

Time cell encoding is decoupled from time perception in deep reinforcement learning agents

Ann Zixiang Huang (zixiang.huang@mail.mcgill.ca)

McGill University, Mila
Montreal, QC, Canada

Dongyan Lin (lindongy@mila.quebec)

McGill University, Mila
Montreal, QC, Canada

Blake Richards (blake.richards@mila.quebec)

McGill University, Mila
Montreal, QC, Canada



Abstract

Time cells are shown to encode the unfolding of time by forming trial-consistent temporal receptive fields that are selective to particular moments. However, it remains unknown whether besides carrying temporal information, time cells directly contribute to the cognitive function of timing. Here, by training deep reinforcement learning (DRL) agents to compare the duration of two sequentially presented stimuli, we show that time cells naturally emerge and encode time elapsed regardless of the cognitive demand of timing. Furthermore, the temporal receptive field of individual cells does not rescale across different stimulus duration nor discriminates between correct and incorrect trials, suggesting a dissociation between time encoding and duration judgment in the DRL agent. Together, our findings posit that time encoding may emerge as an intrinsic circuit phenomenon of recurrent neural networks irrespective of the cognitive function of timing.

Keywords: Time cells; time perception; deep reinforcement learning; recurrent neural network

Time cells are ubiquitously discovered throughout the brain in multiple mammalian species (Akhlagpour et al., 2016; Bakhurin et al., 2017; Cruzado, Tiganj, Brincat, Miller, & Howard, 2020; C. MacDonald, Lepage, Eden, & Eichenbaum, 2011; C. J. MacDonald, Carrow, Place, & Eichenbaum, 2013; Schonhaut, Aghajan, Kahana, & Fried, 2022; Tsao et al., 2018). Yet, it is unknown whether they contribute to the cognitive function of timing. Particularly, Toso et al. (2021) demonstrated a distinction between time coding and time perception in the dorsolateral striatum of rats tasked with comparing the duration of two sequential vibrations (Toso, Reinartz, Pulecchi, & Diamond, 2021).

Recently, several studies have shown that recurrent neural networks could successfully solve behavioral timing tasks (Deverett, Faulkner, Fortunato, Wayne, & Leibo, 2019; Hardy & Buonomano, 2018). During the task, the artificial recurrent units developed time-dependent activity resembling the temporal representations in the brain (C. MacDonald et al., 2011; Janssen & Shadlen, 2005; Jazayeri & Shadlen, 2015), suggesting a shared computational principle between the brain and the artificial neural network for temporal processing. This motivated us to train DRL agents with hidden recurrent connections on a neuroscience-based timing task and study time encoding in the recurrent neural network.

We simulated a Delayed Duration Comparison (DDC) task wherein a location-fixed reinforcement learning agent received two visual stimuli of varying length presented sequentially, separated by a fixed-length delay period (Fig. 1A). The duration of stimulus 1 and stimulus 2 was sampled uniformly among 7 equally spaced stimulus lengths between 10 seconds and 40 seconds (10, 15, 20, and so on) under the constraint that the two stimuli must have different durations. After stimulus presentation, the agent would make a response to indicate which stimulus was perceived longer and received a reward for the correct response. The DRL agent was a deep neural network with a 512-unit LSTM (i.e. recurrent) module

to process sensory input and an actor-critic module to judge the utility of the current sensory state and select actions (Fig. 1B). The agent was trained with policy gradient methods. We recorded the activities from all LSTM recurrent units during the stimulus presentation phase and the delay phase for the last 5,000 trials of training.

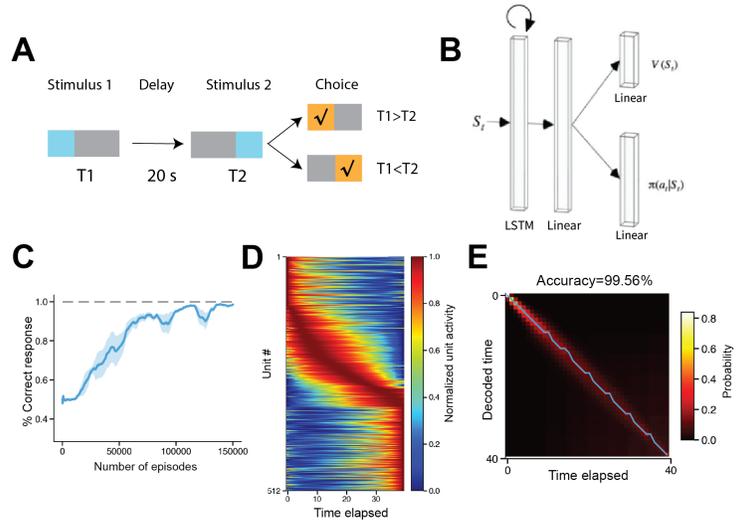


Figure 1: **A)** DDC task structure where T1 and T2 denote the duration of stimulus 1 and stimulus 2, respectively. **B)** Architecture of the DRL agent. S_t represents the sensory state of the environment at time step t , $V(S_t)$ represents the agent’s judgment of the value of the current sensory state, and $\pi(a|S_t)$ selects an action a based on its assessments of action utilities at the current state S_t . **C)** Behavioral performance measured as the fraction of correctly responded episodes. The solid line and shaded area represent the average and standard deviation of performance over four agents. **D)** An example heatmap showing the hidden state activities during the stimulus presentation phase, averaged across all presentation period where the stimulus duration is 40 seconds. Each row shows the trial-averaged activities of a single cell normalized to its minimum (blue) and maximum (red). Rows are sorted by the latency to the cell’s peak activity. **E)** Multi-class logistic regression decoding of elapsed time since stimulus onset from the population hidden state activity at each time step. Heatmap shows the probability of predicted time plotted against actual elapsed time, with superimposed blue lines representing the decoded time with the highest probability estimate.

Results

Time cells emerge from DRL agents optimized on the DDC task. After training, the DRL agent reliably performed the DDC task with almost perfect response accuracy (Fig. 1C). We sorted each cell’s trial-averaged activity by the latency to its peak activity. We found that individual LSTM cells fire selectively at a particular time elapsed, and the population activity of LSTM cells formed a continuous sequence that tiled

the entire duration of the stimulus presentation period (Fig. 1D), resembling the characteristics of time cells in experimental studies (C. MacDonald et al., 2011; C. J. MacDonald et al., 2013). Using multi-class logistic regression, we decoded the external time elapsed purely from the population activity of LSTM cells at each time point during the stimulus presentation phase with 99.56% accuracy (Fig. 1E), demonstrating that LSTM cells encoded the passage of time with high temporal precision during stimulus presentation.

Single unit maintains temporal receptive field regardless of the actual stimulus duration. To assess by which mechanism time cells keep track of the time elapsed, we analyzed the temporal receptive field of each cell during the stimulus presentation phase. We found that the vast majority of the cells were tuned globally to the most prolonged stimulus duration and did not rescale across different stimulus duration (Fig. 2A), which meant that individual cells maintained their temporal receptive field regardless of the actual stimulus duration. It demonstrated that the perception of the stimulus duration did not affect the temporal tuning of each LSTM unit, which implied that the LSTM cells tracked the absolute passage of time independent of the perception of the stimulus duration.

Time cell encoding does not discriminate between correct and error trials. To further test the hypothesis that time encoding is decoupled from time perception in the DRL agent, we examined time encoding and the population activity separately for correct and incorrect trials. When trained on population activity at each time step during stimulus presentation only in correct trials, the multi-class logistic regression decoder predicted the time elapsed during both the correct trials and the incorrect trials with comparable accuracy (Fig. 2B), suggesting that the population activity of time cells did not encode stimulus duration differently on correct versus incorrect trials. Hence, there was a dissociation between the temporal information carried by the population activity of time cells and the agent's perceptual decision regarding the relative duration of sensory stimuli. In other words, during stimulus presentation, time cell encoding can be decoupled from the perceived stimulus duration, which manifest in the agent's behavioral choices. Furthermore, the temporal tuning of time cells at a population level did not differ qualitatively between correct and incorrect trials (Fig. 2C).

Time cells track time passage regardless of the cognitive demand of timing. Finally, we reason that if time cells contribute to time perception in DRL agents, their temporal tuning properties (e.g., temporal resolution) should adapt to the cognitive demand of timing and hence show a privileged representation of task-relevant time span (i.e., the stimulus duration rather than the delay duration). However, during the delay period, where the cognitive demand for timing is eliminated, the trial-averaged activity of LSTM cells still peaked at successive moments and tiles the entire delay period (Fig. 2D). Remarkably, the temporal resolution of time encoding in LSTM cells did not diminish during the delay, as a multi-class logistic re-

gression decoder achieved perfect accuracy when decoding the time elapsed from the delay-period population activity at each time step (Fig. 2E). This suggests that time encoding may emerge as an intrinsic circuit property of recurrent neural networks irrespective of the cognitive demand of timing.

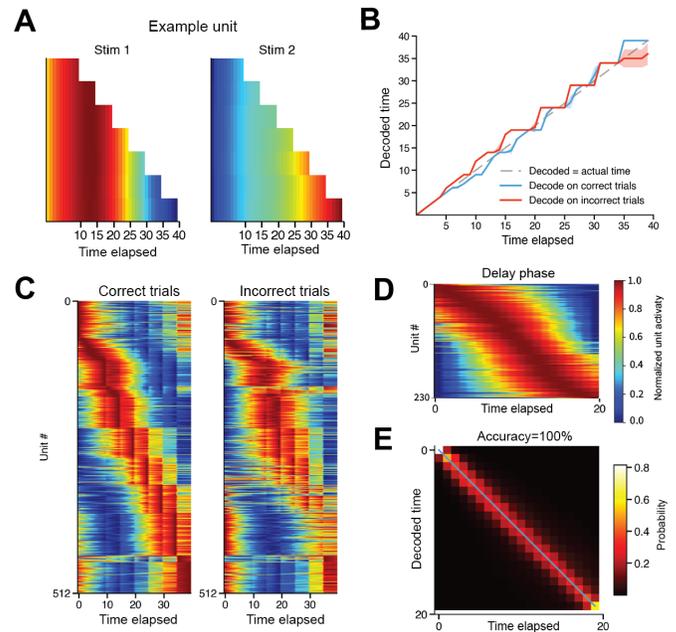


Figure 2: **A)** The temporal receptive field of an example cell. Each row shows the normalized unit activity averaged across all trials wherein the corresponding stimulus has the duration indicated by the x-axis. **B)** The decoded time plotted against the actual elapsed time using a multi-class logistic regression decoder trained only on correct trials. **C)** The population activity averaged across all stimulus duration for correct trials versus incorrect trials. Rows in both panels are sorted according to the cell's peak activity during the correct trials. **D)** The trial-averaged population activity during the delay phase forms a continuous sequence. Only cells with a temporal receptive field other than the beginning or end of the delay phase are shown. **E)** Heatmap showing the decoded time during the delay period.

Discussions

In this work, we trained DRL agents on a delayed duration comparison task and showed a dissociation between the encoding of external time elapsed and the perceived stimulus duration in the hidden recurrent network. Our findings suggest that rather than directly contributing to the cognitive function of duration perception, time cells serve as an intrinsic circuit phenomenon to track the unfolding of time. Our modeling study elucidates the role of time cells for temporal processing in recurrent neural networks trained on behavioral tasks and provides concrete predictions for future neurophysiological studies regarding the role of time cells in time perception.

Acknowledgments

This research is supported by grants to B.A.R from NSERC (Discovery Grant: RGPIN-202017 05105; Discovery Accelerator Supplement: RGPAS-2020-00031), Healthy Brains, Healthy Lives 18 (New Investigator Award: 2bNISU-8), and CIFAR (Canada AI Chair; Learning in Machine and 19 Brains Fellowship). Additionally, D.L. was supported by an IVADO Excellence Scholarship 20 (MSc-2020-0435588519).

References

- Akhlaghpour, H., Wiskerke, J., Choi, J. Y., Taliaferro, J. P., Au, J., & Witten, I. B. (2016). Dissociated sequential activity and stimulus encoding in the dorsomedial striatum during spatial working memory. *Elife*, *5*, e19507. (Publisher: eLife Sciences Publications Limited)
- Bakhurin, K. I., Goudar, V., Shobe, J. L., Claar, L. D., Buonomano, D. V., & Masmanidis, S. C. (2017). Differential encoding of time by prefrontal and striatal network dynamics. *Journal of Neuroscience*, *37*(4), 854–870. (Publisher: Soc Neuroscience)
- Cruzado, N. A., Tiganj, Z., Brincat, S. L., Miller, E. K., & 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. (Publisher: Wiley Online Library)
- Deverett, B., Faulkner, R., Fortunato, M., Wayne, G., & Leibo, J. Z. (2019, December). *Interval timing in deep reinforcement learning agents*. arXiv. (Number: arXiv:1905.13469 arXiv:1905.13469 [cs])
- Hardy, N. F., & Buonomano, D. V. (2018). Encoding time in feedforward trajectories of a recurrent neural network model. *Neural computation*, *30*(2), 378–396. (ISBN: 0899-7667 Publisher: MIT Press One Rogers Street, Cambridge, MA 02142-1209, USA journals-info . . .)
- Janssen, P., & Shadlen, M. N. (2005). A representation of the hazard rate of elapsed time in macaque area LIP. *Nature neuroscience*, *8*(2), 234–241. (Publisher: Nature Publishing Group)
- Jazayeri, M., & Shadlen, M. N. (2015). A neural mechanism for sensing and reproducing a time interval. *Current Biology*, *25*(20), 2599–2609. (Publisher: Elsevier)
- MacDonald, C., Lepage, K., Eden, U., & Eichenbaum, H. (2011, August). Hippocampal “Time Cells” Bridge the Gap in Memory for Discontiguous Events. *Neuron*, *71*(4), 737–749. doi: 10.1016/j.neuron.2011.07.012
- MacDonald, C. J., Carrow, S., Place, R., & Eichenbaum, H. (2013). Distinct hippocampal time cell sequences represent odor memories in immobilized rats. *Journal of Neuroscience*, *33*(36), 14607–14616. (ISBN: 0270-6474 Publisher: Soc Neuroscience)
- Schönhaut, D. R., Aghajan, Z. M., Kahana, M. J., & Fried, I. (2022). A neural code for spatiotemporal context. *bioRxiv*. (Publisher: Cold Spring Harbor Laboratory)
- Toso, A., Reinartz, S., Pulecchi, F., & Diamond, M. E. (2021, November). Time coding in rat dorsolateral striatum. *Neuron*, *109*(22), 3663–3673.e6.
- Tsao, A., Sugar, J., Lu, L., Wang, C., Knierim, J. J., Moser, M.-B., & Moser, E. I. (2018). Integrating time from experience in the lateral entorhinal cortex. *Nature*, *561*(7721), 57–62. (Publisher: Nature Publishing Group)