Characterizing attractor geometry in human decision making via low-dimensional RNNs

Hua-Dong Xiong (hdx@arizona.edu)[1]
Department of Psychology, University of Arizona
Tucson, AZ, 85721, USA

Li Ji-An (jil095@ucsd.edu)[1]
Neurosciences Graduate Program, University of California San Diego
La Jolla, CA, 92093, USA

Marcelo G. Mattar (marcelo.mattar@nyu.edu)[2]
Department of Psychology, New York University
New York, NY, 10003, USA

Robert C. Wilson (bob@arizona.edu)[2]
Department of Psychology and Cognitive Science Program, University of Arizona
Tucson, AZ, 85721, USA

[1] Co-first authors.

[2] Co-senior authors

Abstract

Recurrent neural networks (RNNs) have been widely utilized for modeling biological decision-making behaviors and uncovering underlying cognitive mechanisms. These networks demand less manual engineering and offer a more flexible framework compared to classical cognitive models such as reinforcement learning. However, previous studies have predominantly focused on simple decision making tasks, such as two-armed bandits with discrete rewards. Less is known about the ability of RNNs to uncover novel computational mechanisms in more complex settings, including tasks with multiple phases and continuous rewards. Here, we trained RNNs and classical cognitive models to predict choices of human subjects performing the Horizon task, which employs two phases to examine the human explore-exploit trade-off. Our RNNs substantially outperformed classical cognitive models. We then reverse-engineered these RNNs by distilling them into two-dimensional versions for each individual and analyzing the geometry of their attractors through dynamical systems analysis. We discovered that these RNNs identified a spectrum of correlated value-update rules and reward utilities forms. Our approach reveals diverse strategies employed by individuals that traditional cognitive modeling often overlooks, thereby advancing our understanding of the complex explore-exploit behavioral dynamics inherent in human.

Keywords: recurrent neural network; dynamical system analysis; decision-making; computational modeling

Introduction

Recurrent neural networks (RNNs) have emerged as a powerful paradigm for modeling sequential data, such as neural dynamics and behavioral sequences. Although these models can yield excellent predictive performance, their lack of interpretability often limits deeper insights. Recently, a series of works have demonstrated that a novel RNN modeling framework, which restricts the numbers of dynamical variables in hidden states, can reveal novel cognitive mechanisms from choice behavior of subjects performing decision-making tasks (Ji-An, Benna, & Mattar, 2023; Miller, Eckstein, Botvinick, & Kurth-Nelson, 2023). However, it remains unclear how it will adapt to more complex tasks than those studied in existing works. Here, we apply this framework to a behavioral dataset of human subjects performing the Horizon task, designed to examine human explore-exploit behavior across multiple task phases (Wilson, Geana, White, Ludvig, & Cohen, 2014). By analyzing the attractor geometry of these RNN models, we discovered diverse strategies that are often missed by classical cognitive models.

Results

Horizon Task

In this task, 641 subjects participate in a sequence of games, completing around 600,000 trials collectively. Each game (Fig. 1a) involves choosing between two slot machines with rewards from a Gaussian distribution, requiring exploration to identify the optimal machine. Initially, the first four ”instructed trials” allow only passive observation of rewards. Subsequently, subjects actively decide in either one (short-horizon) or six (long-horizon) ”free trials,” balancing exploration and exploitation.

Fixed points of reinforcement learning models

We implemented one-dimensional (1D) and two-dimensional (2D) model-free reinforcement learning (RL) models (Ji-An et al., 2023), each fitted to individual subjects' behavior (see performance in Fig. 1b). In the 2D RL model (Fig. 1c), the chosen action value $Q_t(a_i)$ is updated by $Q_t(a_i) = Q_{t-1}(a_i) + \alpha(r - Q_{t-1}(a_i))$, where α is the learning rate. When the model consistently selects the same action a_i and receives the same reward r, $Q(a_i)$ converges to r, while the unchosen action value $Q(\bar{a}_i)$ decays to 0, representing a fixed point $(Q^*(a_i), Q^*(\bar{a}_i)) = (r, 0)$. Thus, each action with varying rewards corresponds to a line attractor, positioned at $Q_L = 0$ or $Q_R = 0$, with these two attractors orthogonal to each other. In the 1D RL model (Fig. 1c), the unchosen action value $(Q_t(\bar{a}_i))$ is always completely anti-correlated with the chosen action value $(Q_t(\bar{a}_i) = -Q_t(a_i))$. When consistently selecting action a_i and receiving reward r, the model converges to the fixed point $(Q^*(a_i), Q^*(\bar{a}_i)) = (r, -r)$. Thus, the two line attractors are anti-parallel to each other, positioned at $Q_L = -Q_R$.

Training two-dimensional RNNs for individuals

We first trained a large teacher RNN to predict the choices of all subjects. Subsequently, we trained individual-specific two-dimensional student RNNs (a version of lowrank RNN with a gating mechanism, see (Xiong, Ji-An, Mattar, & Wilson, 2023)) to predict the logits provided by the teacher RNN. This knowledge distillation reduces the number of required trials per participant (Ji-An et al., 2023).

To facilitate interpretability of student RNNs, we used a diagonal matrix for the readout from the two-dimensional recurrent layer to the output layer (h_1 and h_2 corresponding to two actions). Thus, each h_i corresponds to βQ_i in classical cognitive models, where β, the inverse temperature, indicates behavioral stocasticity. Our teacher RNNs and both 1D and 2D student RNNs outperformed the best-known cognitive models in predicting human choices (Fig. 1b; evaluated with nested cross-validation).

Characterize attractor geometry of RNNs

In the RNNs, the hidden state $\mathbf{h}_t$ at time t is updated using the function $\mathbf{h}_{t+1} = F(\mathbf{h}_t, \mathbf{x}_t)$, where $\mathbf{x}_t$ represents the input. Fixed points $h^*(r,a)$ are hidden states that remain approximately stable under the update dynamics when input action a and reward r are constant, such that $\mathbf{h}^* \approx F(\mathbf{h}^*, \mathbf{x})$. We identify these fixed points numerically by minimizing the squared speed of dynamics $|\mathbf{h} - F(\mathbf{h}, \mathbf{x})|^2$ with respect to the hidden states $\mathbf{h}$ (Sussillo & Barak, 2013). This optimization is per-

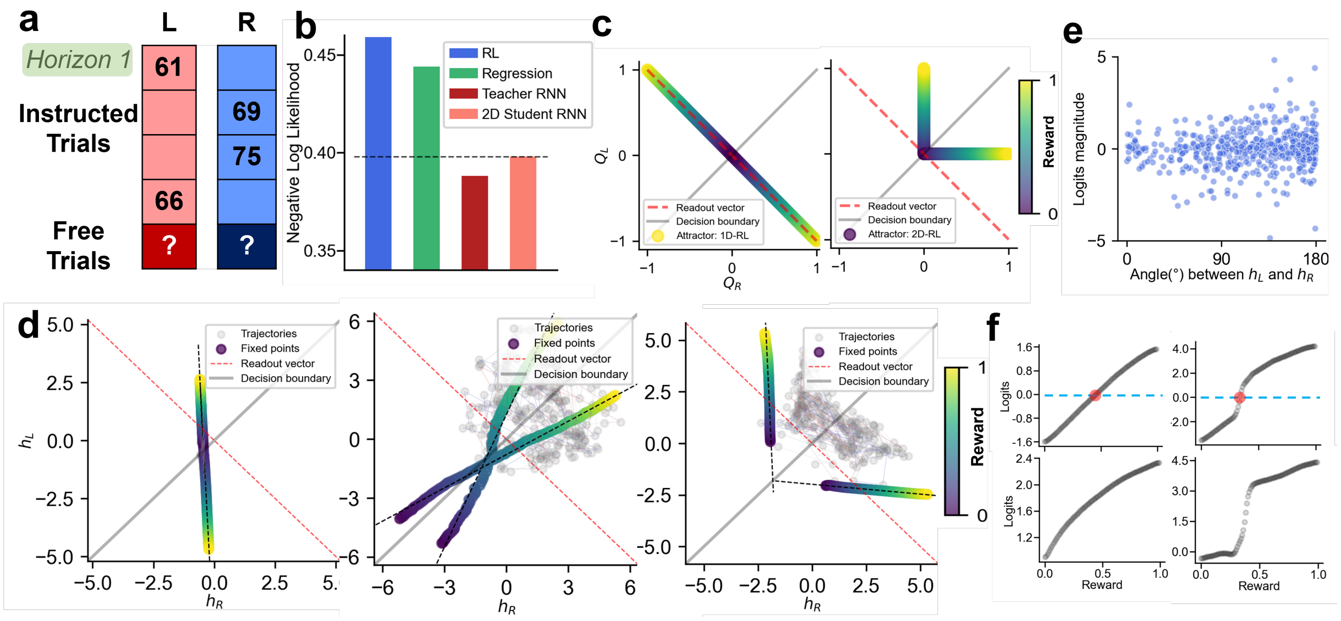

Figure 1: **a.** Schematic of the Horizon task, showing the short-horizon condition (only one free trial). **b.** The predictive performance (test negative log-likelihood in nested cross-validation; lower is better) of RNNs and cognitive models. The logistic regression model is used in the original study (Wilson et al., 2014). **c.** Line attractors of fixed points in one-dimensional and two-dimensional model-free reinforcement learning models. **d.** Line attractors of fixed points in two-dimensional RNNs fitted to individual subject's behavior, showing three example subjects. Each gray point corresponds to a trial actually encountered by the subject. (Left) Anti-parallel updates, as in the one-dimensional model. (Middle) Partially correlated updates. (Right) Approximately orthogonal updates, as in the two-dimensional model. **e.** Action angle versus maximum value of projected logits. Each point is a subject. **f.** Projected logit as a function of reward, showing four example subjects.

formed using Adam optimizer multiple times, with different initializations sampled from hidden states explored by the RNN.

For each subject-specific RNN, we found two line attractors (each corresponding to one of the two actions), consisting of marginally stable fixed points. Each fixed point, associated with a specific reward, is depicted in Fig. 1d.

Action angles between two line attractors reveal correlated value updates

We parameterized fixed points along the line attractor with their first principal direction. We then measured the *action angle* between the principal directions of the two line attractors, representing the correlation between the updates for two actions. A degree of 90° indicates orthogonal updates, as in the 2D RL model; while a degree of 180° indicates completely anti-correlated updates, as in the 1D RL model. Our RNNs revealed that different subjects exhibit distinct action angles (see three example subjects in Fig. 1d), illustrating diverse individual strategies in this task (refer to the x-axis in Fig. 1e).

Parameterizations of fixed points uncover diverse forms of reward utilities

To assess the impact of reward magnitude on action preference, we projected the first principal components of fixed points (their parameterizations) on the readout vector (i.e., the logit axis, $L = h_L - h_R$). The maximum value of the projected logit is akin to the inverse temperature in RL models, capturing the level of behavioral stochasticity (refer to y-axis in Fig. 1e). We plotted the projected logit L as a function of reward r. In a model-free RL model using r to update action values, this correponds to the linear function $L = \beta u(r) = \beta r$, where u is the identity utility function.

Our RNNs uncovered a variety of reward utility functions across different subjects (see four example subjects in Fig. 1f). We offer several key insights: First, the reward at $L = 0$ serves as a reference point (red point in Fig. 1f), above which the received reward favors the chosen action. Second, the slope of this function indicates reward sensitivity. Third, the shapes of these functions are reminiscent of prospect utility theory (Kahneman & Tversky, 1979).

Conclusion

Our study utilizes low-dimensional individual-specific RNNs to model human decision-making in the Horizon task, revealing novel and diverse cognitive strategies that traditional cognitive models often overlook. The attractor geometry via the fixed point analysis of these RNNs revealed a spectrum of correlated value updates and diverse forms of utility functions, providing a deeper insight into the complexities of human exploration-exploitation dynamics. This approach demonstrates the potential of RNNs in revealing intricate cognitive mechanisms through the lens of dynamical systems theory and pave the way for more personalized and accurate cognitive modeling, which could be instrumental for relating cognitive models to psychiatry and neuroscience.

Acknowledgments

This work was funded by Kavli Institute for Brain and Mind Innovative Research Grant #2022-2209 to LJA; a National Institutes of Health grant R01AG061888 to RCW.

References

Ji-An, L., Benna, M. K., & Mattar, M. G. (2023). Automatic discovery of cognitive strategies with tiny recurrent neural networks. *bioRxiv*, 2023–04.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*(2), 263–292.

Miller, K. J., Eckstein, M., Botvinick, M. M., & Kurth-Nelson, Z. (2023). Cognitive model discovery via disentangled rnns. *bioRxiv*, 2023–06.

Sussillo, D., & Barak, O. (2013, March). Opening the Black Box: Low-Dimensional Dynamics in High-Dimensional Recurrent Neural Networks. *Neural Computation*, *25*(3), 626–649. doi: 10.1162/NECO$_{a0}0409$

Wilson, R. C., Geana, A., White, J. M., Ludvig, E. A., & Cohen, J. D. (2014). Humans use directed and random exploration to solve the explore–exploit dilemma. *Journal of Experimental Psychology: General*, *143*(6), 2074.

Xiong, H.-D., Ji-An, L., Mattar, M. G., & Wilson, R. C. (2023). Distilling human decision-making dynamics: a comparative analysis of low-dimensional architectures. In *Neurips 2023 ai for science workshop*.