Learning abstract features with deep RL agents in an evidence accumulation task

James Mochizuki-Freeman **Md Rysul Kabir** **Zoran Tiganj**
Department of Computer Science, Indiana University Bloomington

Abstract

Recent neuroscience studies suggest that the hippocampus encodes a low-dimensional ordered representation of evidence through sequential neural activity. Cognitive modelers have proposed a mechanism by which such sequential activity could emerge through the modulation of the decay rate of neurons with exponentially decaying firing profiles. Through a linear transformation, this representation gives rise to neurons tuned to a specific magnitude of evidence, resembling neurons recorded in the hippocampus. Here we integrated this cognitive model inside reinforcement learning agents and trained the agents to perform an evidence accumulation task designed to mimic a task used in experiments on animals. We found that the agents were able to learn the task and exhibit sequential neural activity as a function of the amount of evidence, similar to the activity reported in the hippocampus.

Keywords: Evidence accumulation; Cognitive model; Deep RL; Neural sequences; Abstract features

Introduction

Evidence accumulation is a central concept in cognitive psychology and neuroscience that describes the process by which the brain integrates sensory information over time to make decisions. A recent study by Nieh et al. (2021) found that neurons in the hippocampus encode the amount of evidence such that different neurons are tuned to a different magnitude of evidence. Specifically, Nieh et al. (2021) trained mice on the "accumulating towers task" where while mice moved along a virtual track, objects (referred to as "towers") appeared on both sides of the track. When they arrived at the end of the track, to earn a reward mice had to choose the left- or right-hand side, depending on which side had more towers. The difference in the number of towers is an abstract latent variable that corresponds to the amount of evidence for either of the two options. Nieh et al. (2021) recorded the activity of hundreds of individual neurons from the dorsal CA1 sub-region of mice hippocampus while they performed the accumulating towers task. The results indicated the existence of cells tuned to a particular difference in the number of towers, such that a population of neurons tiled the entire *evidence* axis. This neural coding scheme based on sequences resembles coding of other variables in the hippocampus, specifically time and space through time cells (Pastalkova, Itskov, Amarasingham, & Buzsaki, 2008; MacDonald, Lepage, Eden, & Eichenbaum, 2011; Salz et al., 2016) and place cells (Bures, Fenton, Kaminsky, & Zinyuk, 1997) respectively (Eichenbaum, 2014).

Here we sought to construct artificial agents that can learn to represent abstract variables, in this case the amount of evidence, using the same coding scheme reported in Nieh et al. (2021). The agents receive pixel inputs and reward signal similar to those of the mice in the accumulating towers task. We base our approach on a computational cognitive framework that proposed a unified representation for coding time, space, and sequences in the hippocampus (Howard et al., 2014). The framework uses a set of neurons that perform leaky integration (**f** layer) of the input, each with a different time constant (approximation of a Laplace transform). Each such neuron has an impulse response that decays exponentially as a function of time. The output of the leaky integrators is transformed into a set of sequentially activated neurons ($\tilde{\mathbf{f}}$ layer) through a linear transformation that resembles lateral inhibition, which is an approximation of an inverse Laplace transform (Shankar & Howard, 2012). Importantly, if the decay rate of the leaky integrators is modulated by a time derivative of some variable, then the decay becomes an exponential function of that variable (Fig. 2). For example, if the modulator is the time derivative of distance (velocity), then the impulse response of the leaky integrators is an exponential function of distance (place cells). Applying the same linear transformation to the exponential functions of distance gives rise to neurons that sequentially activate as a function of distance, not time. An extension of this work (Howard, Luzardo, & Tiganj, 2018) demonstrated that modulating the decay rate by the change in the amount of evidence gives rise to neurons tuned to a particular magnitude of evidence. Previous work (Mochizuki-Freeman, Maini, & Tiganj, 2023) has used this framework to train deep learning agents on a simple version of the accumulating towers task. The simple version of the task did not use a realistic visual environment but directly provided relevant features to the agent. Here we demonstrate that the network can backpropagate through the cognitive model and learn the task-relevant abstract features.

Methods

We used the accumulating towers task environment described in Mochizuki-Freeman, Kabir, Gulecha, and Tiganj (2023) that closely followed the design of the environment used in Nieh et al. (2021). The agent received realistic visual inputs that were then passed through an encoder composed of several convolutional layers and a single fully connected layer. The outputs of the encoder modulated the decay rate of leaky integrators in the evidence accumulation module, the input of which was set to a delta pulse at the beginning of each trial. The output of the leaky integrators underwent the same linear transformation described in (Howard et al., 2014; Shankar & Howard, 2012) and it was followed by another dense layer. Finally, the output of the dense layer was fed into the RL module based on the A2C architecture (Mnih et al., 2016).

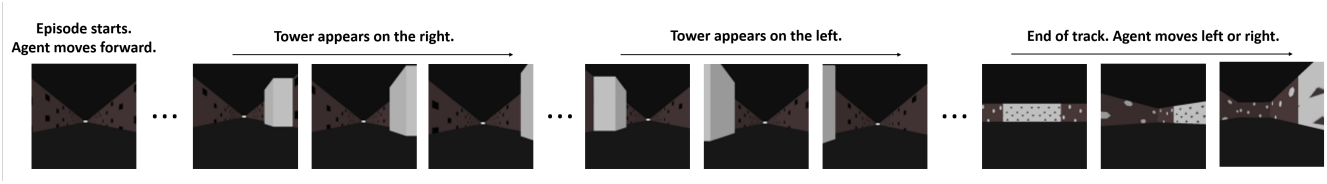

Figure 1: The accumulating towers task. In each trial, the agent moves down a narrow corridor and observes "towers" (white objects) on the left- and right-hand sides of the wall. To obtain the reward, the agent needs to turn left or right at the end of the corridor, depending on which side has more towers.

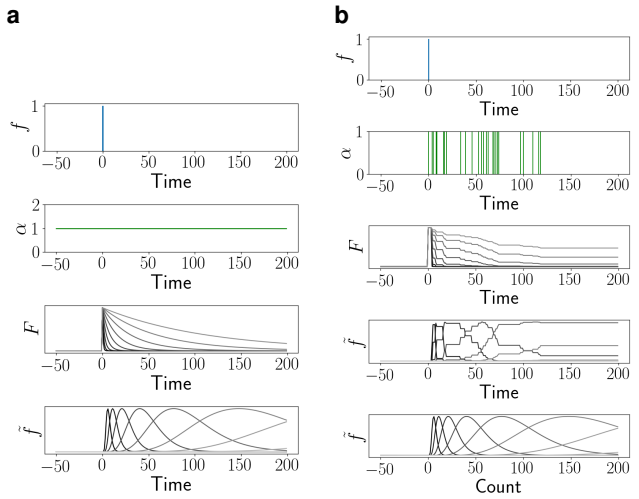

Figure 2: Example of the Laplace and inverse Laplace transform without (a) and with modulatory input α (b). α represents change in the count (e.g., number of "towers") and turns temporal sequences into sequences of count.

Results and Discussion

We trained and evaluated five different groups of agents on the accumulating towers task. Two groups were based on evidence accumulation: one group included the full model with Laplace and inverse Laplace transform ($\tilde{\mathbf{f}}$), and the other included only the Laplace transform (**F**). We also compared three groups of agents based on existing recurrent architectures: a GRU, LSTM, and simple RNN.

Agents based on the cognitive model learned the task as fast as the GRU-based agents despite having much fewer parameters – the evidence accumulation module had 0 trainable parameters while GRU had 1,294,080 trainable parameters (Fig. 3).

We visualized the neural activity of one representative agent for $\tilde{\mathbf{f}}$ and GRU architectures after 10M environment steps of training (Fig. 4). The neurons in the $\tilde{\mathbf{f}}$ agents activated sequentially as a function of evidence, resembling the activity in neural recordings from the hippocampus (Nieh et al., 2021). This indicates that the network successfully learned the temporal derivative of the task-relevant abstract variable. This demonstrates that incorporating the cognitive model into deep RL agents enabled the agents to learn a more efficient neural representation that also resembles the neural representation of accumulated evidence in the hippocampus.

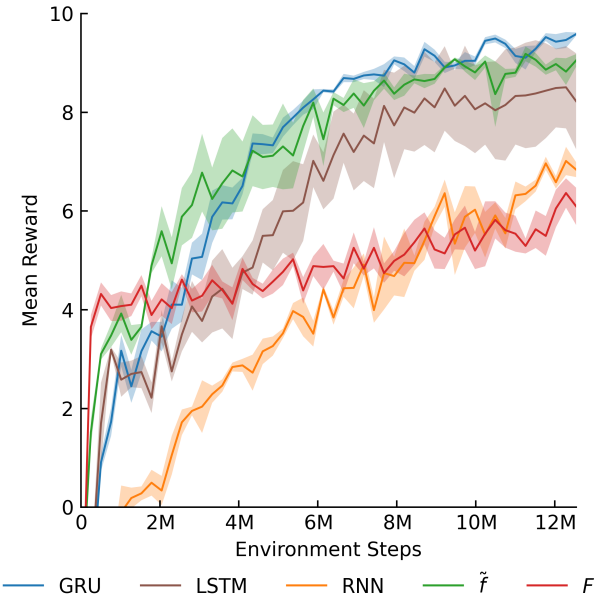

Figure 3: Agent performance on the accumulating towers task. Mean and standard deviation across the four agents.

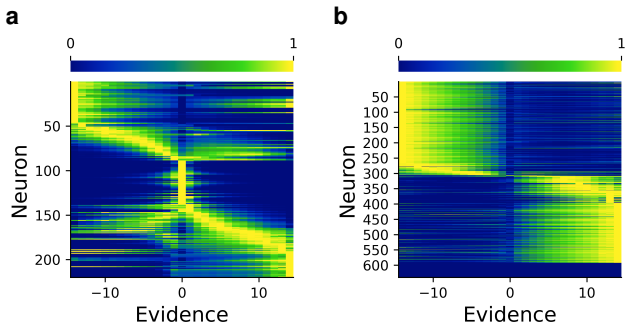

Figure 4: Activity of neurons in the recurrent layer. Neurons are sorted by peak activity. Agent architecture: (a) $\tilde{\mathbf{f}}$, (b) GRU. Neural activity of **F**, LSTM and RNN agents resembled activity of GRU agents. $\tilde{\mathbf{f}}$ resembles results reported in Nieh et al. (2021)

Acknowledgement

We gratefully acknowledge support from the National Institutes of Health's National Institute on Aging, grant 5R01AG076198-02.

References

Bures, J., Fenton, A., Kaminsky, Y., & Zinyuk, L. (1997). Place cells and place navigation. *Proceedings of the National Academy of Sciences*, *94*(1), 343–350.

Eichenbaum, H. (2014). Time cells in the hippocampus: a new dimension for mapping memories. *Nature Reviews Neuroscience*, *15*(11), 732–744.

Howard, M. W., Luzardo, A., & Tiganj, Z. (2018). Evidence accumulation in a laplace domain decision space. *Computational brain & behavior*, *1*(3), 237–251.

Howard, M. W., MacDonald, C. J., Tiganj, Z., Shankar, K. H., Du, Q., Hasselmo, M. E., & Eichenbaum, H. (2014). A unified mathematical framework for coding time, space, and sequences in the hippocampal region. *Journal of Neuroscience*, *34*(13), 4692–4707.

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

Mnih, V., Badia, A. P., Mirza, M., Graves, A., Lillicrap, T., Harley, T., ... Kavukcuoglu, K. (2016). Asynchronous methods for deep reinforcement learning. In *International conference on machine learning* (pp. 1928–1937).

Mochizuki-Freeman, J., Kabir, M. R., Gulecha, M., & Tiganj, Z. (2023). Geometry of abstract learned knowledge in deep rl agents. In *Neurips 2023 workshop on symmetry and geometry in neural representations*.

Mochizuki-Freeman, J., Maini, S. S., & Tiganj, Z. (2023). Characterizing neural activity in cognitively inspired rl agents during an evidence accumulation task. In *2023 international joint conference on neural networks (ijcnn)* (pp. 01–09).

Nieh, E. H., Schottdorf, M., Freeman, N. W., Low, R. J., Lewallen, S., Koay, S. A., ... Tank, D. W. (2021). Geometry of abstract learned knowledge in the hippocampus. *Nature*, 1–5.

Pastalkova, E., Itskov, V., Amarasingham, A., & Buzsaki, G. (2008). Internally generated cell assembly sequences in the rat hippocampus. *Science*, *321*(5894), 1322–1327.

Salz, D. M., Tiganj, Z., Khasnabish, S., Kohley, A., Sheehan, D., Howard, M. W., & Eichenbaum, H. (2016). Time cells in hippocampal area ca3. *Journal of Neuroscience*, *36*(28), 7476–7484.

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