1 should take. Lemmas 1 and 2 are integrals that are used to prove Lemma 3.

Lemma 1. If the constant a is positive and k is a non-negative integer,

$$\int_0^\infty \frac{e^{-a/x}}{x^{k+2}}\,dx = \frac{k!}{a^{k+1}}.$$

Proof. The integral is

$$I = \int_0^\infty \frac{e^{-a/x}}{x^{k+2}}\,dx = \frac{1}{a}\int_0^\infty \frac{ae^{-a/x}}{x^2}x^{-k}\,dx.$$

Noting that $\frac{d}{dx}\left(e^{-a/x}\right) = \frac{a}{x^2}e^{-a/x}$, we integrate by parts:

$$u = x^{-k} \qquad dv = \frac{ae^{-a/x}}{x^2}$$

$$du = (-k)\,x^{-k-1} \qquad v = e^{-a/x},$$

so our integral is now

$$I = \cancel{\left[\frac{1}{a}x^{-k}e^{-a/x}\right]_0^{\infty}} - \frac{-k}{a}\int_0^{\infty}\frac{e^{-a/x}}{x^{k+1}}\,dx = \frac{k}{a}\frac{1}{a}\int_0^{\infty}\frac{ae^{-a/x}}{x^2}x^{-(k-1)}\,dx.$$

If we were to integrate by parts again, we would have:

$$u = x^{-(k-1)} \qquad dv = \frac{ae^{-a/x}}{x^2}$$

$$du = -(k-1)\,x^{-k-1} \qquad v = e^{-a/x}.$$

Each ith iteration reduces the exponent on x in the denominator of the integrand by 1 and introduces a factor of $(k-i+1)/a$, and k iterations are needed to go from having x^{k+2} to x^2 in the denominator of the integrand. Noting that $(k)(k-1)\cdots(2)(1) = k!$, we thus have

$$I = \frac{k!}{a^k}\frac{1}{a}\int_0^{\infty}\frac{ae^{-a/x}}{x^2}\,dx = \frac{k!}{a^k}\frac{1}{a}\left[e^{-a/x}\right]_0^{\infty} = \frac{k!}{a^{k+1}}. \qquad \square$$

Lemma 2. If the constants A, a and b are positive and k and m are positive integers, then

$$\int_0^{\infty}\frac{e^{-a/x}}{x^{k+1}}\log\frac{Ae^{-b/x}}{x^m}\,dx = \frac{(k-1)!}{a^k}\left[m\psi(k) - \frac{kb}{a} + \log\frac{A}{a^m}\right],$$

where $\psi(k)$ is the digamma function.

Proof. The integrand is

$$\frac{e^{-a/x}}{x^{k+1}} \log \frac{Ae^{-b/x}}{x^m} = \frac{e^{-a/x}}{x^{k+1}} \log A - \frac{be^{-a/x}}{x^{k+2}} - m\frac{e^{-a/x}}{x^{k+1}} \log x.$$

We integrate term by term. Applying Lemma 1, the first term is

$$\log A \int_0^\infty \frac{e^{-a/x}}{x^{k+1}}\, dx = (\log A)\, \frac{(k-1)!}{a^k},$$

noting that k above is at least 1, and the second term is

$$-b \int_0^\infty \frac{e^{-a/x}}{x^{k+2}}\, dx = -b\frac{k!}{a^{k+1}}.$$

The third term is

$$-m \int_0^\infty \frac{e^{-a/x}}{x^{k+1}} \log x\, dx, \qquad k \geq 1.$$

Substituting $t = a/x$, $dt = -a/x^2\, dx$, we have

$$\begin{aligned}
-\frac{m}{a^{k-1}} \int_\infty^0 z^{k-1} e^{-t} \log \frac{a}{t} \left(-\frac{1}{a}\right) dt &= \frac{m}{a^k} \int_0^\infty t^{k-1} e^{-t} \log \frac{t}{a}\, dt \\
&= \frac{m}{a^k} \left[(-\log a) \int_0^\infty t^{k-1} e^{-t} dt + \int_0^\infty t^{k-1} e^{-t} \log t\, dt \right] \\
&= \frac{m}{a^k} \left[(-\log a)\, \Gamma(k) + \Gamma'(k) \right] \\
&= \frac{m}{a^k} \Gamma(k) \left[\psi(k) - \log a \right] \\
&= \frac{m}{a^k} (k-1)! \left[\psi(k) - \log a \right],
\end{aligned}$$

where we have applied Eqs. 5.2.1, 5.9.19 and 5.2.2 from DLMF (2021), and used the fact that k is a positive integer. Putting everything together, we have

$$\begin{aligned}\int_0^\infty \frac{e^{-a/x}}{x^{k+1}} \log \frac{Ae^{-b/x}}{x^m}\, dx &= (\log A)\,\frac{(k-1)!}{a^k} - b\frac{k!}{a^{k+1}} + \frac{m}{a^k}\,(k-1)!\,[\psi(k) - \log a] \\ &= \frac{(k-1)!}{a^k}\left\{\log A - \frac{kb}{a} + m\,[\psi(k) - \log a]\right\} \\ &= \frac{(k-1)!}{a^k}\left[m\psi(k) - \frac{kb}{a} + \log\frac{A}{a^m}\right]. \qquad \square\end{aligned}$$

Lemma 3. Let Event i cue Event j with a fixed time interval t_{ij}, and let no event co-occur with Event i. After training, at the time Event i occurs,

$$m^j(\delta) = M_i^j(\delta),$$

for all $\delta > 0$.

Proof. On the right hand side, we have

$$M_i^j(\delta) = K\frac{t_{ij}^k}{\delta^{k+1}} e^{-kt_{ij}/\delta}, \qquad \delta > 0,$$

where $K = \frac{k^{k+1}}{k!}$.

To compute the left hand side, we note that

$$f_\delta^i\left(\overset{*}{\tau}\right) = K\frac{\delta^k}{\overset{*}{\tau}^{k+1}} e^{-k\delta/\overset{*}{\tau}}, \qquad \overset{*}{\tau}, \delta > 0.$$

Then

$$\int f_\delta^i\left(\overset{*}{\tau}\right)\log M_i^j\left(\overset{*}{\tau}\right)d\overset{*}{\tau}=\int_0^\infty K\frac{t_{ij}^k}{\overset{*}{\tau}^{k+1}}e^{-kt_{ij}/\overset{*}{\tau}}\log\left(K\frac{t_{ij}^k}{\overset{*}{\tau}^{k+1}}e^{-kt_{ij}/\overset{*}{\tau}}\right)d\overset{*}{\tau} \tag{A.7}$$

$$=K\delta^k\int_0^\infty\frac{e^{-k\delta/\overset{*}{\tau}}}{\overset{*}{\tau}^{k+1}}\log\left(\frac{Kt_{ij}^k}{\overset{*}{\tau}^{k+1}}e^{-kt_{ij}/\overset{*}{\tau}}\right)d\overset{*}{\tau}. \tag{A.8}$$

Substituting $a=k\delta$, $b=kt_{ij}$, $A=Kt_{ij}^k$, $m=k+1$ into Lemma 2, the above evaluates to

$$\begin{aligned}\int f_\delta^i\left(\overset{*}{\tau}\right)\log M_i^j\left(\overset{*}{\tau}\right)d\overset{*}{\tau}&=\cancel{K\delta^k\frac{(k-1)!}{(k\delta)^k}}\left[(k+1)\,\psi(k)-\frac{k^2t_{ij}}{k\delta}+\log\frac{Kt_{ij}^k}{(k\delta)^{k+1}}\right]\\&=(k+1)\,\psi(k)-\frac{kt_{ij}}{\delta}+\log\frac{Kt_{ij}^k}{\delta^{k+1}}-(k+1)\log k\\&=\log\frac{Kt_{ij}^k}{\delta^{k+1}}-\frac{kt_{ij}}{\delta}-\log\kappa_1\end{aligned}$$

Thus, the left hand side of the lemma is

$$\begin{aligned}m^j(\delta)&=\kappa_1\exp\left(\frac{1}{|\mathcal{E}_t|}\sum_{\alpha\in\mathcal{E}_t}\int f_\delta^\alpha(\overset{*}{\tau};t)\log M_\alpha^j(\overset{*}{\tau})\,d\overset{*}{\tau}\right)\\&=\kappa_1\exp\left(\int f_\delta^i(\overset{*}{\tau};t)\log M_i^j(\overset{*}{\tau})\,d\overset{*}{\tau}\right)\\&=\kappa_1\exp\left(\log\frac{Kt_{ij}^k}{\delta^{k+1}}-\frac{kt_{ij}}{\delta}-\log\kappa_1\right)\\&=K\frac{t_{ij}^k}{\delta^{k+1}}e^{-kt_{ij}/\delta}\\&=M_i^j(\delta).\end{aligned}$$

□

Our choice of $\kappa_1=\left[ke^{-\psi(k)}\right]^{k+1}$ allowed the equality $m^j(\delta)=M_i^j(\delta)$ to hold when Event i cues Event j with a fixed time interval and no event co-occurs with Event i, without

any additional constant of proportionality in the equation.

Remark. The conclusion of Lemma 3 does not hold in general if the delay interval is not fixed. Let i and then j occur at times 1 or 2 apart (with equal probability). Let $k = 2$ (for ease of calculation) so $K = \frac{k^{k+1}}{k!} = 4$. We have

$$\begin{aligned} M_i^j\left(\overset{*}{\tau}\right) &= \frac{\int \tilde{f}_{\delta=0}^i\left(\overset{*}{\tau};t\right) f_j(t)\,dt}{\int f_i(t)\,dt} \\ &= \frac{1}{2}\left(K\frac{1}{\overset{*}{\tau}^{k+1}}\left(e^{-k/\overset{*}{\tau}} + 2^k e^{-2k/\overset{*}{\tau}}\right)\right) \\ &= \frac{2}{\overset{*}{\tau}^3}\left(e^{-2/\overset{*}{\tau}} + 4e^{-4/\overset{*}{\tau}}\right) \\ M_i^j\left(\overset{*}{\tau}=1\right) &= 2\left(e^{-2}+4e^{-4}\right) = \boxed{0.417195}, \end{aligned}$$

and as before

$$f_\delta^i\left(\overset{*}{\tau}\right) = 4\frac{\delta^2}{\overset{*}{\tau}^3}e^{-2\delta/\overset{*}{\tau}}.$$

But when $\delta = 1$,

$$\begin{aligned} \int f_\delta^i\left(\overset{*}{\tau}\right)\log M_i^j\left(\overset{*}{\tau}\right)d\overset{*}{\tau} &= 4\delta^2\int_0^\infty \frac{e^{-2\delta/\overset{*}{\tau}}}{\overset{*}{\tau}^3}\log\left[\frac{2}{\overset{*}{\tau}^3}\left(e^{-2/\overset{*}{\tau}}+4e^{-4/\overset{*}{\tau}}\right)\right]d\overset{*}{\tau} \\ &= -1.51366 \\ m^j(\delta=1) &= \kappa_1\exp(-1.51366) \\ &= 8\exp[-3(1-\gamma)]\exp(-1.51366) \\ &= \boxed{0.495310}. \end{aligned}$$

Since the two boxed values are different, the equality $m^j(\delta) = M_i^j(\delta)$ does not hold in

general.

A.4 Credit Association and Prediction

The agent maintains credit associations $\exp \mathbf{C}$ between each pair of event types, which estimates the multiplier for the agent's belief about the rate of each outcome whenever it sees a potential cue. The agent maintains predictions $\mathbf{p}$, a timeline of future events based on the credit associations. As the agent experiences the world, $\exp \mathbf{C}$ and $\mathbf{p}$ are iteratively updated.

This section motivates the form of the prediction $\mathbf{p}$ (Eqs. 4.14–4.15), by showing that the integral approximately subtracts the time elapsed since each cue, from the time delay between that cue and the outcome of interest. The integral thus produces a function of δ that peaks approximately at the time remaining to the outcome. This section consists of three subsections. In the first subsection, we calculate the projected memory $\tilde{\mathbf{f}}_\delta$, which is an element in the prediction $\mathbf{p}$, in the case of perfect memory. In the second subsection, we calculate the prediction $\mathbf{p}$ in the case of perfect memory. In the last subsection, we discuss the case of fuzzy memory.

A.4.1 Perfect Memory: Projected Memory

We proposed the following form for the prediction:

$$p^\beta(\delta;t) = \Lambda^\beta \exp \sum_{E \in \mathcal{E}} \int C_E^\beta(\overset{*}{\tau}) \tilde{f}_\delta^E(\overset{*}{\tau};t)\, d\overset{*}{\tau},$$

where Λ^{β} denotes the long-term average of event type β, and $\mathcal{E}$ denotes the set of possible cue types. We motivate the above functional form by deriving the prediction in the case of perfect memory (i.e., $k \to \infty$) and discrete events. To do so, we would need to find the projected memory $\tilde{f}_{\delta}^{E}$ for event type E. Note (see Eq. 4.1) that

$$\lim_{k\to\infty} \Phi_k(x) = \delta(x-1),$$

where $\delta(\cdot)$ represents the Dirac delta function. Let the input function $\mathbf{f}$ be a series of discrete events of type e_i occurring at times $t_i < t$,

$$f^{E}(\tau) = \sum_i \delta(t_i - \tau) I(e_i = E).$$

If we imagine that memory were perfect, then from Eq. 4.4, the projected memory would be represented as

$$\begin{aligned}
\tilde{\phi}_{\delta}^{E}(\overset{*}{\tau};t) &= \lim_{k\to\infty} \frac{1}{\overset{*}{\tau}} \int_{-\infty}^{t} f^{E}(\tau)\, \Phi_k\left(\frac{t+\delta-\tau}{\overset{*}{\tau}}\right) d\tau \\
&= \frac{1}{\overset{*}{\tau}} \int_{-\infty}^{t} \sum_i \delta(t_i - \tau) I(e_i = E)\, \delta\left(\frac{t+\delta-\tau}{\overset{*}{\tau}} - 1\right) d\tau \\
&= \sum_i I(e_i = E) \int_{-\infty}^{t} \delta(t_i - \tau) \frac{1}{\overset{*}{\tau}} \delta\left(\frac{t+\delta-\tau-\overset{*}{\tau}}{\overset{*}{\tau}}\right) d\tau \\
&= \sum_i I(e_i = E) \int_{-\infty}^{t} \delta(t_i - \tau)\, \delta\left(t+\delta-\tau-\overset{*}{\tau}\right) d\tau \\
&= \sum_i \delta(t - t_i + \delta - \overset{*}{\tau}) I(e_i = E),
\end{aligned}$$

where $\tilde{\phi}_\delta$ is the analog of $\tilde{\mathbf{f}}_\delta$ when memory is perfect, and we have used the property $\delta(\alpha x) = \delta(x)/|\alpha|$ of the Dirac delta function.

A.4.2 Perfect Memory: Prediction

We are now ready to derive the prediction:

$$\begin{aligned} p^\beta(\delta;t) &= \Lambda^\beta \exp \sum_{E\in\mathcal{E}} \int C_E^\beta(\overset{*}{\tau})\tilde{\phi}_\delta^E(\overset{*}{\tau};t)\, d\overset{*}{\tau} \\ &= \Lambda^\beta \exp \sum_i \sum_{E\in\mathcal{E}} \int C_E^\beta(\overset{*}{\tau})\delta(t-t_i+\delta-\overset{*}{\tau})I(e_i=E)\, d\overset{*}{\tau}. \end{aligned}$$

Consider the summand associated with $i=1$,

$$\sum_{E\in\mathcal{E}} \int C_E^\beta(\overset{*}{\tau})\delta(t-t_1+\delta-\overset{*}{\tau})I(e_1=E)\, d\overset{*}{\tau} = \int C_{e_1}^\beta(\overset{*}{\tau})\delta(t-t_1+\delta-\overset{*}{\tau})\, d\overset{*}{\tau}.$$

The above integral is a convolution of $C_{e_1}^\beta(\delta)$ with $\delta_{-(t-t_1)}(\delta) = \delta(t-t_1+\delta)$. The result is a translation of $C_{e_1}^\beta(\delta)$ by $-(t-t_1)$, the time interval since Event $i=1$:

$$\int C_{e_1}^\beta(\overset{*}{\tau})\delta(t-t_1+\delta-\overset{*}{\tau})\, d\overset{*}{\tau} = C_{e_1}^\beta(t-t_1+\delta).$$

This makes sense because $C_{e_1}^\beta$ peaks at the time that β is expected after having observed e_1 (say, τ). Since e_1 occurred $t-t_1$ ago, the view of $C_{e_1}^\beta$ that matters for the prediction should be translated by $-(t-t_1)$. This view peaks at $\tau-(t-t_1)$, which is the time remaining until β is expected on the basis of e_1. (For example, if β is expected 5 time units after e_1 ($C_{e_1}^\beta(\delta)$ peaks at $\delta=5$) and e_1 occurred 3 time units ago, then β is expected in 2 time units

($C^{\beta}_{e_1}(t-t_1+\delta) = C^{\beta}_{e_1}(3+\delta)$ peaks at $\delta = 2$).)

If there is only one event episode, indexed by $i = 1$,

$$p^{\beta}(\delta;t) = \Lambda^{\beta} \exp C^{\beta}_{e_1}(t-t_1+\delta).$$

Thus, the prediction p^{β} will also peak at $\delta = \tau - (t-t_1)$. (Continuing with the example, the prediction will also peak at $\delta = 2$.) The form of the prediction makes sense. The quantity Λ^{β} is the rate of occurrence of β without any evidence, or the agent's prior belief. The quantity $\exp C^{\beta}_{e_1}$ is the adjustment factor for the rate of β due to e_1. The product is the new, adjusted estimate for the rate of β.

If there are exactly two event types with one episode each,

$$p^{\beta}(\delta;t) = \Lambda^{\beta} \left[\exp C^{\beta}_{e_1}(t-t_1+\delta)\right] \left[\exp C^{\beta}_{e_2}(t-t_2+\delta)\right],$$

and so on. In other words, the adjustment factors for different event types multiply, as would be desired.

A.4.3 Fuzzy Memory

In the earlier subsections, we have assumed that memory is perfect to illustrate the principles behind the form of the prediction **p**. However, memory is fuzzy in real, resource-bounded systems, so the foregoing comments only apply approximately. The form of the integral in the prediction,

$$\int C^{\beta}_{E}(\overset{*}{\tau}) \tilde{f}^{E}_{\delta}(\overset{*}{\tau};t)\, d\overset{*}{\tau},$$

is not a convolution since the width of $\tilde{f}_{\delta}^{E}$ increases with δ. However, the purpose of the integral is to approximate a convolution, so that the relevant view of C_{E}^{β} can be used to generate a prediction. For example, if E occurred 3 time units ago, $\tilde{f}_{\delta}^{E}(\overset{*}{\tau})$ would peak at $\overset{*}{\tau} = \delta + 3$ units (as δ increases, $\tilde{f}_{\delta}^{E}(\overset{*}{\tau})$ becomes increasingly wider). If E cues β with a delay of 5 time units, $C_{E}^{\beta}(\overset{*}{\tau})$ would peak at around 5 time units. The integral attains its largest value approximately when the peaks of $\tilde{f}_{\delta}^{E}(\overset{*}{\tau})$ and $C_{E}^{\beta}(\overset{*}{\tau})$ are aligned, which is around $\delta = 2$. This makes sense, because β is expected in 2 time units from the present.

In summary, we have shown that the integral in Eq. 4.15 approximately subtracts the time elapsed since each cue, from the time delay between that cue and the outcome of interest. The integral thus produces a function of δ that peaks approximately at the time remaining to the outcome.

A.5 Worked Examples

In Sec. 4.3.2, we noted that fuzzy memory induces fuzziness in the prediction **p**, and noted that the latter could have equivalently been induced by intrinsic temporal uncertainty in the input in an agent with perfect memory, for every given snapshot in time. In Sec. 4.3.4, we noted that other things equal, cues closer in time to outcomes tend to receive more credit for those outcomes, in the case of fuzzy memory. We provided one example each for illustration. This section consists of two subsections. In the first subsection, we provide the mathematical details for the example in Sec. 4.3.2. In the second subsection, we do so for the example in Sec. 4.3.4.

A.5.1 Worked Example 1: Forward Conditioning

For this example, Event X cues Event Y with a fixed delay of time interval τ. We work out the agent's prediction for Y at the time X occurs.

The pairwise association is given in Eq. 4.6,

$$M_{\mathrm{X}}^{\mathrm{Y}}(\overset{*}{\tau}) = \Phi_k(\tau/\overset{*}{\tau})/\overset{*}{\tau}.$$

By Lemma 3, when X occurs,

$$m^{\mathrm{Y}}(\overset{*}{\tau}) = M_{\mathrm{X}}^{\mathrm{Y}}(\overset{*}{\tau}) = \Phi_k(\tau/\overset{*}{\tau})/\overset{*}{\tau}.$$

When X occurs, the memory is empty, so

$$(p^-)^{\mathrm{Y}} = \Lambda^{\mathrm{Y}}.$$

Thus, after training, the credit due to X for Y is

$$\exp C_{\mathrm{X}}^{\mathrm{Y}}(\overset{*}{\tau}) = \frac{p^+(\overset{*}{\tau})}{p^-(\overset{*}{\tau})} = \frac{\Phi_k(\tau/\overset{*}{\tau})}{\Lambda^{\mathrm{Y}}\overset{*}{\tau}}.$$

The projected memory for X when X occurs is

$$\tilde{f}_{\delta}^{\mathrm{X}}(\overset{*}{\tau}) = \Phi_k(\delta/\overset{*}{\tau})/\overset{*}{\tau}.$$

The prediction for Y at the time X occurs is

$$\begin{aligned} p^{\mathrm{Y}}(\delta) &= \Lambda^{\mathrm{Y}} \exp \sum_{E \in \mathcal{E}} \int C_E^{\mathrm{Y}}(\overset{*}{\tau}) \tilde{f}_\delta^E(\overset{*}{\tau})\, d\overset{*}{\tau} \\ &= \Lambda^{\mathrm{Y}} \exp \int C_{\mathrm{X}}^{\mathrm{Y}}(\overset{*}{\tau}) \tilde{f}_\delta^{\mathrm{X}}(\overset{*}{\tau})\, d\overset{*}{\tau}. \end{aligned} \tag{A.9}$$

The integral is

$$\begin{aligned} \int C_{\mathrm{X}}^{\mathrm{Y}}(\overset{*}{\tau}) \tilde{f}_\delta^{\mathrm{X}}(\overset{*}{\tau})\, d\overset{*}{\tau} &= \int \frac{\Phi_k(\delta/\overset{*}{\tau})}{\overset{*}{\tau}} \log \frac{\Phi_k(\tau/\overset{*}{\tau})}{\Lambda^{\mathrm{Y}} \overset{*}{\tau}}\, d\overset{*}{\tau} \\ &= \kappa_0 \delta^k \int \frac{e^{-k\delta/\overset{*}{\tau}}}{\overset{*}{\tau}^{k+1}} \log \frac{\kappa_0 \tau^k e^{-k\tau/\overset{*}{\tau}}}{\Lambda^{\mathrm{Y}} \overset{*}{\tau}^{k+1}}\, d\overset{*}{\tau}. \end{aligned} \tag{A.10}$$

By Lemma 2, we have

$$\int_0^\infty \frac{e^{-a/x}}{x^{k+1}} \log \frac{A e^{-b/x}}{x^m}\, dx = \frac{(k-1)!}{a^k} \left[m\psi(k) - \frac{kb}{a} + \log \frac{A}{a^m} \right],$$

so we substitute $a = k\delta$, $A = \kappa_0 \tau^k / \Lambda^{\mathrm{Y}}$, $b = k\tau$ and $m = k+1$ to find

$$\int_0^\infty \frac{e^{-k\delta/\overset{*}{\tau}}}{\overset{*}{\tau}^{k+1}} \log \frac{\kappa_0 \tau^k e^{-k\tau/\overset{*}{\tau}}}{\Lambda^{\mathrm{Y}} \overset{*}{\tau}^{k+1}}\, d\overset{*}{\tau} = \frac{(k-1)!}{(k\delta)^k} \left[(k+1)\,\psi(k) - \frac{k\tau}{\delta} + \log \frac{\kappa_0 \tau^k}{\Lambda^{\mathrm{Y}} (k\delta)^{k+1}} \right].$$

Therefore,

$$\begin{aligned} \int C_{\mathrm{X}}^{\mathrm{Y}}(\overset{*}{\tau}) \tilde{f}_\delta^{\mathrm{X}}(\overset{*}{\tau})\, d\overset{*}{\tau} &= \kappa_0 \delta^k \frac{(k-1)!}{(k\delta)^k} \left[(k+1)\,\psi(k) - \frac{k\tau}{\delta} + \log \frac{\tau^k}{\Lambda^{\mathrm{Y}} k! \delta^{k+1}} \right] \\ &= (k+1)\,\psi(k) - \frac{k\tau}{\delta} + \log \frac{\tau^k}{\Lambda^{\mathrm{Y}} k! \delta^{k+1}} \end{aligned}$$

and the prediction for Y at the time X occurs is

$$\begin{aligned}
p^{\mathrm{Y}}(\delta) &= \Lambda^{\mathrm{Y}} \exp \int C_{\mathrm{X}}^{\mathrm{Y}}(\overset{*}{\tau}) \tilde{f}_{\delta}^{\mathrm{X}}(\overset{*}{\tau})\, d\overset{*}{\tau} \\
&= \Lambda^{\mathrm{Y}} \exp \left[(k+1)\,\psi(k) - \frac{k\tau}{\delta} + \log \frac{\tau^k}{\Lambda^{\mathrm{Y}} k! \delta^{k+1}} \right] \\
&= \Lambda^{\mathrm{Y}} \exp \left[(k+1)\,\psi(k) \right] \frac{\tau^k}{\Lambda^{\mathrm{Y}} k! \delta^{k+1}} e^{-k\tau/\delta} \\
&= \frac{\exp \left[(k+1)\,\psi(k) \right]}{k!} \frac{\tau^k}{\delta^{k+1}} e^{-k\tau/\delta} \\
&= \frac{\exp \left[(k+1)\,\psi(k) \right]}{k^{k+1}} \frac{1}{\delta} \kappa_0 \left(\frac{\tau}{\delta} \right)^k e^{-k\tau/\delta} \\
&= \frac{\kappa_1^{-1}}{\delta} \Phi_k \left(\frac{\tau}{\delta} \right).
\end{aligned}$$

This equation is Eq. 4.20. Sec. 4.3.2 makes use of this result.

A.5.2 Worked Example 2: Credit and Temporal Proximity

For this example, Event X cues Event Y and Event Z with a fixed delay of time interval $2 - t_{\mathrm{YZ}}$ and 2 respectively (relative to the time of occurrence of X). We work out $\exp C_{\mathrm{Y}}^{\mathrm{Z}}$ as a function of t_{YZ}. The significance of this result is discussed in Sec. 4.3.4.

We need $(p^+)_{\mathrm{Y}}^{\mathrm{Z}}$, which we write as shorthand for "the degree to which Z appears in $\mathbf{p}^+$ due to Y" and $(p^-)^{\mathrm{Z}}$ at the time Y occurs.

$$\left(p^-\right)^{\mathrm{Z}} = \Lambda^{\mathrm{Z}} \exp \sum_{E \in \mathcal{E}} \int C_E^{\mathrm{Z}}(\overset{*}{\tau}) \tilde{f}_{\delta}^{E}(\overset{*}{\tau}; t_{\mathrm{Y}}^-)\, d\overset{*}{\tau},$$

where t_{Y}^- refers to the moment just before Y occurs. At that time, the memory contains only

X which occurred $(2-t_{\mathrm{YZ}})$ ago, so

$$\left(p^{-}\right)^{\mathrm{Z}}=\Lambda^{\mathrm{Z}}\exp\int C_{\mathrm{X}}^{\mathrm{Z}}(\overset{*}{\tau})\tilde{f}_{\delta}^{\mathrm{X}}(\overset{*}{\tau};t_{\mathrm{Y}}^{-})\,d\overset{*}{\tau}. \tag{A.11}$$

The credit due to X for Z is

$$\exp C_{\mathrm{X}}^{\mathrm{Z}}(\delta)=\frac{\Phi_k(2/\overset{*}{\tau})}{\Lambda^{\mathrm{Z}}\overset{*}{\tau}},$$

and the projected memory for X when Y occurs is

$$\tilde{f}_{\delta}^{\mathrm{X}}(\overset{*}{\tau})=\Phi_k\left((\delta+2-t_{\mathrm{YZ}})/\overset{*}{\tau}\right)/\overset{*}{\tau}.$$

Thus, the integral is

$$\int C_{\mathrm{X}}^{\mathrm{Z}}(\overset{*}{\tau})\tilde{f}_{\delta}^{\mathrm{X}}(\overset{*}{\tau};t_{\mathrm{Y}}^{-})\,d\overset{*}{\tau}=\int\frac{1}{\overset{*}{\tau}}\Phi_k\left(\frac{\delta+2-t_{\mathrm{YZ}}}{\overset{*}{\tau}}\right)\log\frac{\Phi_k(2/\overset{*}{\tau})}{\Lambda^{\mathrm{Z}}\overset{*}{\tau}}\,d\overset{*}{\tau}.$$

This integral is the same as Eq. A.10 in the previous example with δ, τ and Λ^{Y} replaced with $\delta+2-t_{\mathrm{YZ}}$, 2 and Λ^{Z} respectively. The same relationship holds between $(p^{-})^{\mathrm{Z}}$ here (Eq. A.11) and p^{Y} in the previous example (Eq. A.9). Referring to the previous result, we thus have

$$\left(p^{-}\right)^{\mathrm{Z}}(\delta)=\frac{\kappa_1^{-1}}{\delta+2-t_{\mathrm{YZ}}}\Phi_k\left(\frac{2}{\delta+2-t_{\mathrm{YZ}}}\right).$$

On the other hand,

$$\left(p^{+}\right)_{\mathrm{Y}}^{\mathrm{Z}}(\delta)=m^{\mathrm{Z}}(\delta)=M_{\mathrm{Y}}^{\mathrm{Z}}(\delta)=\Phi_k(t_{\mathrm{YZ}}/\delta)/\delta.$$

Thus, after training, the credit due to Y for Z is

$$\begin{aligned}
\exp C_{\mathrm{Y}}^{\mathrm{Z}}(\delta) &= \frac{(p^+)_{\mathrm{Y}}^{\mathrm{Z}}(\delta)}{(p^-)^{\mathrm{Z}}(\delta)} \\
&= \frac{\Phi_k(t_{\mathrm{YZ}}/\delta)/\delta}{\frac{\kappa_1^{-1}}{\delta+2-t_{\mathrm{YZ}}}\Phi_k\left(\frac{2}{\delta+2-t_{\mathrm{YZ}}}\right)} \\
&= \frac{(t_{\mathrm{YZ}}/\delta)^k \exp\left(-kt_{\mathrm{YZ}}/\delta\right)/\delta}{\frac{\kappa_1^{-1}}{\delta+2-t_{\mathrm{YZ}}}\left(\frac{2}{\delta+2-t_{\mathrm{YZ}}}\right)^k \exp\left(-\frac{2k}{\delta+2-t_{\mathrm{YZ}}}\right)} \\
&= \kappa_1 \left(\frac{t_{\mathrm{YZ}}}{2}\right)^k \left(\frac{\delta+2-t_{\mathrm{YZ}}}{\delta}\right)^{k+1} \exp\left(-k\left[\frac{t_{\mathrm{YZ}}}{\delta} - \frac{2}{\delta+2-t_{\mathrm{YZ}}}\right]\right).
\end{aligned}$$

Fig. 4.8 plots this expression for various values of t_{YZ}. See Sec. 4.3.4 for a discussion.

A.6 Demonstration: Methods

Given a time-ordered set of events $[e_1, e_2, ..., e_n]$, where each $e_i = (x_i, t_i)$ comprises a discrete-valued type and real-valued timestamp, we are interested in predicting the type x_{n+1} of the next event given its time of occurrence t_{n+1}. In the demonstration, we apply the prediction algorithm ("$\mathcal{C}$-based") to a *superposition* of independent MRPs and compare its predictions to those of a pairwise event association model ("$\mathcal{M}$-based"). In the simulation, both the $\mathcal{C}$- and $\mathcal{M}$-based predictors have memories spanning 10^{-5} to 80 time units into the past, each covered by 200 log-spaced memory nodes.

Within each MRP, the probability of the type *and* time of an event depends solely on the type of the most recent past event, i.e., for MRP k,

$$P((x_{n+1}^k, t_{n+1}^k) | \{(x_m^k, t_m^k)\}_{m \le n}) = P((x_{n+1}^k, t_{n+1}^k) | x_n^k),$$

where $t_{n+1}^k > t_n^k$. The set of event types within each MRP is discrete and finite, while transition times $\Delta t_{n+1} = t_{n+1} - t_n > 0$ are real and strictly positive; this allows only one event to occur at a given time. Within each MRP, the probability of the type of the next event is given by the transition matrix

$$P_{ij} = P(x_{n+1}^k = j | x_n^k = i) = \begin{pmatrix} 0.05 & 0.75 & 0.2 \\ 0.2 & 0.05 & 0.75 \\ 0.75 & 0.2 & 0.05 \end{pmatrix}.$$

The transition times from i to j in MRP k follows a truncated normal distribution $\mathcal{N}(\mu_{ij}^k, \sigma_{ij}^{2k})$, with a lower bound cutoff of 10^{-5} (to ensure positivity).

We use two approaches which generate superposed processes differently. We discuss the first approach, used for Fig. 4.9c. The means μ_{ij}^k and variances σ_{ij}^{2k} of the transition time distributions are drawn uniformly from the intervals $(0,10)$ and $(0,2)$, respectively. The same values are used across all six runs of the simulation. For each run, we generate exactly seven MRPs, labeled $k = 1,\ldots,7$, each with 500 event episodes. We then construct seven superposed processes from the aforementioned MRPs as follows. The first superposed process consists of one MRP, namely, the MRP $k = 1$; the second superposed process consists of two MRPs, namely the MRPs with $k = 1$ and $k = 2$; and so on. Each component MRP has three types of events, so the total number of event types in the superposed process is $3N$, where N is the number of MRPs superposed.

We now discuss the second approach, used for Fig. 4.9d. We draw exactly one set of transition time distribution parameters μ_{ij} and σ_{ij}^2 as before. This same set of parameters

is used across all six runs of the simulation, and for all MRPs $k = 1, \ldots, 7$. We generate exactly seven MRPs of 20,000 event episodes each, and generate seven superposed processes therefrom by incrementally superposing the MRPs as in the first approach. Every MRP has three types of events ($\mathrm{U}, \mathrm{V}, \mathrm{W}$). In the superposed processes, the event types are not distinguished according to the MRP of origin (e.g., a U from one MRP and a U from another MRP are both of type U in the superposed process). Thus, in contrast to the previous approach, the algorithms only observe three types of events in the superposed MRPs.

In both Fig. 4.9c and 4.9d, 80% of each superposed process is used for training and the rest for testing. For the $\mathcal{C}$-based prediction, accuracy on the test set is computed by checking if, at every time t_n that an event occurs, the prediction evaluated at t_{n+1}, the time of the next event, $\mathrm{argmax}_i\ p^i(\delta = \Delta t_{n+1}; t = t_n)$ matches the event that actually occurs at that time. For the $\mathcal{M}$-based prediction, the computation is analogous, except the prediction is found *via* $\mathrm{argmax}_i\ m^i(\delta = \Delta t_{n+1}; t = t_n)$, where $j = x_n$, the type of the event at t_n. The simulation is run 6 times and the average accuracy is reported.

REFERENCES

Altmann, E. G., Cristadoro, G., and Esposti, M. D. (2012). On the origin of long-range correlations in texts. *Proceedings of the National Academy of Sciences*, 109(29):11582–7.

Astrand, E., Ibos, G., Duhamel, J.-R., and Ben Hamed, S. (2015). Differential dynamics of spatial attention, position, and color coding within the parietofrontal network. *Journal of Neuroscience*, 35(7):3174–3189.

Babcock, S. W., Howard, M. W., and McGuire, J. T. (2020). Time-conjunctive representations of future events. *Memory & Cognition*, 48:672–682.

Balsam, P. D., Drew, M. R., and Gallistel, C. R. (2010). Time and associative learning. *Comparative Cognition & Behavior Reviews*, 5:1–22.

Balsam, P. D. and Gallistel, C. R. (2009). Temporal maps and informativeness in associative learning. *Trends in Neuroscience*, 32(2):73–78.

Bateson, M. (2003). *Interval timing and optimal foraging*, pages 113–141. Functional and neural mechanisms of interval timing. CRC Press/Routledge/Taylor & Francis Group, Boca Raton, FL, US.

Ben-Yakov, A., Rubinson, M., and Dudai, Y. (2014). Shifting gears in hippocampus: Temporal dissociation between familiarity and novelty signatures in a single event. *Journal of Neuroscience*, 34(39):12973–12981.

Bernacchia, A., Seo, H., Lee, D., and Wang, X. J. (2011). A reservoir of time constants for memory traces in cortical neurons. *Nature Neuroscience*, 14(3):366–72.

Bolding, K. A., Nagappan, S., Han, B. X., Wang, F., and Franks, K. M. (2020). Recurrent circuitry is required to stabilize piriform cortex odor representations across brain states. *eLife*, 9:e53125.

Bolkan, S. S., Stujenske, J. M., Parnaudeau, S., Spellman, T. J., Rauffenbart, C., Abbas, A. I., Harris, A. Z., Gordon, J. A., and Kellendonk, C. (2017). Thalamic projections sustain prefrontal activity during working memory maintenance. *Nature Neuroscience*, 20(7):987–996.

Brenner, N., Strong, S. P., Koberle, R., Bialek, W., and de Ruyter van Steveninck, R. R. (2000). Synergy in a neural code. *Neural Computation*, 12(7):1531–1552.

Bright, I. M., Meister, M. L. R., Cruzado, N. A., Tiganj, Z., Buffalo, E. A., and Howard, M. W. (2020). A temporal record of the past with a spectrum of time constants in the monkey entorhinal cortex. *Proceedings of the National Academy of Sciences*, 117:20274–20283.

Buzsáki, G. (2002). Theta oscillations in the hippocampus. *Neuron*, 33(3):325–40.

Cardinal, R. N., Pennicott, D. R., Lakmali, C., Sugathapala, Robbins, T. W., and Everitt, B. J. (2001). Impulsive choice induced in rats by lesions of the nucleus accumbens core. *Science*, 292(5526):2499–2501.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36(03):181–204.

Cohen, R., Erez, K., Ben-Avraham, D., and Havlin, S. (2000). Resilience of the internet to random breakdowns. *Physical Review Letters*, 85(21):4626.

Cont, R. (2005). Long range dependence in financial markets. In Lévy-Véhel, J. and Lutton, E., editors, *Fractals in Engineering*, pages 159–179, London. Springer.

Crowe, D. A., Averbeck, B. B., and Chafee, M. V. (2010). Rapid sequences of population activity patterns dynamically encode task-critical spatial information in parietal cortex. *Journal of Neuroscience*, 30(35):11640–11653.

Cruzado, N. A., Tiganj, Z., Brincat, S. L., Miller, E. K., and Howard, M. W. (2020). Conjunctive representation of what and when in monkey hippocampus and lateral prefrontal cortex during an associative memory task. *Hippocampus*, 30(12):1332–1346.

Daw, N. D., Courville, A. C., and Touretzky, D. S. (2006). Representation and timing in theories of the dopamine system. *Neural Computation*, 18(7):1637–1677.

Dayan, P. (1993). Improving generalization for temporal difference learning: The successor representation. *Neural Computation*, 5(4):613–624.

DLMF (2021). NIST Digital Library of Mathematical Functions. http://dlmf.nist.gov/, Release 1.1.2 of 2021-06-15. F. W. J. Olver, A. B. Olde Daalhuis, D. W. Lozier, B. I. Schneider, R. F. Boisvert, C. W. Clark, B. R. Miller, B. V. Saunders, H. S. Cohl, and M. A. McClain, eds.

Dong, C., Madar, A. D., and Sheffield, M. E. J. (2021). Distinct place cell dynamics in CA1 and CA3 encode experience in new environments. *Nature Communications*, 12(1):2977.

Eichenbaum, H. (2014). Time cells in the hippocampus: a new dimension for mapping memories. *Nature Reviews Neuroscience*, 15(11):732–44.

Eichenbaum, H., Sauvage, M., Fortin, N., Komorowski, R., and Lipton, P. (2012). Towards a functional organization of episodic memory in the medial temporal lobe. *Neuroscience & Biobehavioral Reviews*, 36(7):1597–1608. Memory Formation.

Einstein, G. O., McDaniel, M. A., Richardson, S. L., Guynn, M. J., and Cunfer, A. R. (1995). Aging and prospective memory: Examining the influences of self-initiated retrieval processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21(4):996–1007.

Fiorillo, C. D., Newsome, W. T., and Schultz, W. (2008). The temporal precision of reward prediction in dopamine neurons. *Nature Neuroscience*, 11:966–973.

Frank, L. M., Stanley, G. B., and Brown, E. N. (2004). Hippocampal plasticity across multiple days of exposure to novel environments. *Journal of Neuroscience*, 24(35):7681–7689.

Fried, I., MacDonald, K. A., and Wilson, C. L. (1997). Single neuron activity in human hippocampus and amygdala during recognition of faces and objects. *Neuron*, 18(5):753–765.

Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, 11:127–138.

Frost, R. and McNaughton, N. (2017). The neural basis of delay discounting: A review and preliminary model. *Neuroscience & Biobehavioral Reviews*, 79:48–65.

Gallistel, C., Craig, A. R., and Shahan, T. A. (2019). Contingency, contiguity, and causality in conditioning: Applying information theory and Weber's law to the assignment of credit problem. *Psychological Review*, 126(5):761.

Glimcher, P. W. (2011). Understanding dopamine and reinforcement learning: The dopamine reward prediction error hypothesis. *Proceedings of the National Academy of Sciences*, 108(Supplement 3):15647–15654.

Goh, W. Z., Ursekar, V., and Howard, M. W. (In press). Predicting the future with a scale-invariant temporal memory for the past. *Neural Computation*.

Gretenkord, S., Kostka, J. K., Hartung, H., Watznauer, K., Fleck, D., Minier-Toribio, A., Spehr, M., and Hanganu-Opatz, I. L. (2019). Coordinated electrical activity in the olfactory bulb gates the oscillatory entrainment of entorhinal networks in neonatal mice. *PLoS Biology*, 17(1):e2006994.

Hainmueller, T. and Bartos, M. (2020). Dentate gyrus circuits for encoding, retrieval and discrimination of episodic memories. *Nature Reviews Neuroscience*, 21(3):153–168.

Hasselmo, M. E., Bodelón, C., and Wyble, B. P. (2002). A proposed function for hippocampal theta rhythm: Separate phases of encoding and retrieval enhance reversal of prior learning. *Neural Computation*, 14:793–817.

Hollerman, J. R. and Schultz, W. (1998). Dopamine neurons report an error in the temporal prediction of reward during learning. *Nature Neuroscience*, 1(4):304–309.

Imamura, F., Ito, A., and LaFever, B. J. (2020). Subpopulations of projection neurons in the olfactory bulb. *Frontiers in Neural Circuits*, 14:561822.

Iravani, B., Arshamian, A., Lundqvist, M., Kay, L. M., Wilson, D. A., and Lundström, J. N. (2021). Odor identity can be extracted from the reciprocal connectivity between olfactory bulb and piriform cortex in humans. *NeuroImage*, 237:118130.

K Namboodiri, V. M. (2021). What is the state space of the world for real animals? *bioRxiv*.

K Namboodiri, V. M. and Stuber, G. D. (2021). The learning of prospective and retrospective cognitive maps within neural circuits. *Neuron*, 109(22):3552–3575.

Kahana, M. J., Seelig, D., and Madsen, J. R. (2001). Theta returns. *Current Opinion in Biology*, 11(6):739–44.

Kamiński, J., Sullivan, S., Chung, J. M., Ross, I. B., Mamelak, A. N., and Rutishauser, U. (2017). Persistently active neurons in human medial frontal and medial temporal lobe support working memory. *Nature Neuroscience*, 20(4):590.

King, J. and Dehaene, S. (2014). Characterizing the dynamics of mental representations: the temporal generalization method. *Trends in Cognitive Sciences*, 18(4):203–210.

Kurth-Nelson, Z. and Redish, A. D. (2009). Temporal-difference reinforcement learning with distributed representations. *PLoS One*, 4(10):e7362.

Larkin, M. C., Lykken, C., Tye, L. D., Wickelgren, J. G., and Frank, L. M. (2014). Hippocampal output area CA1 broadcasts a generalized novelty signal during an object-place recognition task. *Hippocampus*, 24(7):773–783.

Ledoit, O. and Wolf, M. (2004). A well-conditioned estimator for large-dimensional covariance matrices. *Journal of Multivariate Analysis*, 88(2):365–411.

Lee, S. W., O'Doherty, J. P., and Shimojo, S. (2015). Neural computations mediating one-shot learning in the human brain. *PLoS Biology*, 13(4):e1002137.

Leszczynski, M. (2011). How does hippocampus contribute to working memory processing? *Frontiers in Human Neuroscience*, 5:168.

Liu, S. (2020). Dopamine suppresses synaptic responses of fan cells in the lateral entorhinal cortex to olfactory bulb input in mice. *Frontiers in Cellular Neuroscience*, 14:181.

Ludvig, E. A., Sutton, R. S., and Kehoe, E. J. (2008). Stimulus representation and the timing of reward-prediction errors in models of dopamine system. *Neural Computation*, 20:3034–3054.

Ludvig, E. A., Sutton, R. S., and Kehoe, E. J. (2012). Evaluating the TD model of classical conditioning. *Learning & Behavior*, 40(3):305–319.

MacDonald, C. J., Lepage, K. Q., Eden, U. T., and Eichenbaum, H. (2011). Hippocampal "time cells" bridge the gap in memory for discontiguous events. *Neuron*, 71(4):737–749.

Major, G. and Tank, D. (2004). Persistent neural activity: prevalence and mechanisms. *Current Opinion in Neurobiology*, 14(6):675–84.

Masse, N. Y., Rosen, M. C., and Freedman, D. J. (2020). Reevaluating the role of persistent neural activity in short-term memory. *Trends in Cognitive Sciences*, 24(3):242–258.

McGuire, J. T. and Kable, J. W. (2013). Rational temporal predictions can underlie apparent failures to delay gratification. *Psychological Review*, 120(2):395–410.

Mello, G. B., Soares, S., and Paton, J. J. (2015). A scalable population code for time in the striatum. *Current Biology*, 25(9):1113–1122.

Millot, J. L., Laurent, L., and Casini, L. (2016). The influence of odors on time perception. *Frontiers in Psychology*, 7:181.

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., Graves, A., Riedmiller, M., Fidjeland, A. K., Ostrovski, G., et al. (2015). Human-level control through deep reinforcement learning. *Nature*, 518(7540):529–533.

Momennejad, I. and Howard, M. W. (2018). Predicting the future with multi-scale successor representations. *bioRxiv*.

Morcos, A. S. and Harvey, C. D. (2016). History-dependent variability in population dynamics during evidence accumulation in cortex. *Nature Neuroscience*, 19(12):1672–1681.

Morris, G., Arkadir, D., Nevet, A., Vaadia, E., and Bergman, H. (2004). Coincident but distinct messages of midbrain dopamine and striatal tonically active neurons. *Neuron*, 43(1):133–143.

Mozer, M. C. (1992). Induction of multiscale temporal structure. In *Advances in neural information processing systems*, pages 275–282.

Muchnik, C., Hildesheimer, M., Rubinstein, M., Sadeh, M., Shegter, Y., and Shibolet, B. (1985). Minimal time interval in auditory temporal resolution. *Journal of Auditory Research*, 25(4):239–246.

Murray, J. D., Bernacchia, A., Roy, N. A., Constantinidis, C., Romo, R., and Wang, X.-J. (2017). Stable population coding for working memory coexists with heterogeneous neural dynamics in prefrontal cortex. *Proceedings of the National Academy of Sciences*, 114(2):394–399.

Nakahara, H., Itoh, H., Kawagoe, R., Takikawa, Y., and Hikosaka, O. (2004). Dopamine neurons can represent context-dependent prediction error. *Neuron*, 41(2):269–280.

Pan, W. X., Schmidt, R., Wickens, J. R., and Hyland, B. I. (2005). Dopamine cells respond to predicted events during classical conditioning: evidence for eligibility traces in the reward-learning network. *Journal of Neuroscience*, 25(26):6235–42.

Pastalkova, E., Itskov, V., Amarasingham, A., and Buzsaki, G. (2008). Internally generated cell assembly sequences in the rat hippocampus. *Science*, 321(5894):1322–7.

Pilkiw, M., Insel, N., Cui, Y., Finney, C., Morrissey, M. D., and Takehara-Nishiuchi, K. (2017). Phasic and tonic neuron ensemble codes for stimulus-environment conjunctions in the lateral entorhinal cortex. *eLife*, 6:e28611.

Post, E. (1930). Generalized differentiation. *Transactions of the American Mathematical Society*, 32:723–781.

Potocanac, Z. and Duysens, J. (2017). Online adjustments of leg movements in healthy young and old. *Experimental Brain Research*, 235(8):2329–2348.

Rasmussen, J. G. (2018). Lecture notes: Temporal point processes and the conditional intensity function. *arXiv preprint 1806.00221*.

Rigotti, M., Barak, O., Warden, M. R., Wang, X.-J., Daw, N. D., Miller, E. K., and Fusi, S. (2013). The importance of mixed selectivity in complex cognitive tasks. *Nature*, 497(7451):585–90.

Salz, D. M., Tiganj, Z., Khasnabish, S., Kohley, A., Sheehan, D., Howard, M. W., and Eichenbaum, H. (2016). Time cells in hippocampal area CA3. *Journal of Neuroscience*, 36:7476–7484.

Schultz, W. (2006). Behavioral theories and the neurophysiology of reward. *Annual Review of Psychology*, 57:87–115.

Schultz, W., Dayan, P., and Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, 275:1593–1599.

Shankar, K. H. and Howard, M. W. (2012). A scale-invariant internal representation of time. *Neural Computation*, 24(1):134–193.

Shankar, K. H. and Howard, M. W. (2013). Optimally fuzzy temporal memory. *Journal of Machine Learning Research*, 14:3753–3780.

Shankar, K. H., Singh, I., and Howard, M. W. (2016). Neural mechanism to simulate a scale-invariant future. *Neural Computation*, 28:2594–2627.

Shimbo, A., Izawa, E.-I., and Fujisawa, S. (2021). Scalable representation of time in the hippocampus. *Science Advances*, 7(6):eabd7013.

Shintaki, R., Tanaka, D., Suzuki, S., Yoshimoto, T., Sadato, N., Chikazoe, J., and Jimura, K. (2021). Anticipatory dynamics in the human brain guide foraging for primary rewards. *bioRxiv*.

Silver, D., Hubert, T., Schrittwieser, J., Antonoglou, I., Lai, M., Guez, A., Lanctot, M., Sifre, L., Kumaran, D., Graepel, T., et al. (2018). A general reinforcement learning algorithm that masters chess, shogi, and go through self-play. *Science*, 362(6419):1140–1144.

Stokes, M. G., Kusunoki, M., Sigala, N., Nili, H., Gaffan, D., and Duncan, J. (2013). Dynamic coding for cognitive control in prefrontal cortex. *Neuron*, 78(2):364–375.

Sutton, R. S. (1988). Learning to predict by the methods of temporal differences. *Machine Learning*, 3(1):9–44.

Sutton, R. S. and Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT Press.

Suzuki, W. A., Miller, E. K., and Desimone, R. (1997). Object and place memory in the macaque entorhinal cortex. *Journal of Neurophysiology*, 78(2):1062–1081.

Tanninen, S. E., Morrissey, M. D., and Takehara-Nishiuchi, K. (2013). Unilateral lateral entorhinal inactivation impairs memory expression in trace eyeblink conditioning. *PLoS One*, 8(12):e84543.

Tano, P., Dayan, P., and Pouget, A. (2020). A local temporal difference code for distributional reinforcement learning. *Advances in Neural Information Processing Systems*, 33:13662–13673.

Taxidis, J., Pnevmatikakis, E. A., Dorian, C. C., Mylavarapu, A. L., Arora, J. S., Samadian, K. D., Hoffberg, E. A., and Golshani, P. (2020). Differential emergence and stability of sensory and temporal representations in context-specific hippocampal sequences. *Neuron*, 108(5):984–998.e9.

Tiganj, Z., Cromer, J. A., Roy, J. E., Miller, E. K., and Howard, M. W. (2018). Compressed timeline of recent experience in monkey lPFC. *Journal of Cognitive Neuroscience*, 30:935–950.

Tiganj, Z., Gershman, S. J., Sederberg, P. B., and Howard, M. W. (2019). Estimating scale-invariant future in continuous time. *Neural Computation*, 31(4):681–709.

Tishby, N., Pereira, F. C., and Bialek, W. (2000). The information bottleneck method. *arXiv preprint physics/0004057*.

Toda, K., Lusk, N. A., Watson, G. D., Kim, N., Lu, D., Li, H. E., Meck, W. H., and Yin, H. H. (2017). Nigrotectal stimulation stops interval timing in mice. *Current Biology*, 27(24):3763–3770.e3.

Tsao, A., Sugar, J., Lu, L., Wang, C., Knierim, J. J., Moser, M.-B., and Moser, E. I. (2018). Integrating time from experience in the lateral entorhinal cortex. *Nature*, 561:57–62.

van der Meer, M. A. A. and Redish, A. D. (2011). Theta phase precession in rat ventral striatum links place and reward information. *Journal of Neuroscience*, 31(8):2843–54.

Vanderveldt, A., Oliveira, L., and Green, L. (2016). Delay discounting: Pigeon, rat, human—does it matter? *Journal of Experimental Psychology: Animal Learning and Cognition*, 42(2):141–162.

Waelti, P., Dickinson, A., and Schultz, W. (2001). Dopamine responses comply with basic assumptions of formal learning theory. *Nature*, 412(6842):43–8.

Watkins, C. J. and Dayan, P. (1992). Q-learning. *Machine Learning*, 8(3-4):279–292.

Weger, U. W. and Pratt, J. (2008). Time flies like an arrow: Space-time compatibility effects suggest the use of a mental timeline. *Psychonomic Bulletin & Review*, 15(2):426–430.

Woods, N. I., Stefanini, F., Apodaca-Montano, D. L., Tan, I. M., Biane, J. S., and Kheirbek, M. A. (2020). The dentate gyrus classifies cortical representations of learned stimuli. *Neuron*, 107:173–184.

Wouterlood, F. and Nederlof, J. (1983). Terminations of olfactory afferents on layer II and III neurons in the entorhinal area: Degeneration-golgi-electron microscopic study in the rat. *Neuroscience Letters*, 36(2):105–110.

Xu, M., Zhang, S.-y., Dan, Y., and Poo, M.-m. (2014). Representation of interval timing by temporally scalable firing patterns in rat prefrontal cortex. *Proceedings of the National Academy of Sciences*, 111(1):480–485.

Yue, Z., Gao, T., Chen, L., and Wu, J. (2016). Odors bias time perception in visual and auditory modalities. *Frontiers in Psychology*, 7:535.

Zhang, X., Yan, W., Wang, W., Fan, H., Hou, R., Chen, Y., Chen, Z., Ge, C., Duan, S., Compte, A., and Li, C. T. (2019). Active information maintenance in working memory by a sensory cortex. *eLife*, 8:e43191.

Zhou, B., Feng, G., Chen, W., and Zhou, W. (2017). Olfaction warps visual time perception. *Cerebral Cortex*, 28(5):1718–1728.

Zhou, P., Resendez, S. L., Rodriguez-Romaguera, J., Jimenez, J. C., Neufeld, S. Q., Giovannucci, A., Friedrich, J., Pnevmatikakis, E. A., Stuber, G. D., Hen, R., Kheirbek, M. A., Sabatini, B. L., Kass, R. E., and Paninski, L. (2018). Efficient and accurate extraction of in vivo calcium signals from microendoscopic video data. *eLife*, 7:e28728.

CURRICULUM VITAE