Fast and robust Bayesian inference for modular combinations of dynamic learning and decision models

Krishn Bera (krishn_bera@brown.edu)

Alexander Fengler (alexander_fengler@brown.edu)

Michael J. Frank (michael_frank@brown.edu)
Dept. of Cognitive, Linguistic & Psychological Sciences
Carney Institute for Brain Sciences
Brown University, RI (US)

Abstract

In cognitive neuroscience, there has been growing interest in adopting sequential sampling models (SSM) as the choice function for reinforcement learning (RLSSM), opening up new avenues for exploring generative processes that can jointly account for decision dynamics within and across trials. To date, such approaches have been limited by computational tractability, e.g., due to lack of closed-form likelihoods for the decision process and expensive trial-by-trial evaluation of complex reinforcement learning (RL) processes. We enable hierarchical Bayesian parameter estimation for a broad class of RLSSM models, using Likelihood Approximation Networks (LANs) in conjunction with differentiable RL likelihoods to leverage fast gradient-based inference methods including Hamiltonian Monte Carlo or Variational Inference (VI). By exploiting the differentiability of RL likelihoods, this method improves scalability and enables faster convergence for complex combinations of RL and decision processes. To showcase these methodological advantages, we consider multiple interacting generative learning processes with the Reinforcement Learning - Working Memory (RLWM) task and model. This RLWM model is then combined with SSMs via LANs. When combined with hierarchical variational inference, this approach can accurately recover the posterior parameter distributions in complex RLSSM paradigms, and moreover, that in comparison, fitting data with the equivalent choice only RLWM model yields a biased estimator of the true generative process.

Keywords: Hierarchical Bayesian Inference; Variational Inference; Reinforcement Learning; Sequential Sampling Models; Likelihood-Free Inference

Introduction

Reinforcement Learning – Sequential Sampling Models (RLSSM) are a powerful and expressive class of models that are naturally suited for computational modeling of cognitive tasks where the learning process informs the decision-making process. However, to date empirical data analysis has been mostly limited to basic instances of RLSSM that employ Drift Diffusion Models or simple race models and n armed bandits (Pedersen, Frank, & Biele, 2017; Fontanesi, Gluth, Spektor, & Rieskamp, 2019; Miletić et al., 2021). Great theoretical interest in more elaborate models is established, however Bayesian parameter inference for such models is hampered by computational complexity due to a lack of closed-form likelihood for the decision process, a complex reinforcement learning (RL) process, or a combination of both (Fengler, Bera, Pedersen, & Frank, 2022).

Here, we show how Variational Inference (VI) methods (Blei, Kucukelbir, & McAuliffe, 2017; Jordan, Ghahramani, Jaakkola, & Saul, 1998; Liu & Wang, 2016) in combination with LANs and differentiable complex RL dynamics, can be leveraged for computationally efficient and scalable Bayesian inference treatment of a broad class of RLSSM models. While we focus on a specific example, and link to empirical findings of independent interest, we stress that the methods explored are much more generally applicable. Using empirical and synthetic datasets of a widely used cognitive task, the Reinforcement Learning - Working Memory (RLWM) task (Collins & Frank, 2012), we show that this approach can yield fast, tractable inference with a high-dimensional model ($D = 10$ for each participant, with ~ 900 free model parameters) on a large dataset ($> 31k$ trials) in hierarchical settings to recover the posterior distribution over parameters accurately.

Method

Participants: All participants were recruited from five different US locations as part of the Cognitive Neuroscience Test Reliability and Computational Applications for Schizophrenia Consortium (CNTRaCS). For this study, we model the behavioral data from the control group (n=87) to experiment with and refine the methodological advances proposed in this work.

Reinforcement Learning Working Memory (RLWM) task and model: The RLWM task (Collins & Frank, 2012; Collins, Brown, Gold, Waltz, & Frank, 2014; McDougle & Collins, 2021) is designed to disentangle the contributions of reinforcement learning and working memory (WM) to stimulus-response learning. Participants learn stimulus-response associations through trial-and-error feedback in a 3-alternative forced-choice paradigm. In addition to incremental RL, the task systematically varies WM demands by manipulating the set size (the number of unique stimuli, ranging from 2 to 5 per block) and delay (the number of trials before re-encountering a stimulus). Participants complete 10 training blocks for 360 trials total. We compare two computational models that are slight variants of existing RLWM models – RLWM with Softmax decision process (choice-only model) and RLWM with collapsing bound Linear Ballistic Accumulator (LBA-Angle) decision process (choice and reaction time model). We refer the reader to the cited papers for details about computational modeling. Since RTs show systematic effects of WM load, the inclusion of RTs in modeling could constrain and help interpret the generative RL and decision process.

Variational Inference: VI approximates the posterior by optimizing a parameterized family of distributions to minimize divergence from the true posterior. The combination of differentiable surrogate likelihoods via LANs and differentiable RL implementation allows computation of evidence lower bound (ELBO) gradients with respect to variational parameters, enabling efficient updates via gradient-based methods for VI. In this work, we used Automatic Differentiation Variational Inference (ADVI; Kucukelbir, Tran, Ranganath, Gelman, and Blei (2017)), as implemented in PyMC (Abril-Pla et al., 2023). All experiments were performed by recovering the parameters hierarchically with the same non-informative priors. The best-fit run (out of 20 inference runs) as determined by the log probability of fit was used for all further analysis. The ELBO loss was monitored for convergence and stability.

The likelihood was implemented in a JAX (Bradbury et al., 2018) function and wrapped in PyTensor (Developers, 2024) to interface with PyMC for all autodiff/numerical computations. We trained a multilayer perceptron with five layers to approximate the likelihoods of the joint choices and decision dynamics (RT distributions for each of the three choices), as per the procedures outlined in Fengler, Govindarajan, Chen, and Frank (2021).

Results

VI significantly boosts the speed of inference: model fits are achieved in under 60min. Attempting the same analysis with MCMC, took between one and two orders of magnitude more time.

RLSSM model captures both choice and RT distributions: Figure 1 (e) shows a comparison of common participant-level parameters between the RLWM LBA-Angle model and RLWM Softmax models. The RLWM Softmax overestimates the WM reliance parameter (ρ), WM capacity parameter (C) as well as RL learning rate (α) relative to the RLWM LBA-Angle model. Indeed, if one simulates RTs from the LBA-Angle model using the αs estimated by the Softmax model, the RTs become progressively faster at a higher rate than observed empirically (Figure 1 b). In contrast, the RLSSM jointly models the choice and RT distributions, and hence, imposes further constraints on identifying the generative RL parameters separable from WM. Posterior predictive checks confirmed this interpretation. We generated data using estimated parameters (sub-sampled by a factor of 50 from the posterior) and compared the observed and simulated results. As predicted, the RL-SSM model more appropriately captures both choice and RT dynamics and moreover, specifically at higher set sizes when RL is more dominant (Figure 1 a-b).

RLSSM model recovers the parameters better than the RL-only model: We performed parameter recovery (Figure 1 c-d) to further assess whether these differences allow RLSSM parameters to be more identifiable than the Softmax version. We simulate a synthetic dataset based on the parameters that best-fit to our empirical dataset and attempted to recover the true data-generating parameters. Figure 1 (c) shows that RLWM LBA-Angle model was able to recover all parameters with reasonable accuracy, even recovered all common parameters significantly better than the RLWM Softmax model (d). This suggests that the RLWM LBA-Angle is a better generative model for the empirical data and thus highlighting the potential of joint models of choice and RT data to improve parameter identification in cognitive models (Ballard & McClure, 2019). Importantly, the RLWM Softmax model overestimates WM reliance (ρ), WM capacity (C) and, notably, the RL learning rate (α). The parameter recovery results recapitulate the tendency of the RLWM Softmax model to overestimate certain parameters (Figure 1 e). The choice-only Softmax model can give biased estimates of the model parameters especially when the RT distributions are informative of the underlying cognitive processes. Our results highlight the ease by which we may be lead astray concerning the generative mechanisms underlying our data, if we do not take into account all information present in our dataset (RTs and choices). The combination of methods we illustrate here, allows a broad class of RLSSM models to be tested on large scale experimental datasets, and hence to properly exploit the availability of RT data where traditionally computational considerations imposed hard limits.

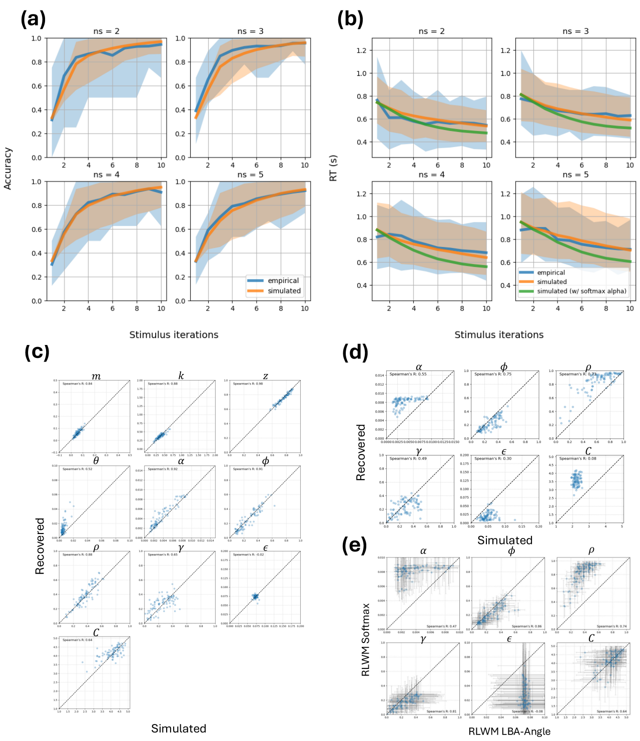

Figure 1: Posterior predictive checks for (a) accuracy and (b) reaction times grouped by set size (ns). The solid lines indicate mean across participants and simulations. The shaded region corresponds to 94% HDI ranges. Parameter recovery results with (c) RLWM LBA-Angle and (d) RLWM Softmax model. (e) Comparison of common participant-level parameters recovered between RLWM LBA-Angle and RLWM Softmax models. The error bars indicate 94% HDI range.

Acknowledgements

This work was supported by the Brainstorm Program at the Robert J. Nancy D. Carney Institute for Brain Science, NIMH grants NIMH P50MH119467 and NIMH R01 MH084840, and ONR grant ONR MURI Award N00014-23-1-2792.

References

Abril-Pla, O., Andreani, V., Carroll, C., Dong, L., Fonnesbeck, C. J., Kochurov, M., ... Zinkov, R. (2023). PyMC: a modern, and comprehensive probabilistic programming framework in Python. *PeerJ Computer Science*, *9*, e1516.

Ballard, I. C., & McClure, S. M. (2019). Joint modeling of reaction times and choice improves parameter identifiability in reinforcement learning models. *Journal of Neuroscience Methods*, *317*, 37–44.

Blei, D. M., Kucukelbir, A., & McAuliffe, J. D. (2017). Variational Inference: A Review for Statisticians. *Journal of the American Statistical Association*, *112*(518), 859–877.

Bradbury, J., Frostig, R., Hawkins, P., Johnson, M. J., Leary, C., Maclaurin, D., ... Zhang, Q. (2018). *JAX: composable transformations of Python+NumPy programs.* Retrieved from `http://github.com/jax-ml/jax`

Collins, A. G. E., Brown, J. K., Gold, J. M., Waltz, J. A., & Frank, M. J. (2014). Working Memory Contributions to Reinforcement Learning Impairments in Schizophrenia. *Journal of Neuroscience*, *34*(41), 13747–13756.

Collins, A. G. E., & Frank, M. J. (2012). How much of reinforcement learning is working memory, not reinforcement learning? A behavioral, computational, and neurogenetic analysis. *European Journal of Neuroscience*, *35*(7), 1024–1035.

Developers, P. (2024). *pytensor.* Retrieved from `https://github.com/pymc-devs/pytensor`

Fengler, A., Bera, K., Pedersen, M. L., & Frank, M. J. (2022). Beyond Drift Diffusion Models: Fitting a Broad Class of Decision and Reinforcement Learning Models with HDDM. *Journal of Cognitive Neuroscience*, *34*(10), 1780–1805.

Fengler, A., Govindarajan, L. N., Chen, T., & Frank, M. J. (2021). Likelihood approximation networks (LANs) for fast inference of simulation models in cognitive neuroscience. *eLife*, *10*, e65074.

Fontanesi, L., Gluth, S., Spektor, M. S., & Rieskamp, J. (2019). A reinforcement learning diffusion decision model for value-based decisions. *Psychonomic Bulletin & Review*, *26*(4), 1099–1121.

Jordan, M. I., Ghahramani, Z., Jaakkola, T. S., & Saul, L. K. (1998). An Introduction to Variational Methods for Graphical Models. In M. I. Jordan (Ed.), *Learning in Graphical Models* (pp. 105–161). Springer Netherlands.

Kucukelbir, A., Tran, D., Ranganath, R., Gelman, A., & Blei, D. M. (2017). Automatic Differentiation Variational Inference. *Journal of Machine Learning Research*, *18*(14), 1–45.

Liu, Q., & Wang, D. (2016). *Stein Variational Gradient Descent: A General Purpose Bayesian Inference Algorithm.* arXiv.

McDougle, S. D., & Collins, A. G. E. (2021). Modeling the influence of working memory, reinforcement, and action uncertainty on reaction time and choice during instrumental learning. *Psychonomic Bulletin & Review*, *28*(1), 20–39.

Miletić, S., Boag, R. J., Trutti, A. C., Stevenson, N., Forstmann, B. U., & Heathcote, A. (2021). A new model of decision processing in instrumental learning tasks. *eLife*, *10*, e63055.

Pedersen, M. L., Frank, M. J., & Biele, G. (2017). The drift diffusion model as the choice rule in reinforcement learning. *Psychonomic Bulletin & Review*, *24*(4), 1234–1251.