Neuro-Nav: A Library for Neurally-Plausible Reinforcement Learning

Arthur Juliani
Microsoft Research
New York City, NY USA
awjuliani@gmail.com

Samuel A. Barnett
Princeton University
Princeton, NJ USA
samuelab@princeton.edu

Brandon Davis
MIT
Cambridge, MA USA
davisb@mit.edu

Margaret Sereno
University of Oregon
Eugene, OR USA
msereno@uoregon.edu

Ida Momennejad
Microsoft Research
New York City, NY USA
idamo@microsoft.com

Abstract

In this work we propose Neuro-Nav, an open-source library for neurally plausible reinforcement learning (RL). RL is among the most common modeling frameworks for studying decision making, learning, and navigation in biological organisms. In utilizing RL, cognitive scientists often handcraft environments and agents to meet the needs of their particular studies. On the other hand, artificial intelligence researchers often struggle to find benchmarks for neurally and biologically plausible representation and behavior (e.g., in decision making or navigation). In order to streamline this process across both fields with transparency and reproducibility, Neuro-Nav offers a set of standardized environments and RL algorithms drawn from canonical behavioral and neural studies in rodents and humans. We demonstrate that the toolkit replicates relevant findings from a number of studies across both cognitive science and RL literatures, and can furthermore be extended with novel algorithms and environments to address future research needs of the field.

Keywords: reinforcement learning, open source, navigation, neuroscience

Introduction

Understanding how animals make non-reactive decisions in their environments remains a longstanding interest of psychologists and neuroscientists alike (Tolman, 1948). Applying the formalism of reinforcement learning (RL) to decision making has accelerated the progress in this field within recent decades (Sutton & Barto, 2018). Indeed, there is neuroscientific evidence that various forms of learning in humans and other animals can be modeled using the formalism of RL, and thus be described by classical algorithms from the RL literature (Niv, 2009; Momennejad, 2020).

Progress in this research program is made possible by abstracting the underlying complexity of the real environment-agent system into an idealized model. This process of abstraction is largely non-standardized, and many of the concrete implementations of these models are not available in open-source forms, change when translated from one coding language to another, or the code has been lost altogether. As a result, the ability to replicate or build on previous work has been limited, slowing progress in the field.

In this work, we present Neuro-Nav[1], an open-source library providing a standardized set of benchmark environments and RL algorithms which can be used to both replicate previous findings, test the biological plausibility of new models, and provide scaffolding for future experimental work in the field. The benchmarks focus on domains that are the basis of many experimental studies: spatial navigation and associative learning in 2D mazes (Russek, Momennejad, Botvinick, Gershman, & Daw, 2017)(Figure 1 Top) and abstract graphs (Momennejad et al., 2017) (Figure 1 Bottom). In addition, we provide standardized implementations of classical RL algorithms which can be evaluated using the benchmark environments. We provide these all within the context of a documented and tested open-source repository, which can be used and developed further by the broader research community.

[1] https://github.com/awjuliani/neuro-nav

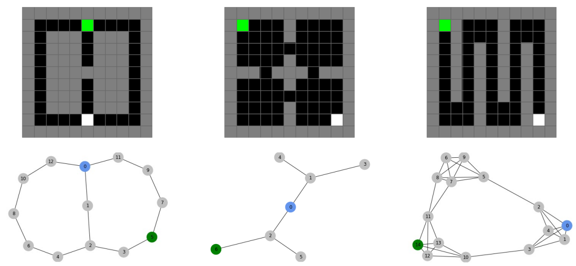

Figure 1: **Top**: Examples of maze environments. White square corresponds to agent location. Green square corresponds to goal location. Grey squares correspond to walls. **Bottom**: Examples of graph environments. Blue circle corresponds to agent location. Green circle corresponds to goal location.

Neuro-Nav Library

Neuro-Nav is an open source python library consisting of three components: a set of benchmark environments, a toolkit containing artificial reinforcement learning agents and algorithms, and a set of interactive notebooks for replicating experimental findings in the literature.

Benchmark Environments

Neuro-Nav contains of a set of graph and maze navigation tasks. In both cases, the underlying environment dynamics consists of a graph of 'state' nodes connected by 'action' edges. The agent occupies one of the nodes in the graph. The graph may also contain nodes that provide reward upon the agent occupying them, as well as a goal node, which both provides a reward as well as terminates the episode. Maze environments may be seen as a specific class of graph environments where the structure is mapped to a 2D grid, and the action space always consists of four actions: movement in each of the cardinal directions. Both types of environments are implemented using the OpenAI Gym interface, thus allowing for integration with a wide number of pre-existing open source RL codebases (Brockman et al., 2016).

For each class of environments, we provide both an API for defining and interacting with the environment as well as a set of pre-defined topographies and tasks, which can be utilized for evaluation. These tasks are often taken from existing literature, and serve to aid in the partial or complete replication of various findings from those original papers. In particular, we draw from tasks defined in the context of neuroscience (Schapiro, Rogers, Cordova, Turk-Browne, & Botvinick, 2013; Stachenfeld, Botvinick, & Gershman, 2017), classical RL (Sutton & Barto, 2018), and cognitive science (Momennejad et al., 2017; Russek et al., 2017). See Figure 1 for examples of a representative subset of the included maze and graph topographies.

Algorithms Toolkit

In addition to graph and maze environments, we provide a set canonical algorithms from the RL literature. We focus on tabular rather than deep learning algorithms, as this has likewise been the focus of much of the literature at the intersection of neuroscience and RL (Russek et al., 2017). We include algorithms implemented either as model-free or model-based learning algorithms, as well as hybrid algorithms such as Dyna (Sutton & Barto, 2018). See Table 1 for a description of the provided agent types.

Algorithm Name	Function(s)
TD-Q	$Q(s,a)$
TD-AC	$V(s)$, $\pi(a\|s)$
TD-SR	$\psi(s,a)$, $\omega(s)$
Dyna-Q	$Q(s,a)$
Dyna-SR	$\psi(s,a)$, $\omega(s)$
Dyna-AC	$V(s)$, $\pi(a\|s)$
MB-V	$Q(s,a)$, $T(s'\|s,a)$
MB-SR	$\psi(s,a)$, $\omega(s)$, $Q(s,a)$, $T(s'\|s,a)$

Table 1: Reinforcement learning algorithms included in the Neuro-Nav toolkit.

Experimental Results

We validate Neuro-Nav by replicating known results from human and rodent navigation and decision-making experiments. In particular, we replicate two classes of results: behavioral, which compares the observable behavior of biological and artificial agents, and representational, comparing learned representations of biological agents and artificial agents.

In order to validate Neuro-Nav for the replication of behavioral results, we focused on a recent study in the literature, (Momennejad et al., 2017). In it, agent performance is compared to that of humans in a set of graph navigation tasks, where aspects of the environments are changed partway during learning. We present the results in Figure 2, which are consistent with those of the original study.

The internal content and structure of information utilized for decision making is also of interest to many researchers. Here we focus on two relevant internal representations: place fields and grid fields. We present Figure 3, which contains examples of each of these learned from an open field maze environment, replicating the results of Stachenfeld et al. (2017).

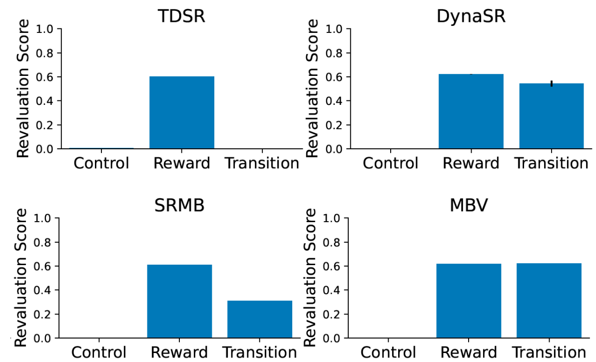

Figure 2: Replication of results from revaluation task described in Experiment 1 from (Momennejad et al., 2017). Higher revaluation score corresponds to greater change in behavior after the re-learning phase.

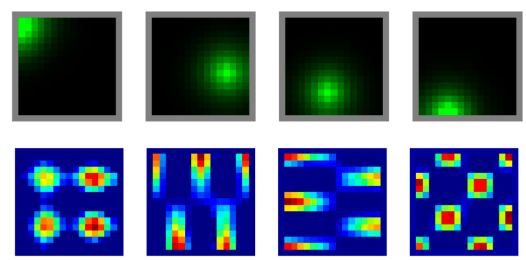

Figure 3: **Top**: Values of a selection of units in the $\psi(s)$ function in a maze environment after learning. **Bottom**: PCA of $\phi(s)$ function in a maze environment after learning.

Discussion

Here we presented Neuro-Nav, an open-source library for performing learning experiments on decision making and navigation tasks. With this project, we aim to empower reproducibility and standardization of evaluation within research in neurally plausible RL models of navigation and decision making.

The library can be extended in various ways. The first is by creating novel environment topographies and tasks to evaluate specific navigation, associative learning, or decision making experiments. The second is by defining novel observation spaces for the agents which capture relevant paradigms in the literature, such as using images as observations in nodes, or providing Euclidean distance from maze walls (Schapiro et al., 2013; de Cothi & Barry, 2020). The third is by implementing additional learning algorithms, such as those in the deep RL literature (Arulkumaran, Deisenroth, Brundage, & Bharath, 2017). While we plan to continue developing the project, we believe the current version of Neuro-Nav offers a promising step toward biologically plausible benchmarks, and thus serves as a potentially significant contributions to the field.

References

Arulkumaran, K., Deisenroth, M. P., Brundage, M., & Bharath, A. A. (2017). Deep reinforcement learning: A brief survey. *IEEE Signal Processing Magazine*, *34*(6), 26–38.

Brockman, G., Cheung, V., Pettersson, L., Schneider, J., Schulman, J., Tang, J., & Zaremba, W. (2016). Openai gym. *arXiv preprint arXiv:1606.01540*.

de Cothi, W., & Barry, C. (2020). Neurobiological successor features for spatial navigation. *Hippocampus*, *30*(12), 1347–1355.

Momennejad, I. (2020). Learning structures: predictive representations, replay, and generalization. *Current Opinion in Behavioral Sciences*, *32*, 155–166.

Momennejad, I., Russek, E. M., Cheong, J. H., Botvinick, M. M., Daw, N. D., & Gershman, S. J. (2017). The successor representation in human reinforcement learning. *Nature human behaviour*, *1*(9), 680–692.

Niv, Y. (2009). Reinforcement learning in the brain. *Journal of Mathematical Psychology*, *53*(3), 139–154.

Russek, E. M., Momennejad, I., Botvinick, M. M., Gershman, S. J., & Daw, N. D. (2017). Predictive representations can link model-based reinforcement learning to model-free mechanisms. *PLoS computational biology*, *13*(9), e1005768.

Schapiro, A. C., Rogers, T. T., Cordova, N. I., Turk-Browne, N. B., & Botvinick, M. M. (2013). Neural representations of events arise from temporal community structure. *Nature neuroscience*, *16*(4), 486–492.

Stachenfeld, K. L., Botvinick, M. M., & Gershman, S. J. (2017). The hippocampus as a predictive map. *Nature neuroscience*, *20*(11), 1643–1653.

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT press.

Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological review*, *55*(4), 189.