Discovering Cognitive Models in a Competitive Mixed-strategy Game

Peiyu Liu[1], Kevin J. Miller[2], Hyojung Seo[1]

Yale SCHOOL OF MEDICINE

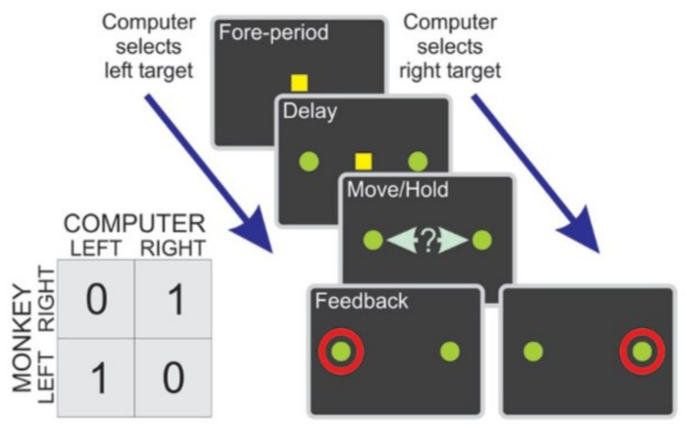
1. Department of Psychiatry, Yale University School of Medicine 2. Google Deepmind Neuroscience Lab; Sainsbury Wellcome Centre and Institute of Ophthalmology, University College London

Google DeepMind

Overview

- Complex behavioral tasks are crucial for cognitive neuroscience, but challenging to model computationally
- **Disentangled Recurrent neural networks (disRNN)** offer new opportunities for model discovery
- We apply DisRNN to data from monkeys playing the **Matching Pennies** game, which has been widely studied but for which no adequate computational model exists
- We **reveal interpretable cognitive models**, in which behavior is a mixture of long-term perseveration, medium-term reward-following, and a novel short-term choice decorrelation strategy

Matching pennies (MP) task

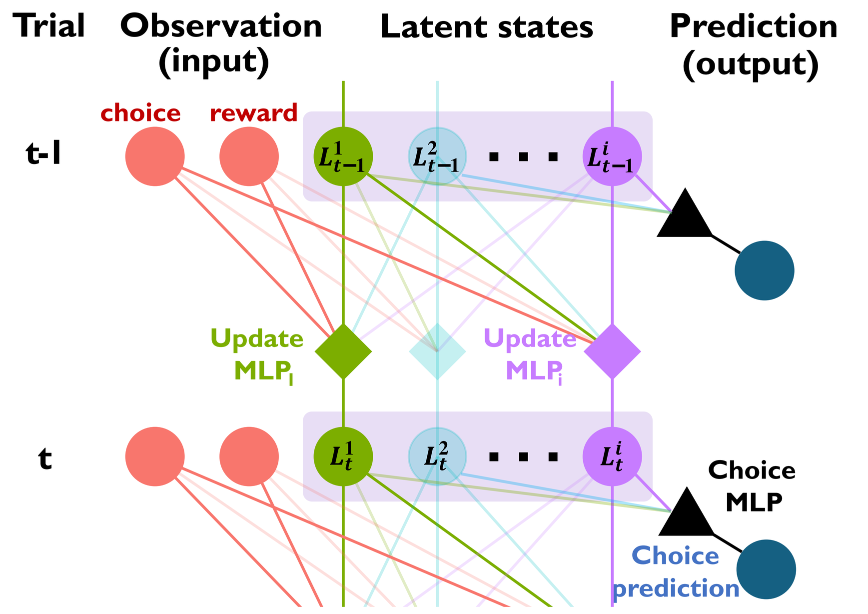


Matching Pennies is a competitive mixed-strategy game in which two players each attempt to predict the choice of the other. It can be thought of as a two-alternative variant of rock-paper-scissors.

In our dataset, monkeys play the matching pennies task against a computer opponent, which tracks and exploits patterns in monkey behavior. We consider data from three example monkeys (1: 79 sessions, 2: 111 sessions, 3: 47 sessions, each session is capped at 500 trials)

(Lee et al. 2004; Barraclough et al., 2004)

disRNN architecture



Information bottleneck

$$\tilde{x}_t \sim \mathcal{N}(mx_t, \sigma)$$

$$L_b = \sum_t D_{KL}(\mathcal{N}(m_b x_{b,t}, \sigma_b) || \mathcal{N}(0,1))$$

Latent update

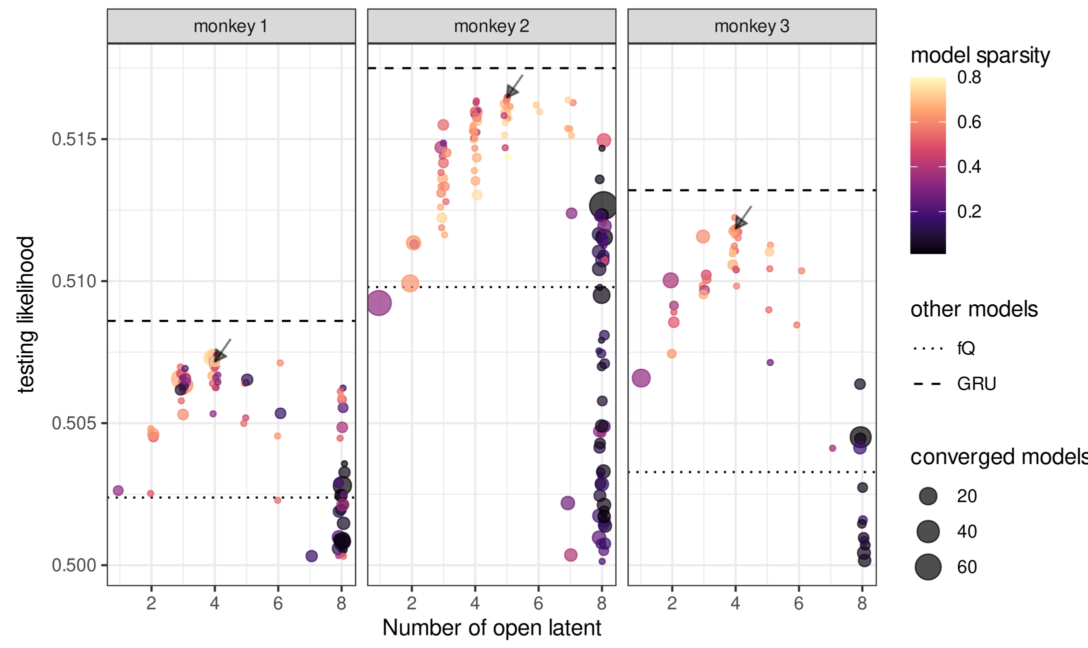
$$w_t^i, u_t^i = \mathrm{MLP}^i_{\mathrm{update}}(\mathrm{bottleneck}(\tilde{\mathbf{z}}_{t-1}, \mathbf{o}_t))$$

$$z_t^i = (1 - w_t^i)\tilde{z}_{t-1}^i + w_t^i u_t^i$$

disRNN is a recurrent network designed for learning sparse and interpretable representations to explain behavioral data. The model processes an agent's observations to predict their next choice, optimizing for prediction accuracy while penalizing model complexity using an "information bottleneck" at connections between latent units. The latent units update similarly to Q-learning, aiding in the interpretation of their functions.

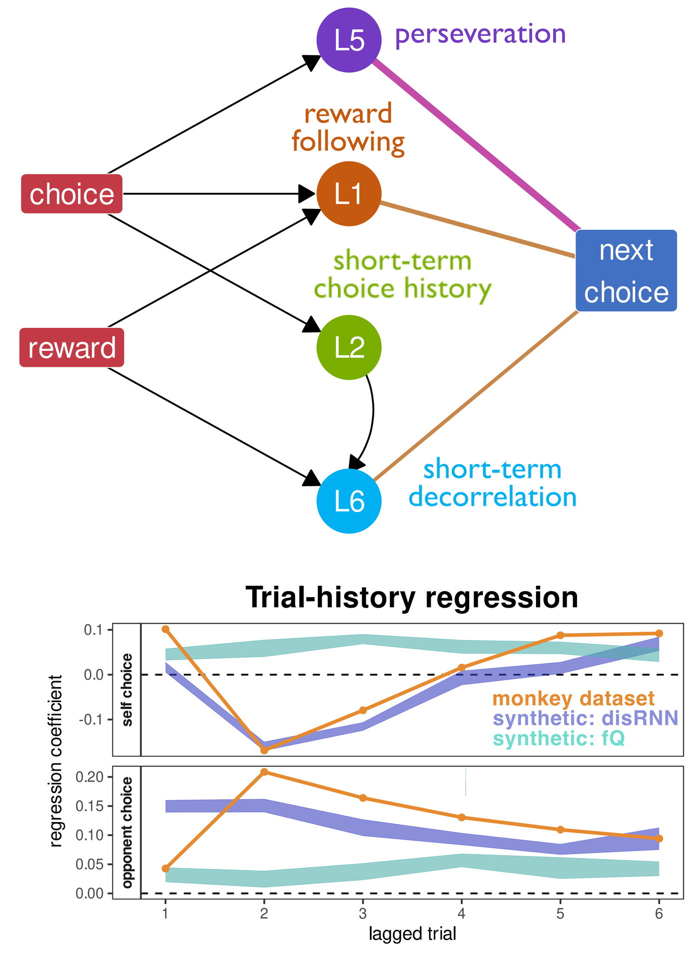
(Miller et al. 2023)

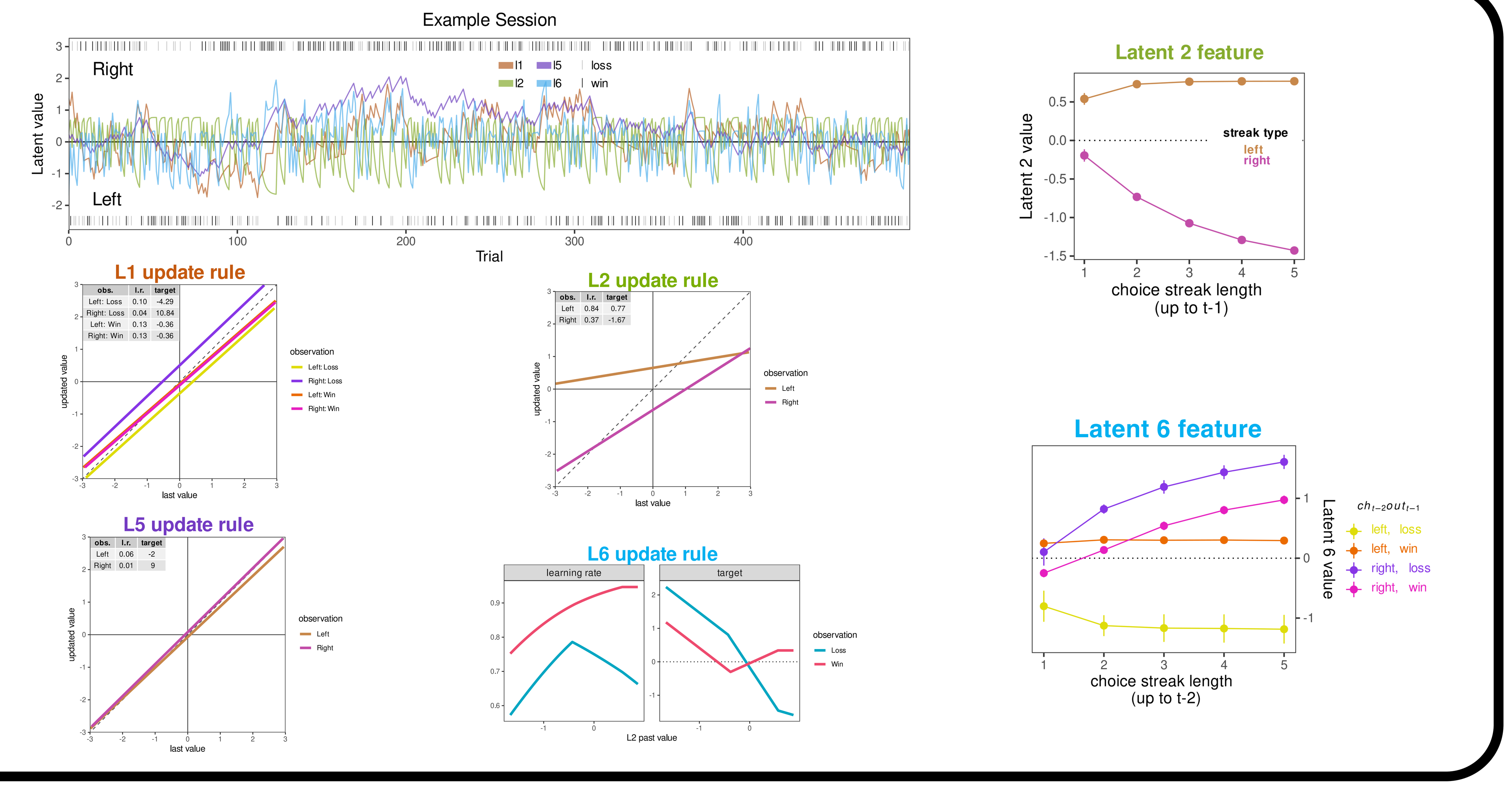
Identifying the best model



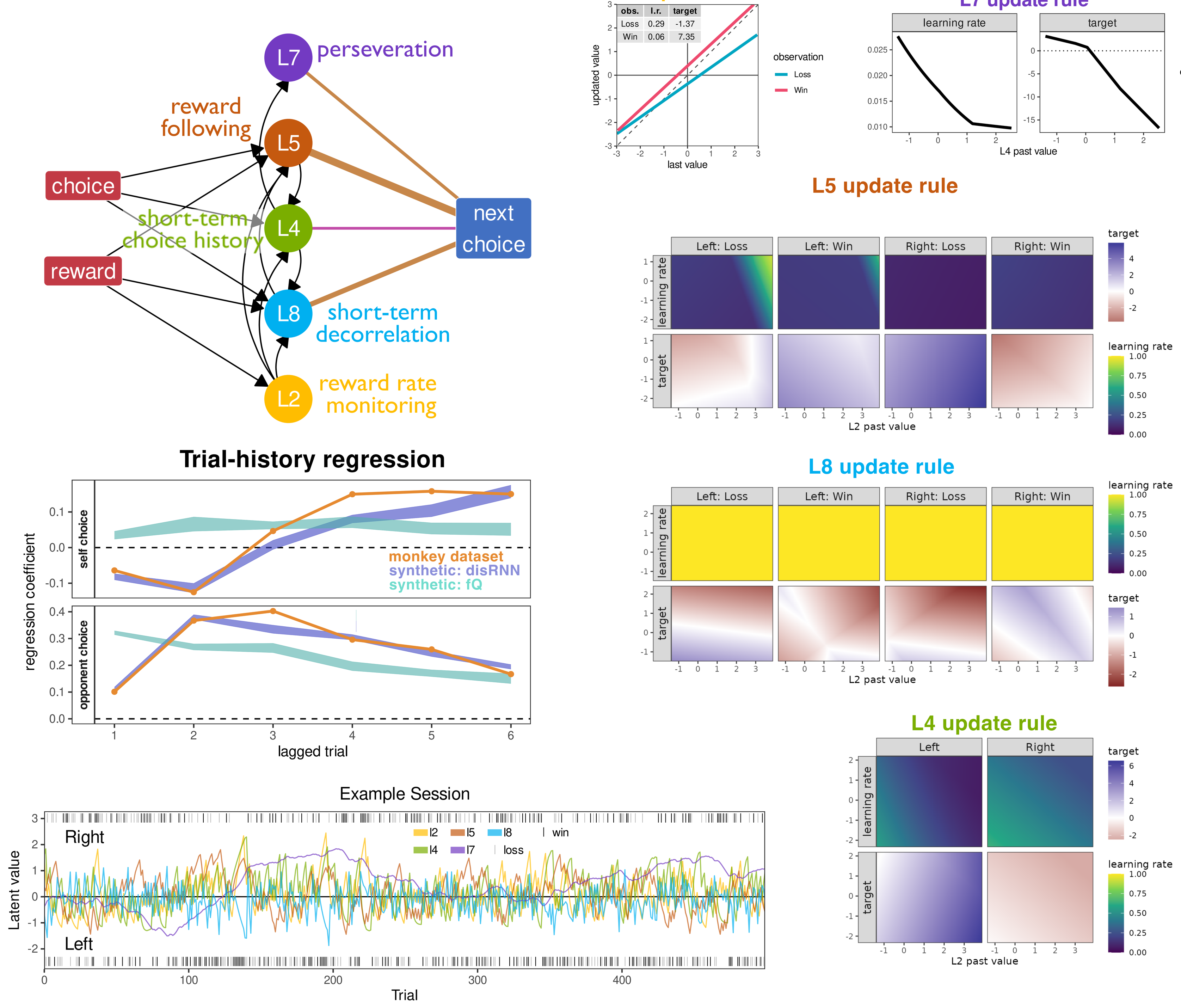
We have trained numerous disRNN models across a grid of combinations of information bottleneck penalty scales. The best disRNN model outperforms the top forgetting Q-learning (fQ) model and approaches the performance of the best Gated recurrent unit (GRU) network model. Then, we interpret the models that balance performance and simplicity for each monkey dataset. Arrows indicate models to be interpreted in the following sections.

Example model: monkey 1

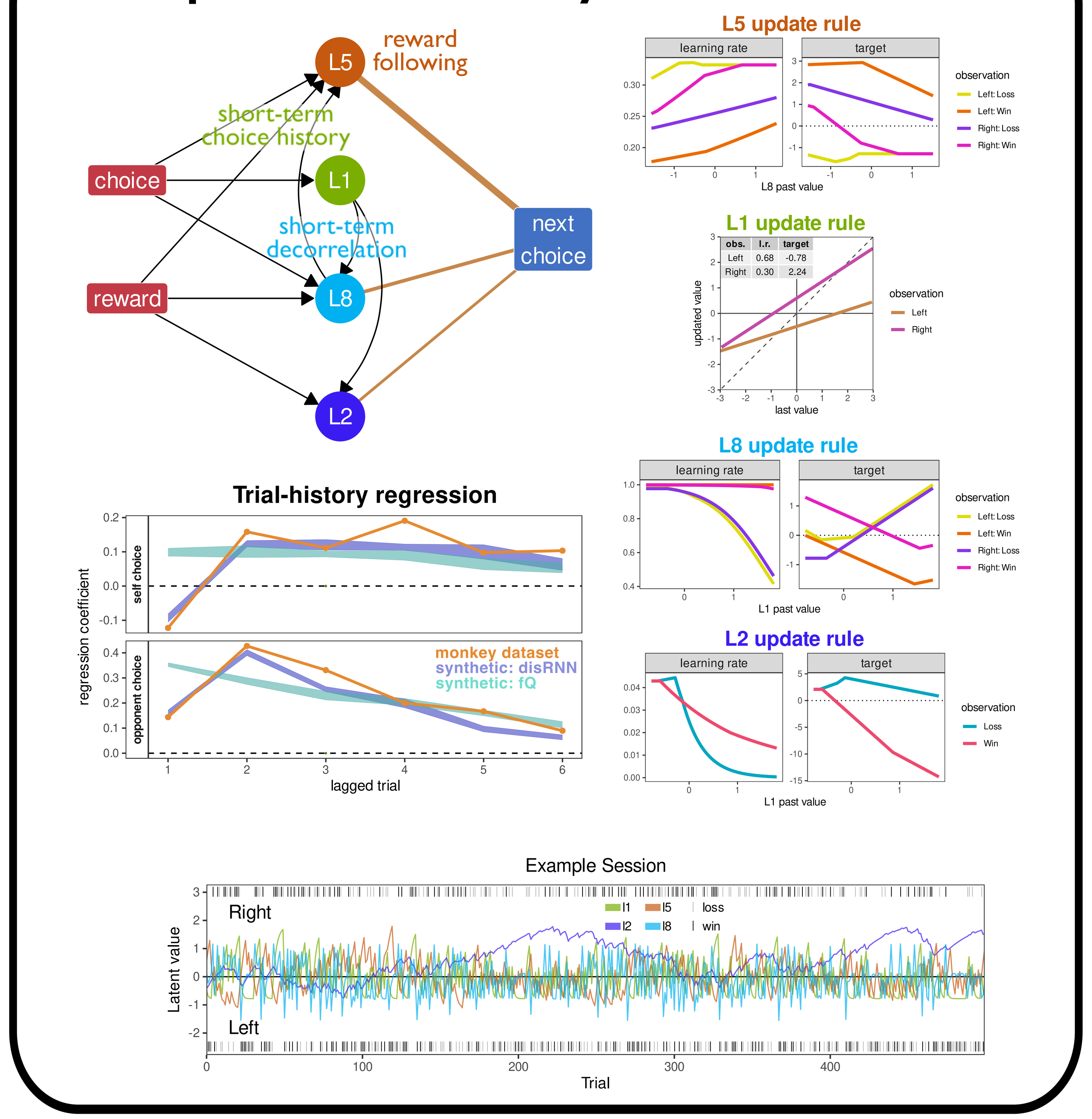




Example model: monkey 2



Example model: monkey 3



Conclusion

- Monkey behavior can be best described as a combination of several components, with some components commonly shared across monkeys, such as **choice perseveration, reward following, and short-term decorrelations**.
- Components in one model span a **wide range of learning rates**, suggesting monkeys integrate information over both short-term and long-term timescales.
- The **short-term decorrelation** component likely captures the critical and idiosyncratic strategy that monkeys use to counteract a computer opponent's exploitation.

Future directions

- Examining behavior from disRNN models with certain components turned off would allow us to more precisely determine the function of each component.
- disRNN models provide novel hypotheses about monkeys' cognitive models. It would be interesting to see if these hypotheses can be confirmed with neural correlate activity.