Behavioral timescale synaptic plasticity, replay, and emergent behavioral choice

Suhee Cho (suheecho@stanford.edu)
Department of Psychology,
Stanford University,
Stanford, 94305, USA

James L. McClelland (jlmcc@stanford.edu)
Department of Psychology,
Stanford University,
Stanford, 94305, USA

Abstract

Animals can adapt to novel environments with minimal exploration and rapidly adjust to change. While the hippocampus is thought to encode spatial and other behaviorally relevant information to support this, how it does so with such efficiency remains unclear. Here, we show that behavioral time-scale synaptic plasticity (BTSP) coupled with replay can allow the hippocampus to learn a predictive map—known as the successor representation (SR)—after minimal exploration. Reward-induced BTSP events bias this map to over-represent reward locations, biasing behavior based on this representation toward reward, and replay events extend this bias to locations farther from the reward. Notably, the representation dynamically adjusts when reward locations shift, supporting rapid behavioral adaptation. Together, our findings offer a biologically plausible account of how BTSP and replay jointly enable quick adaptation of the SR, supporting efficient and adaptive learning.

Keywords: Reinforcement learning, Decision making, Hippocampus, Replay, Successor representation

Animals can be remarkably fast and efficient reinforcement learners. In maze tasks (Fig. 1a), rodents rapidly learn both the structure of the environment and the location of the reward, often within just a few trials (Dong, Madar, & Sheffield, 2021; Sosa & Giocomo, 2021). When the reward is moved to a new location, they quickly begin navigating toward the new reward location (Ambrose, Pfeiffer, & Foster, 2016; Carey, Tanaka, & van der Meer, 2019; Gauthier & Tank, 2018).

The hippocampus is thought to be a key brain region supporting efficient learning. A hallmark observation is the rapid, often one-shot, formation of place cells—neurons that fire at specific locations in an environment (O'Keefe, 1976). A synaptic learning rule known as behavioral time-scale plasticity (BTSP) has recently been proposed to underlie this rapid place field formation (Fig. 1b) (Bittner, Milstein, Grienberger, Romani, & Magee, 2017). BTSP operates over seconds-long timescales and induces large, one-shot synaptic changes. In CA1, BTSP is modulated by reward and exhibits temporal asymmetry (Fig. 1b), suggesting a role in reward prediction (Li, Briguglio, Romani, & Magee, 2024). Beyond spatial coding during exploration, the hippocampus generates internally structured sequences of place cell activity during rest or immobility (Gupta, van der Meer, Touretzky, & Redish, 2010). These replays are believed to support offline learning by consolidating experiences even in the absence of physical movement (Dong et al., 2021; Frank, Stanley, & Brown, 2004).

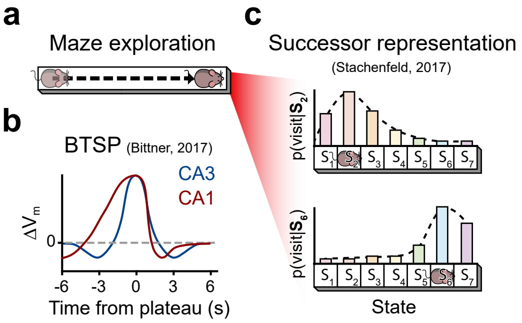

Figure 1: a, Simple linear maze task. A rat explores the maze from left to right. b, BTSP-induced synaptic updates at a CA3 (blue) and CA1 (red) synapse. The y-axis indicates the change in membrane potential of a synapse onto an initially silent cell, before and after BTSP. c, Predicted probability of visit under the SR framework during linear maze exploration. The animal's direction on movement reflects the SR, based on its travel history.

Here, we link these ideas to the proposal that place cells encode a successor representation (SR)–the expected future occupancy of states given the current one (Fig. 1c) (Stachenfeld, Botvinick, & Gershman, 2017). The SR allows the structure of state transitions to be learned independently of rewards, and can later integrate reward signals to support fast, flexible value-based behavior. We show that even without explicitly encoding value, BTSP in CA1 can lead to a reward-biased SR which, when combined with replay, drives rapid shifts in behavior toward previously rewarded locations.

Methods

We modeled a hippocampus with two layers, CA3 and CA1, each with 2,100 neurons (Fig. 2a), drawing on biophysically informed models (Ecker et al., 2022). CA3 included recurrent connections and projected feedforward to CA1. Neurons could produce action potentials (APs) and dendritic

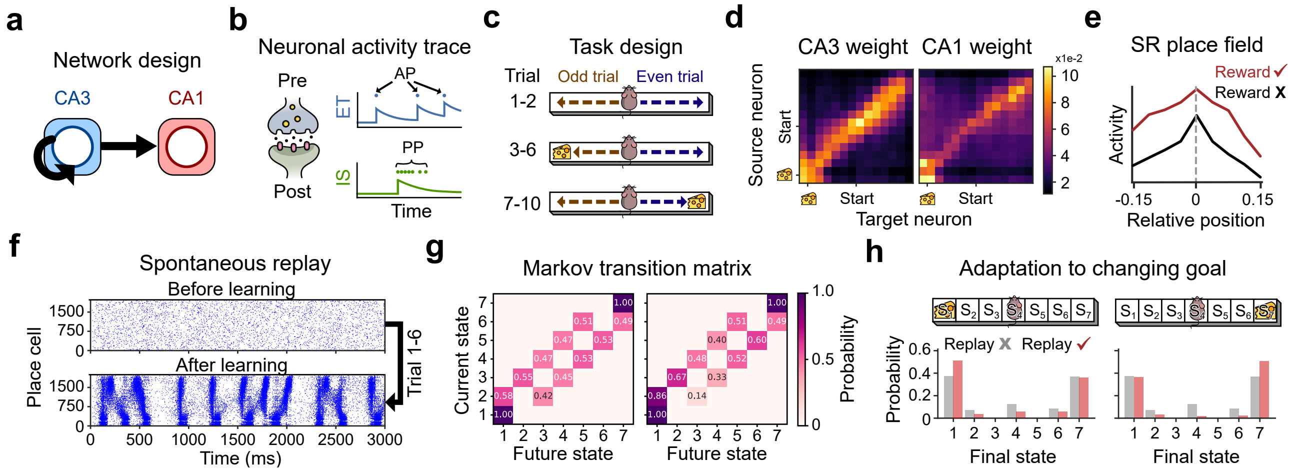

Figure 2: a, Schematic of the hippocampal network b, Schematic of pre- and post-synaptic neuronal activity trace. c, Schematic of linear maze task. d, CA3 and CA1 weight matrices after 6 trials. e, Skewed place fields of representative reward-sensitive and non-reward place cells. f, Spontaneous replay events during 3 seconds of offline activity. g, Markov transition matrices before (left) and after (right) replay. h, 10-step state transition probabilities before (left) and after reversal (right).

plateau potentials (PPs); APs were driven by presynaptic input, while PPs were triggered stochastically by depolarization, and, in CA1, the presence of reward (Milstein et al., 2021). APs created time-decaying eligibility traces (ETs) on outgoing synapses, and PPs produced time-decaying 'instructive signals' (ISs) (Fig. 2b). When a PP occurred, incoming synapses were updated based on the time-integrated product of each incoming synapse's ET and the neuron's IS (Milstein et al., 2021). This causes the weights between pre-and post-synaptic neurons with different temporal offsets to be structured as in Fig. 1b. The longer decay constant for the IS in CA1 resulted in temporally asymmetric BTSP kernel (Fig. 1b). Both recurrent and feedforward projections were randomly initialized with small weights. While 80% of CA3 neurons were designated as place cells and received hard-wired spatial input, CA1 neurons received input only from CA3 via random weights and were expected to develop place-specific responses through maze traversals. We used a linear maze and simulated traversals from the center to each end over a series of trials (Fig. 2c). The first two trials (one in each direction) served as a familiarization phase without reward. Reward was positioned at the left end for the next four trials, then switched to the right end for the last four trials. On rewarded trials, the simulated agent stayed at the reward position briefly, consistent with animals' behavior in rewarded arms of a maze (Gillespie et al., 2021). To examine replay during rest, we introduced low-intensity, random input to CA3 for 3 seconds after trials 6 and 10. During this phase, we used adaptive integrate-and-fire models which support replay (Ecker et al., 2022), and continued updating weights using the BTSP rule.

Results

We found that BTSP enables an SR-like cognitive map, connecting place cells with neighboring place fields. After trials 1-6, we sorted cells into 14 spatial bins based on their acquired place-field locations, averaging synaptic weights to and from cells in each group (Fig. 2d). The CA3 network developed a robust SR-like map, with generally symmetric weights connecting nearby locations, over-representing the reward site from the longer time spent there. In CA1, the asymmetric BTSP kernel led to a directional bias: synapses from start position cells toward terminal position cells were stronger than the reverse, especially near the reward site, which was very strongly represented (Fig. 2e). We next examined replay during the offline phase (Fig. 2f). After trial 6, the network generated SWRs and sequential firing patterns traversing the maze. Neither feature was exhibited prior to learning.

Next, we examined whether the SR map biases the agent's behavior toward reward, and how replay affects this. We discretized the maze into 7 states, S_1 (left) to S_7 (right) (Fig. 2h, top). We simulated the agent starting at S_4 (center), moving until reaching either end, with state transitions based on the center positions of the CA1 place fields of cells active when the agent was positioned at the center of each discrete state (Fig. 2g). Using the learned network after two reward exposures but prior to the offline phase, the agent preferred the reward state S_1 when near it (at S_2), but this preference weakened farther away (from S_3) (Fig. 2g, left). After the offline phase, the updated transition matrix revealed a stronger tendency to move toward the reward, even from earlier positions (Fig. 2g, right). Replay significantly increased the likelihood of reaching the reward site within 10 steps (Fig. 2h, left), measured by taking the transition matrix to the 10th power.

Finally, we tested rapid adaptation to the new reward position. After two traversals in both directions, with the support of replay, the agent did exhibit rapid adaptation, showing a higher probability of reaching the new reward site at S_7 (Fig. 2h, right).

Conclusion

Combining BTSP and replay provides a biologically plausible mechanism for rapidly forming and adapting cognitive maps in the hippocampus. BTSP enables both rapid formation of place fields and replay, rapidly forming an SR that adapts to shifting rewards and enhances prospective decision-making. We see these findings as a step toward a unified framework connecting cellular plasticity, network dynamics, and reinforcement learning, advancing our understanding of how the hippocampus supports fast, flexible learning.

Acknowledgments

The first author is supported by the Emergent Cognitive Functions Fund, Department of Psychology, Stanford University.

References

Ambrose, R. E., Pfeiffer, B. E., & Foster, D. J. (2016). Reverse replay of hippocampal place cells is uniquely modulated by changing reward [Journal Article]. *Neuron*, *91*(5), 1124-1136. Retrieved from `https://doi.org/10.1016/j.neuron.2016.07.047` (doi: 10.1016/j.neuron.2016.07.047) doi: 10.1016/j.neuron.2016.07.047

Bittner, K. C., Milstein, A. D., Grienberger, C., Romani, S., & Magee, J. C. (2017). Behavioral time scale synaptic plasticity underlies ca1 place fields [Journal Article]. *Science*, *357*(6355), 1033-1036. Retrieved from `https://doi.org/10.1126/science.aan3846` (doi: 10.1126/science.aan3846) doi: 10.1126/science.aan3846

Carey, A. A., Tanaka, Y., & van der Meer, M. A. A. (2019). Reward revaluation biases hippocampal replay content away from the preferred outcome [Journal Article]. *Nature Neuroscience*, *22*(9), 1450-1459. Retrieved from `https://doi.org/10.1038/s41593-019-0464-6` doi: 10.1038/s41593-019-0464-6

Dong, C., Madar, A. D., & Sheffield, M. E. J. (2021). Distinct place cell dynamics in ca1 and ca3 encode experience in new environments [Journal Article]. *Nature Communications*, *12*(1), 2977. Retrieved from `https://doi.org/10.1038/s41467-021-23260-3` doi: 10.1038/s41467-021-23260-3

Ecker, A., Bagi, B., Vértes, E., Steinbach-Németh, O., Karlócai, M. R., Papp, O. I., ... Káli, S. (2022, jan). Hippocampal sharp wave-ripples and the associated sequence replay emerge from structured synaptic interactions in a network model of area ca3. *eLife*, *11*, e71850. Retrieved from `https://doi.org/10.7554/eLife.71850` doi: 10.7554/eLife.71850

Frank, L. M., Stanley, G. B., & Brown, E. N. (2004). Hippocampal plasticity across multiple days of exposure to novel environments [Journal Article]. *The Journal of Neuroscience*, *24*(35), 7681. Retrieved from `http://www.jneurosci.org/content/24/35/7681.abstract` doi: 10.1523/JNEUROSCI.1958-04.2004

Gauthier, J. L., & Tank, D. W. (2018). A dedicated population for reward coding in the hippocampus [Journal Article]. *Neuron*, *99*(1), 179-193.e7. Retrieved from `https://www.sciencedirect.com/science/article/pii/S089662` doi: https://doi.org/10.1016/j.neuron.2018.06.008

Gillespie, A. K., Astudillo Maya, D. A., Denovellis, E. L., Liu, D. F., Kastner, D. B., Coulter, M. E., ... Frank, L. M. (2021). Hippocampal replay reflects specific past experiences rather than a plan for subsequent choice [Journal Article]. *Neuron*, *109*(19), 3149-3163.e6. Retrieved from `https://www.sciencedirect.com/science/article/pii/S089662` doi: https://doi.org/10.1016/j.neuron.2021.07.029

Gupta, A. S., van der Meer, M. A. A., Touretzky, D. S., & Redish, A. D. (2010). Hippocampal replay is not a simple function of experience [Journal Article]. *Neuron*, *65*(5), 695-705. Retrieved from https://www.sciencedirect.com/science/article/pii/S0896627310000607 doi: https://doi.org/10.1016/j.neuron.2010.01.034

Li, Y., Briguglio, J. J., Romani, S., & Magee, J. C. (2024). Mechanisms of memory-supporting neuronal dynamics in hippocampal area ca3 [Journal Article]. *Cell*, *187*(24), 6804-6819.e21. Retrieved from https://doi.org/10.1016/j.cell.2024.09.041 (doi: 10.1016/j.cell.2024.09.041) doi: 10.1016/j.cell.2024.09.041

Milstein, A. D., Li, Y., Bittner, K. C., Grienberger, C., Soltesz, I., Magee, J. C., & Romani, S. (2021). Bidirectional synaptic plasticity rapidly modifies hippocampal representations [Journal Article]. *eLife*, *10*, e73046. Retrieved from https://doi.org/10.7554/eLife.73046 doi: 10.7554/eLife.73046

O'Keefe, J. (1976). Place units in the hippocampus of the freely moving rat [Journal Article]. *Experimental Neurology*, *51*(1), 78-109. Retrieved from https://www.sciencedirect.com/science/article/pii/0014488676900558 doi: https://doi.org/10.1016/0014-4886(76)90055-8

Sosa, M., & Giocomo, L. M. (2021). Navigating for reward [Journal Article]. *Nature Reviews Neuroscience*, *22*(8), 472-487. Retrieved from https://doi.org/10.1038/s41583-021-00479-z doi: 10.1038/s41583-021-00479-z

Stachenfeld, K. L., Botvinick, M. M., & Gershman, S. J. (2017). The hippocampus as a predictive map [Journal Article]. *Nature Neuroscience*, *20*(11), 1643-1653. Retrieved from https://doi.org/10.1038/nn.4650 doi: 10.1038/nn.4650