Adaptive learning using attractor switches in recurrent neural networks

Qin He (qin_he@brown.edu)
Brown University Dept. of Neuroscience
Providence, RI 02906 USA

Daniel N. Scott (daniel_scott@brown.edu)
Brown University Dept. of Neuroscience
Providence, RI 02906 USA

Matthew R. Nassar * **(matthew_nassar@brown.edu)**
Brown University Dept. of Neuroscience
Providence, RI 02906 USA

Cristian B. Calderon * **(cristian.buc@cenia.cl)**
Centro Nacional de Inteligencia Artificial (CENIA)
Macul, Chile

*Co-senior authors

Abstract

Behaving adaptively requires determining when to create new state-action associations, and when to modify existing ones. Normative probabilistic models can accomplish this, but are computationally demanding and require strong assumptions. Which approximations to normative models brains use to avoid these difficulties remains unclear, however. Drawing inspiration from work showing thalamo-cortico-basal-ganglia loop involvement in adaptive learning, we develop and characterize a neural network model that builds new state-action associations via Hebbian learning between cortex and striatum when surprise (computed as the entropy of neural responses) is elevated. We test our model on a predictive inference task including change-points, and show that it captures statistics of normative models, human behavior, and individual differences. The mechanisms in our model may therefore support state-action representation dynamics in-vivo, and differences in them may account for individual differences in adaptive behavior.

Keywords: adaptive learning; recurrent neural network; attractor network; basal ganglia

Previous research put forward normative models based on Bayesian probability theory to explain how people adjust the impact of new information on their beliefs, a phenomenon known as adaptive learning. However, these models rely on knowing the generative process producing both observed data (Fearnhead & Liu, 2007) and any abrupt changes in latent variables (Steyvers & Brown, 2005). Full Bayesian inference models are computationally costly (Wilson et al., 2010), but their performance can be approximated via error-driven learning using dynamic learning rates (Nassar et al., 2010, 2012). Unfortunately, approximate models still set learning rates based on change-point probabilities and latent-state estimation uncertainty, which require generative process knowledge. This suggests that learning rate signals are context-agnostic, in contrast with previous work indicating the opposite (D'Acremont & Bossaerts, 2016; Nassar et al., 2019).

To account for context-dependent learning, recent work suggests that adaptive learning can be re-framed as a trade-off between updating a current belief and establishing a new one, making it a form of structure learning (Linda et al., 2021). The authors suggest that orthogonalizing state-action mappings by recruiting distinct neuronal populations for novel states would avoid interference between learning episodes. Inspired by such a proposal, previous work (Razmi & Nassar, 2022) used a two-layer feed-forward network to build state-action pairs with distinct neurons to model human adaptive learning. However, previous work has not tackled *how* the brain may recruit novel neuronal population to build new state-action pairs, and what would be the emergent behavioral properties of such a mechanism.

However, empirical observations may provide a framework for answering the question above. First, task-relevant representations are often coded in neural population attractors (Ebitz & Hayden, 2021), which can emerge in recurrent neural networks (Brunel, 2003; Hopfield, 1982; Litwin-Kumar & Doiron, 2014; Maes et al., 2020; Recanatesi et al., 2022). Second, extensive work shows context information in hierarchically structured tasks is available within cortico-basal ganglia loops, and that mnemonic information in working memory is updated using corticostriatal circuits (Chatham et al., 2014). Finally, thalamocortical projections enable switches between neural representation in the cortex, as shown by experimental studies (Remington et al., 2018; Wang et al., 2018) and supported by theoretical work (Calderon et al., 2022; Recanatesi et al., 2022).

Results

Inspired by the cortico-basal-ganglia loop and incorporating the different ideas discussed above, we propose LEIA (**L**earning as **E**ntropy-**I**nduced **A**ttractor state switches), a biologically inspired model that can learn to represent new state-action pairs, without needing knowledge of specific generative models. In everyday experience, adaptive learning typically unfolds sequentially: we first produce actions then observe outcomes which we learn from. Iterating this process we adapt our learning based on the statistics of our environments. Our model follows the same logic by first producing an action and then observing the environmental response, potentially producing a prediction error to be learned from. Crucially, the magnitude of the prediction error is interpreted in terms of observation statistics to determine a variable learning rate.

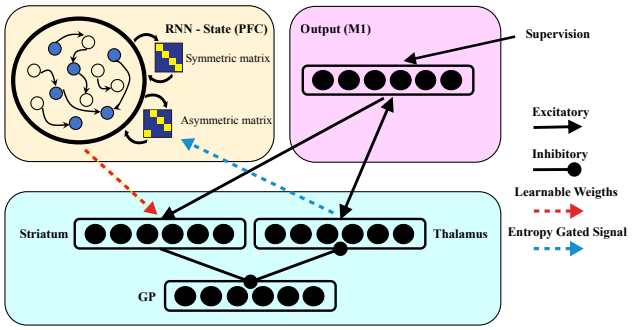

Figure 1: LEIA structure. Our model is composed of three modules: The RNN-state module (yellow area) represents the context as a fixed point attractor state. The Action Selection module (cyan area). The motor output module representing action execution (purple area).

To produce actions, a latent state representation in cortex, modeled as an RNN, is read by a basal ganglia network. The recurrent network encodes latent state representations as attractor states (Fig. 1). We use a symmetrical matrix with positive diagonal elements encoding orthogonal states and negative off-diagonal elements for inhibition to implement these attractors. The RNN projects to the basal ganglia, and in particular to "Go" cells in a striatal layer, which represent the action currently associated with the RNN's hidden state. This action

representation is propagated via the BG through the connections from striatal "Go" cells to the globus pallidus (GP), and then from GP to the motor thalamus, in line with known neurobiology. The motor thalamus then projects to the motor cortex, which executes the action.

Observation-based feedback (supervision) arrives at the motor cortex, which projects back to the striatum and thalamus. When the supervised action is different from the predicted one, the supervised representation in motor thalamus is far from the proposed one, and as a result, surprise, computed as the entropy of the layer's firing rate is also high. Surprise is then used to induce hidden state transitions in the RNN, implemented using an asymmetric recurrent weight matrix. Thus, high entropy induces an attractor state transition in the RNN, nonlinearly with respect to prediction error, therefore implementing an adaptive learning rate. The neurons in a given attractor state become coupled via Hebbian learning in cortico-basal-ganglia projections to the currently supervised actions, reinforcing existing state-action pairs when prediction errors are low or creating new ones when prediction errors are high.

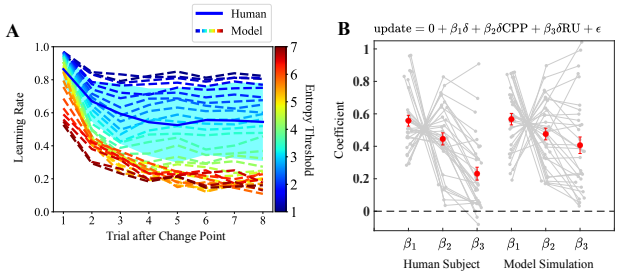

Figure 2: A. Model and human learning rate dynamics. The solid blue line represents the average human learning rate curve after change-points. Dashed lines represent model learning rates using different entropy thresholds. Light blue shading shows participant variability (SDs). B. Coefficients from our regression model that capture prediction error driven (β_1), change-point probability driven (β_2), and relative uncertainty driven learning (β_3).

We tested our model on a predictive inference task and compared it with data collected from 32 participants doing the same task (McGuire et al., 2014). The task required predicting upcoming data points given previous observations. All data points were produced by a generative process with statistics that changed abruptly over time. As in previous work, we computed learning rates for participant and model predictions as functions of post-change-point trial numbers (Razmi & Nassar, 2022). We observed that participants' average learning rates decreased after change-points (Fig 2. solid blue line). We applied our model using different entropy thresholds and performed the same adaptive learning rate analysis (Fig. 2). The results show that LEIA can capture not only trends in adaptive learning rates, but also individual differences. To better characterize both model and human behavior, we ran a regression model predicting learning rates (equation above Fig 2.) (McGuire et al., 2014). The model includes parameters quantifying prediction error (PE, β_1), change-point probability (CPP, β_2) and relative uncertainty (RU, β_3) driven learning. Fig. 2B shows that LEIA captures the rank order of human coefficients and their approximate values.

One shortcoming of our model, as described thus far, is that it is forced to create new states whenever large prediction errors are observed. However, in many settings, these may reflect returns to previous states. To address this, we extended the model by introducing a learnable projection from motor cortex to the RNN gated by surprise. When large prediction errors cause surprise, but the supervised action suggests the latent state is old (trial 20) then the model will switch into previously seen attractors instead of learning anew. The impact on behavior is illustrated in fig. 3A.

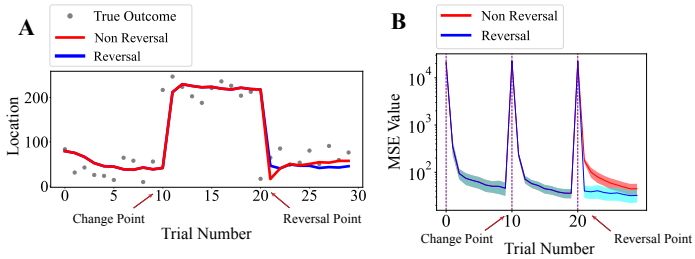

Figure 3: A. An illustrative example of the reversal learning model. The red line and blue line show predictions of the basic model and reversal learning model, respectively. Both models update their predictions after a change-point. However, the reversal model can jump back to previous states and use learned associations after a reversal-point. B. Quantitative analysis. The X-axis shows trial numbers, and the Y-axis shows the average mean square error (MSE) between predictions and the generative means for 5 processes with different noise levels. The shaded area shows the SEM across generative processes.

To summarize, our model shows how simple, established biological mechanisms (attractor networks, Hebbian learning, and surprise) can reproduce human latent-state update dynamics. Specifically, it shows that splitting representations when events are surprising, and incrementally updating them when events are not, can account for subjects' behaviors. These mechanisms can be tested using neural data from regions such as OFC, which are involved in state representation. They also predict that state representations could be continuously modified, such that neurons representing one state could gradually come to represent a completely distinct one if observations changed slowly. Likewise, our results predict that heightened surprise could pathologically fractionate individual states, leading to poor generalization over observations generated by individual latent states.

Acknowledgments

This work was funded by the National Center for Artificial Intelligence CENIA FB210017, Basal ANID.

References

Brunel, N. (2003). Dynamics and plasticity of stimulus-selective persistent activity in cortical network models. *Cerebral Cortex*, *13*(11), 1151–1161.

Calderon, C. B., Verguts, T., & Frank, M. J. (2022). Thunderstruck: The acdc model of flexible sequences and rhythms in recurrent neural circuits. *PLoS computational biology*, *18*(2), e1009854.

Chatham, C. H., Frank, M. J., & Badre, D. (2014). Corticostriatal output gating during selection from working memory. *Neuron*, *81*(4), 930–942.

D'Acremont, M., & Bossaerts, P. (2016). Neural mechanisms behind identification of leptokurtic noise and adaptive behavioral response. *Cerebral Cortex*, *26*(4), 1818–1830.

Ebitz, R. B., & Hayden, B. Y. (2021). The population doctrine in cognitive neuroscience. *Neuron*, *109*(19), 3055–3068.

Fearnhead, P., & Liu, Z. (2007). On-line inference for multiple changepoint problems. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *69*(4), 589–605.

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the national academy of sciences*, *79*(8), 2554–2558.

Linda, Q. Y., Wilson, R. C., & Nassar, M. R. (2021). Adaptive learning is structure learning in time. *Neuroscience & Biobehavioral Reviews*, *128*, 270–281.

Litwin-Kumar, A., & Doiron, B. (2014). Formation and maintenance of neuronal assemblies through synaptic plasticity. *Nature communications*, *5*(1), 5319.

Maes, A., Barahona, M., & Clopath, C. (2020). Learning spatiotemporal signals using a recurrent spiking network that discretizes time. *PLoS computational biology*, *16*(1), e1007606.

McGuire, J. T., Nassar, M. R., Gold, J. I., & Kable, J. W. (2014). Functionally dissociable influences on learning rate in a dynamic environment. *Neuron*, *84*(4), 870–881.

Nassar, M. R., Bruckner, R., & Frank, M. J. (2019). Statistical context dictates the relationship between feedback-related eeg signals and learning. *elife*, *8*, e46975.

Nassar, M. R., Rumsey, K. M., Wilson, R. C., Parikh, K., Heasly, B., & Gold, J. I. (2012). Rational regulation of learning dynamics by pupil-linked arousal systems. *Nature neuroscience*, *15*(7), 1040–1046.

Nassar, M. R., Wilson, R. C., Heasly, B., & Gold, J. I. (2010). An approximately bayesian delta-rule model explains the dynamics of belief updating in a changing environment. *Journal of Neuroscience*, *30*(37), 12366–12378.

Razmi, N., & Nassar, M. R. (2022). Adaptive learning through temporal dynamics of state representation. *Journal of Neuroscience*, *42*(12), 2524–2538.

Recanatesi, S., Pereira-Obilinovic, U., Murakami, M., Mainen, Z., & Mazzucato, L. (2022). Metastable attractors explain the variable timing of stable behavioral action sequences. *Neuron*, *110*(1), 139–153.

Remington, E. D., Egger, S. W., Narain, D., Wang, J., & Jazayeri, M. (2018). A dynamical systems perspective on flexible motor timing. *Trends in cognitive sciences*, *22*(10), 938–952.

Steyvers, M., & Brown, S. (2005). Prediction and change detection. *Advances in neural information processing systems*, *18*.

Wang, J., Narain, D., Hosseini, E. A., & Jazayeri, M. (2018). Flexible timing by temporal scaling of cortical responses. *Nature neuroscience*, *21*(1), 102–110.

Wilson, R. C., Nassar, M. R., & Gold, J. I. (2010). Bayesian online learning of the hazard rate in change-point problems. *Neural computation*, *22*(9), 2452–2476.