rewardGym - a framework for streamlining experiments in cognitive neuroscience

Simon R. Steinkamp (simons@drcmr.dk)
Danish Research Centre for Magnetic Resonance, Department of Radiology and Nuclear Medicine,
Copenhagen University Hospital - Amager and Hvidovre, Copenhagen, Denmark

David Meder (davidm@drcmr.dk)
Danish Research Centre for Magnetic Resonance, Department of Radiology and Nuclear Medicine,
Copenhagen University Hospital - Amager and Hvidovre, Copenhagen, Denmark

Oliver J. Hulme (oliverh@drcmr.dk
Danish Research Centre for Magnetic Resonance, Department of Radiology and Nuclear Medicine
Copenhagen University Hospital - Amager and Hvidovre, Copenhagen, Denmark;
London Mathematical Laboratory
London, United Kingdom;
Department of Psychology, University of Copenhagen
Copenhagen, Denmark

Abstract

Despite significant efforts, psychological and cognitive sciences currently lack standardized frameworks for setting up tasks and models in ways that can be integrated with data, generalized over tasks, and shared between researchers. Typically, researchers program an experimental task, assess how a theoretical model would solve it, and analyze behavioral data, all using custom code. Due to a lack of standardization, cognitive models are often experiment-specific, making it difficult to test the generalizability of models across different tasks and increasing the burden of testing existing models on new tasks. Here, we present a standardized framework that addresses these challenges. We build on the Gymnasium standard from reinforcement learning (RL), which defines how artificial agents interact with computational environments. This standard helps us establish a common graphical language for different tasks that captures the connections between states, actions, and rewards. This representation is further inspired by neuro-nav, which extends the Gymnasium framework to classical neuroscience experiments and focuses on neurally plausible RL. By expressing tasks in a formal language, it is possible to build libraries of models where agents can perform each task. This allows standardized software to perform parameter inference and model comparison on real and synthetic data. What distinguishes our framework is its focus on running experiments. We provide a high degree of control over the environments (e.g., stimulus order) and a direct way to augment the graphical representations with stimulus information. This allows for a direct transition from simulations with artificial agents to running experiments with human participants using PsychoPy. Additionally, we provide logging utilities that save data in BIDS format, which is standard in neuroimaging. With this framework, we hope to make psychological and cognitive science more reproducible and robust.

Keywords: reinforcement learning; cognitive science; software; experiments

Introduction

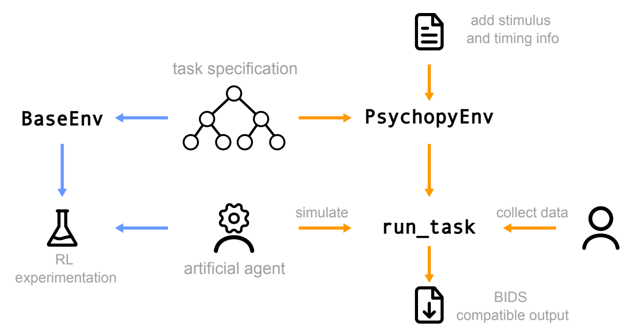

Figure 1: **Overview of the framework.** At its core is the task specification, which utilises a graph structure. Using this structure, the user can do classical RL experiments (left) using the `BaseEnv` class. By augmenting the graph with stimulus information, the `PsychopyEnv` can collect data from human participants and deploy artificial agents to simulate data. For this, we implemented the convenience function `run_task`, which stores simulated and real data in BIDS format.

Improved research practices in cognitive and psychological sciences in response to the reproducibility crisis (Korbmacher et al., 2023) have also revealed a lack of standardization: Experiments are often customized and use non-free, proprietary software (Borghi and Gulick, 2021). Although significant efforts have been made to standardise experiments (Sochat et al., 2016; Sochat, 2018) and provide standards by which to describe them (Poldrack et al., 2011), these efforts often stop before the data analysis stage.

Here we present `rewardGym`[1] as an approach to streamlining experiments from a computational modelling side. We

[1] https://github.com/rewardMap/rewardGym

build on the Gymnasium[2] framework, developed by the RL community, which provides a standardized interface by which artificial agents can interact with computational environments (e.g., games, control tasks, etc.) (Brockman et al., 2016; Towers et al., 2024). While Gymnasium and other frameworks can render the implemented environments and thus allow for human interaction (Towers et al., 2024; Juliani et al., 2022), they might not fulfil the needs of psychology and neuroscience researchers running experiments. We aim to close this gap by providing tools that augment basic environments, to allow for data collection via PsychoPy (Peirce et al., 2019) and giving a high degree of control over the experiment. The collected data is stored in compliance with the Brain Imaging Data Structure (BIDS) developed by the neuroimaging community to facilitate data sharing and collaboration (Gorgolewski et al., 2016).

We present a short description of the software and provide a Jupyter notebook on Binder with a worked example [3].

Software Description

rewardGym is aimed primarily at cognitive neuroscientists and psychologists, who study learning and decision-making in humans with simple experiments.

We build on Gymnasium to set up our environments, relying on the abstraction of the (partially observable) Markov decision process (POMDP) that became a standard in RL. As in Gymnasium, an environment has a reset method that samples the initial observation and a step method, which takes an action and returns an observation and the obtained reward (Towers et al., 2024).

This basic functionality is implemented in the BaseEnv class (left of Fig. 1). This interface allows for classical RL analyses such as the benchmarking of algorithms similar to neuro-nav and Gymnasium (Juliani et al., 2022; Towers et al., 2024). Abstracting an experiment this way allows us to have a common language for different tasks. In our case, and inspired by neuro-nav, we use a directed graph representation (see Fig. 2). Each graph consists of nodes and edges. A node is either a decision point or a terminal state and can be associated with a reward function, whereas the edges represent how an action leads to the next state.

An advantage of graphs is that they can be augmented, e.g., expanding the BaseEnv by adding stimulus displays and timings to the nodes. The PsychopyEnv utilizes the additional information for rendering and response collection with PsychoPy[4] (Peirce et al., 2019).

Finally, the run_task function provides a convenience wrapper to run the whole experiment. It has been adapted to the idiosyncrasies of the six tasks implemented so far (two-step, risk-sensitive, human connectome gambling, monetary incentive delay, gonogo, and posner tasks). The function saves the collected data (in BIDS format) and can be used with artificial agents to simulate task behavior and replicating the output files collected from human participants. This allows researchers to simulate and test the putative behavioral consequences of their computational theories and prepare analyses before data collection.

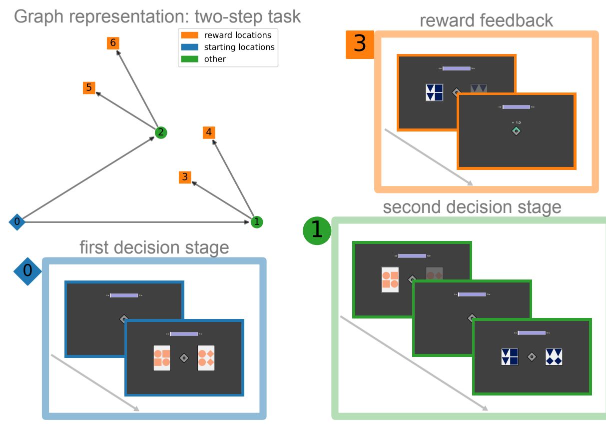

Figure 2: **Implementation of the two-step task (Daw et al., 2011).** Participants in this task need to make a first-stage decision between two images (node 0). This first stage choice determines which second-stage stimulus pair will be displayed (node 1 or 2). Each of the four images in the second stage is associated with a different reward probability (nodes 3 - 6). In rewardGym we can now augment these nodes with stimulus information. Node 0, for example, will show an initial fixation cross and then the first-stage decision. The program will then wait for a response.

Limitations So far, our framework allows only for tabular RL and does not support deep RL. This decision was deliberate to make rewardGym compatible with standalone PsychoPy and because the cognitive tasks we aim to use have a simple structure. Other frameworks will be more suitable for more complex environments (e.g., Juliani et al., 2022).

Since we are developing the framework for a larger neuroimaging project, we needed to make several pragmatic decisions that might provide challenges for downstream maintainability – which we aim to address in the future. For example, we use task-specific conditions in the run_task and use an ad hoc implementation to allow for the high degree of control researchers need over their tasks.

Discussion

With rewardGym we hope to inspire further steps towards standardization in psychology and cognitive neuroscience by building upon standards developed in neuroimaging and RL communities. These standards should facilitate developments in the field. Using a common language for different tasks, we can test if and how the same artificial agents can perform multiple tasks. This is essential for testing how computational theories of cognition generalize over tasks—presenting a much more stringent test of theories.

[2] https://gymnasium.farama.org/
[3] https://mybinder.org/v2/gh/rewardMap/exampleWorkflow/binder ?urlpath=%2Fdoc%2Ftree%2FCCNRLDMsims.ipynb
[4] https://psychopy.org/

First, our framework makes it possible to recycle the code behind a single model to simulate and evaluate its performance on different tasks. With this model in place, it becomes more straightforward and efficient to simultaneously test task assumptions and expected behavior. Second, we can use the same code again for parameter and model inference. Again, this substantially reduces the burden and complexity of writing customized code. These steps are essential due diligence for assessing whether an experiment is capable of meaningfully answering the theoretical questions that the experiment aims to answer. We plan to release another package implementing such reproducible and standardized workflows for model inference and model diagnostics. Finally, our framework can be used for simulating realistic log files, a small but important feature, allowing researchers to carefully plan their analyses before data collection.

Using `rewardGym` as a basis, we hope that more researchers from cognitive psychology and neuroscience will adopt standards from RL for their research, and we hope that these ideas are adopted for other studies, thus allowing for the greater adoption of shareable and generalizable models.

Acknowledgments

SRS was funded by a Lundbeck Foundation postdoctoral grant (ref R402-2022-1411). DM was funded by a Sapere Aude DFF-starting grant from the Independent Research Fund Denmark (Grant No. 1052-00054B). OJH was funded by a Novo Nordisk Foundation Exploratory Interdisciplinary Synergy Grant (ref NNF20OC0064869).

References

Borghi, J. A. and Gulick, A. E. V. (2021). Data management and sharing: Practices and perceptions of psychology researchers. *PLOS ONE*, 16(5):e0252047.

Brockman, G., Cheung, V., Pettersson, L., Schneider, J., Schulman, J., Tang, J., and Zaremba, W. (2016). OpenAI Gym.

Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., and Dolan, R. J. (2011). Model-Based Influences on Humans' Choices and Striatal Prediction Errors. *Neuron*, 69(6):1204–1215.

Gorgolewski, K. J., Auer, T., Calhoun, V. D., Craddock, R. C., Das, S., Duff, E. P., Flandin, G., Ghosh, S. S., Glatard, T., Halchenko, Y. O., Handwerker, D. A., Hanke, M., Keator, D., Li, X., Michael, Z., Maumet, C., Nichols, B. N., Nichols, T. E., Pellman, J., Poline, J.-B., Rokem, A., Schaefer, G., Sochat, V., Triplett, W., Turner, J. A., Varoquaux, G., and Poldrack, R. A. (2016). The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. *Scientific Data*, 3(1):160044.

Juliani, A., Barnett, S., Davis, B., Sereno, M., and Momennejad, I. (2022). Neuro-Nav: A Library for Neurally-Plausible Reinforcement Learning.

Korbmacher, M., Azevedo, F., Pennington, C. R., Hartmann, H., Pownall, M., Schmidt, K., Elsherif, M., Breznau, N., Robertson, O., Kalandadze, T., Yu, S., Baker, B. J., O'Mahony, A., Olsnes, J. Ø.-S., Shaw, J. J., Gjoneska, B., Yamada, Y., Röer, J. P., Murphy, J., Alzahawi, S., Grinschgl, S., Oliveira, C. M., Wingen, T., Yeung, S. K., Liu, M., König, L. M., Albayrak-Aydemir, N., Lecuona, O., Micheli, L., and Evans, T. (2023). The replication crisis has led to positive structural, procedural, and community changes. *Communications Psychology*, 1(1):1–13.

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., and Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, 51(1):195–203.

Poldrack, R., Kittur, A., Kalar, D., Miller, E., Seppa, C., Gil, Y., Parker, D., Sabb, F., and Bilder, R. (2011). The Cognitive Atlas: Toward a Knowledge Foundation for Cognitive Neuroscience. *Frontiers in Neuroinformatics*, 5.

Sochat, V. V. (2018). The Experiment Factory: Reproducible Experiment Containers. *Journal of Open Source Software*, 3(22):521.

Sochat, V. V., Eisenberg, I. W., Enkavi, A. Z., Li, J., Bissett, P. G., and Poldrack, R. A. (2016). The Experiment Factory: Standardizing Behavioral Experiments. *Frontiers in Psychology*, 7:610.

Towers, M., Kwiatkowski, A., Terry, J., Balis, J. U., Cola, G. D., Deleu, T., Goulão, M., Kallinteris, A., Krimmel, M., KG, A., Perez-Vicente, R., Pierré, A., Schulhoff, S., Tai, J. J., Tan, H., and Younis, O. G. (2024). Gymnasium: A Standard Interface for Reinforcement Learning Environments.