

# RLPy: A Value-Function-Based Reinforcement Learning Framework for Education and Research

Alborz Geramifard\*, Christoph Dann, Robert H. Klein, William Dabney\*, Jonathan P. How

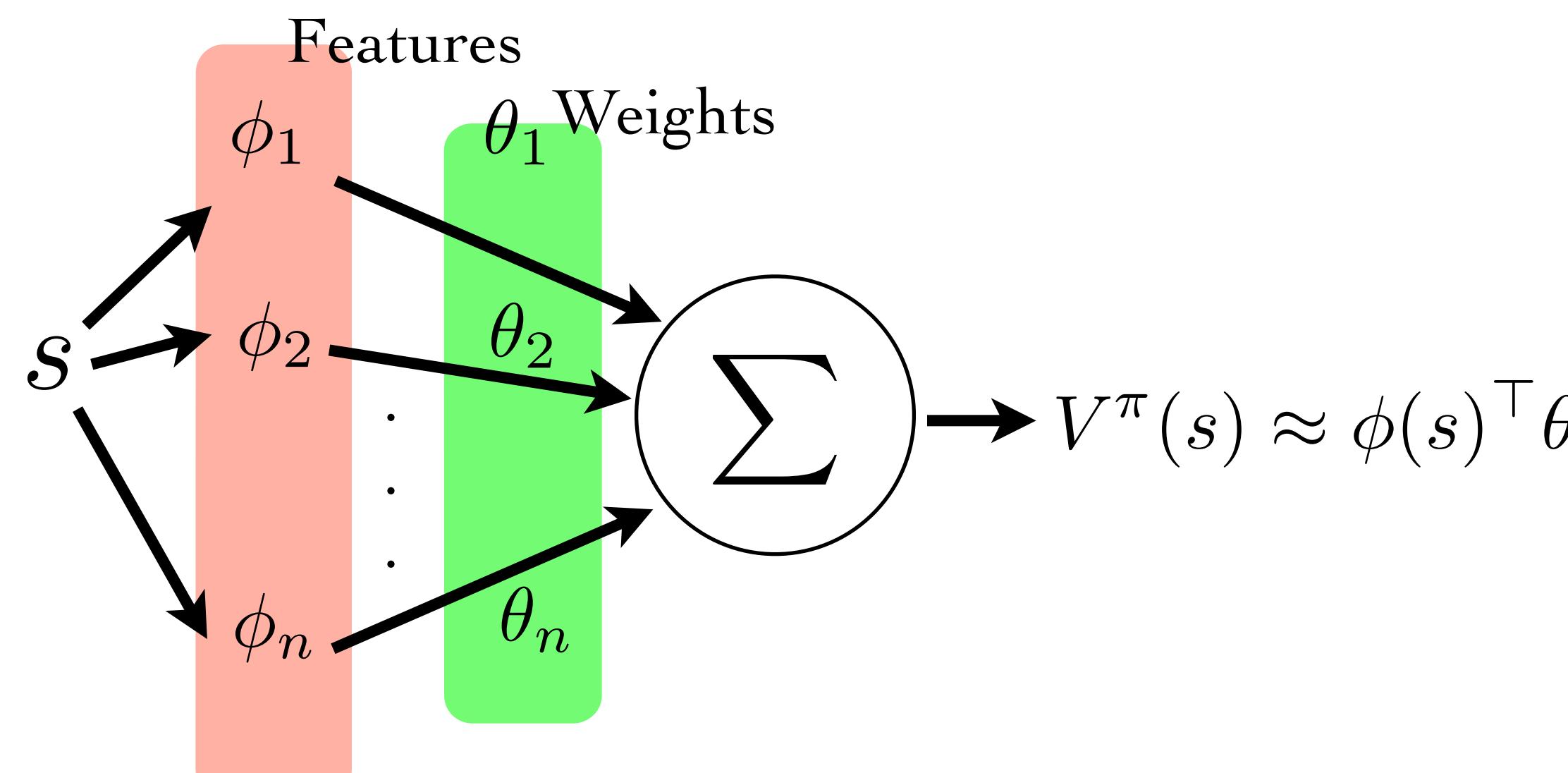


## Abstract

RLPy is an **object-oriented** reinforcement learning (RL) software package with focus on value- function-based methods using linear function approximation and discrete actions. The framework was designed for both **education** and **research** purposes. It provides a rich library of fine-grained, easily exchangeable components for learning agents (e.g., policies or representations of value functions), facilitating recent increased specialization in RL. RLPy is written in **Python** to allow fast prototyping but is also suitable for **large-scale experiments** through its inbuilt support for optimized numerical libraries and parallelization. **Code profiling**, **domain visualizations**, and **data analysis** are integrated in a self-contained package available under the Modified BSD License. All these properties allow users to compare various RL algorithms with **little effort**.

## Problem

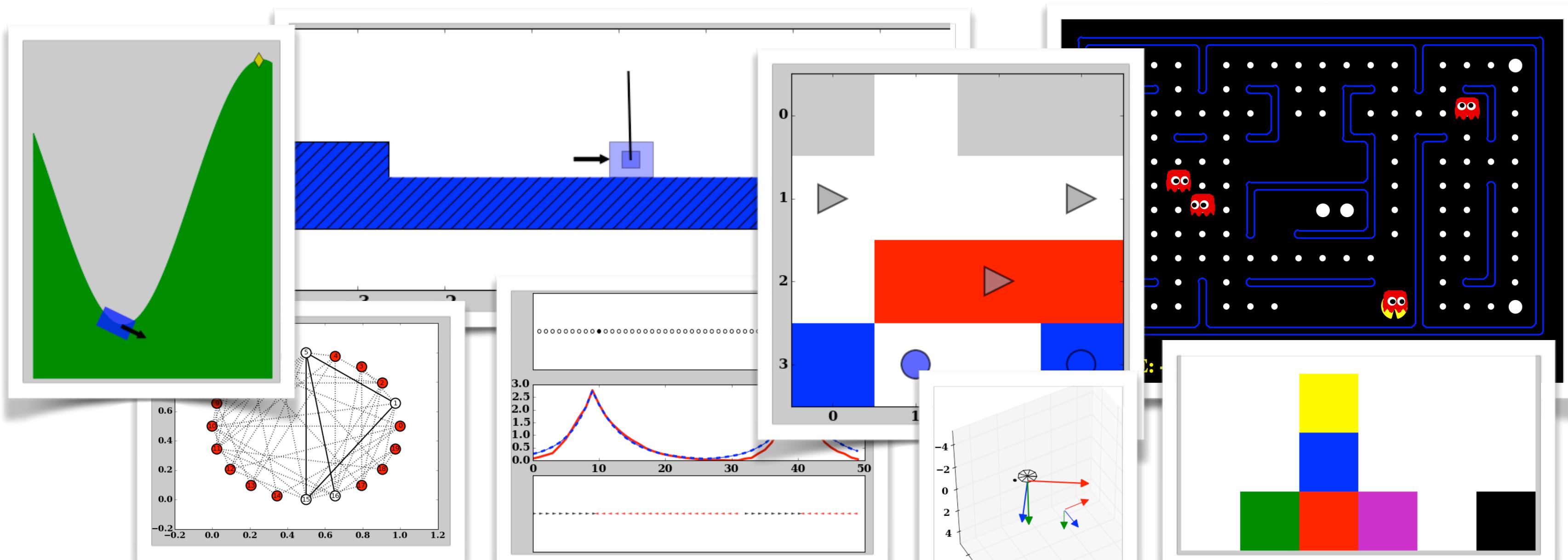
- An easy to use RL framework for both **research** and **education**



- Interest in Value-based RL with **linear function approximation** [1-4]

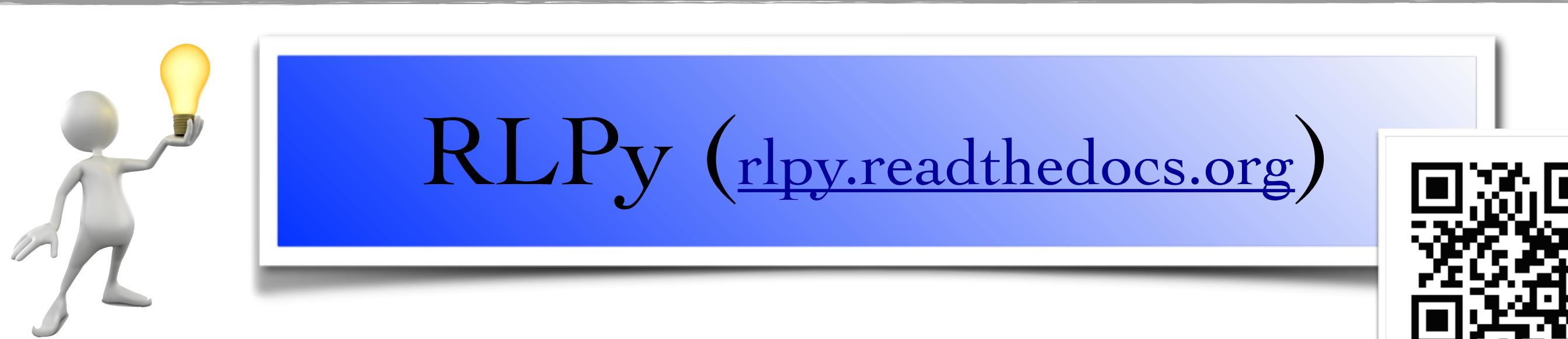
- Increased Specialization in RL → More Granular Framework [3-4]

- Comparison with State-of-the-art Techniques in **Various Domains**



## Existing Gap

- Lack of **granularity** to accommodate recent advances in RL [6]
- Challenging for entry level due to programming languages (e.g. C++) [6-7]
- Not **Self-contained** [5]



RLPy ([rlpy.readthedocs.org](http://rlpy.readthedocs.org))

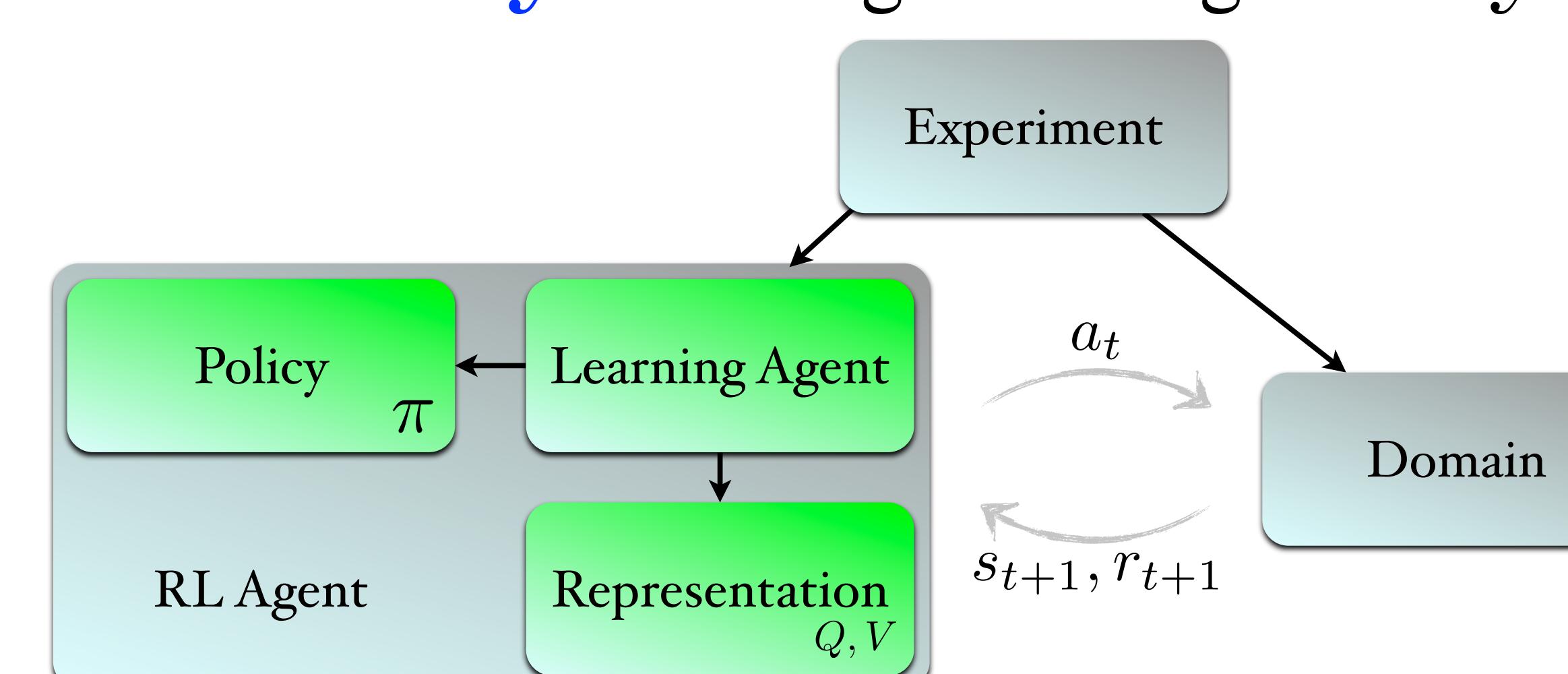
### Easy Installation

> pip install -U RLPy

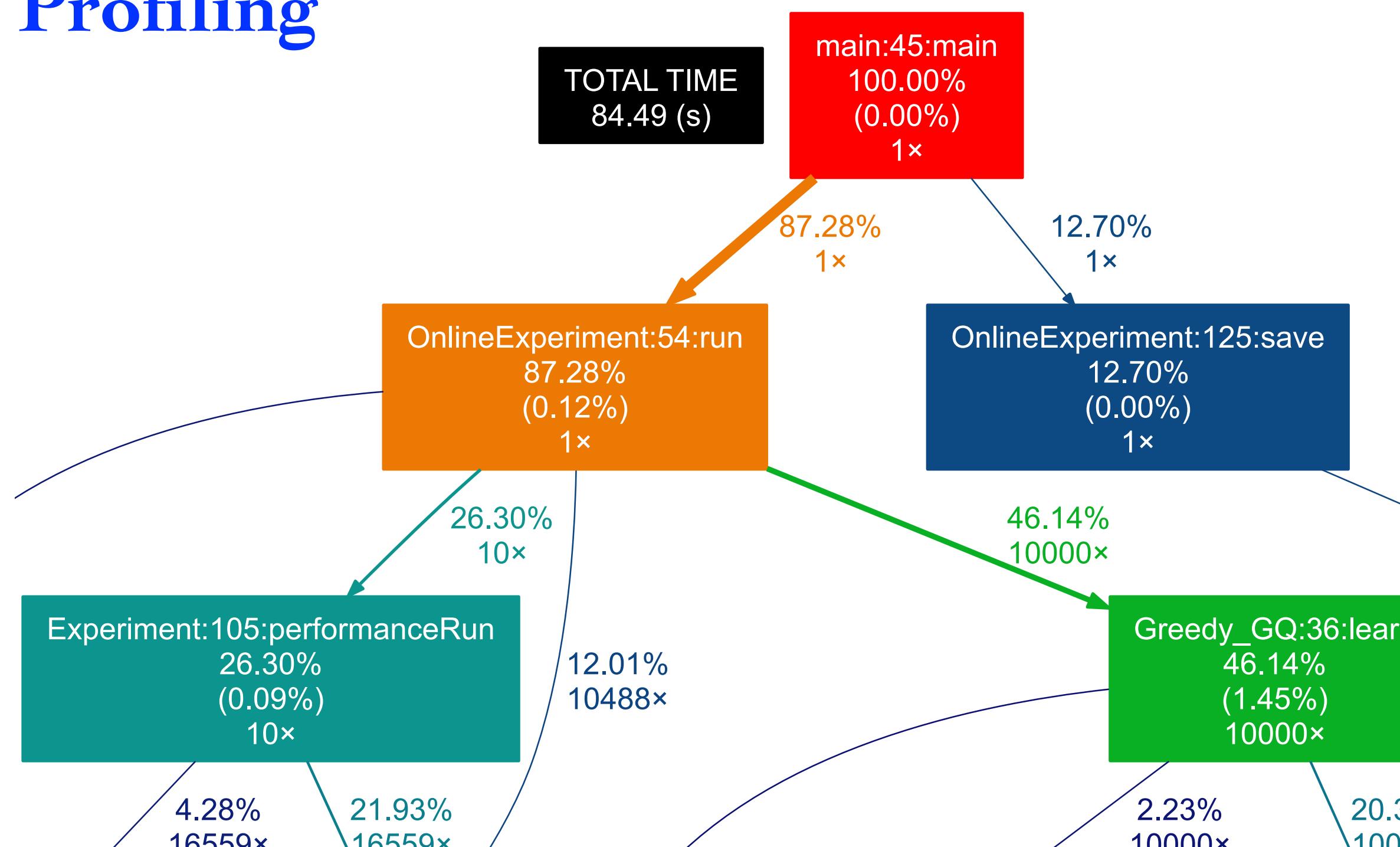
### Rapid Prototyping in Python

```
import rlpy
#### Domain ####
domain = rlpy.Domains.InfCartPoleBalance()
#### Agent ####
representation = rlpy.Representations.Tabular(domain, discretization=20)
policy = rlpy.Policies.eGreedy(representation, epsilon=0.1)
agent = rlpy.Agents.SARSA(policy, representation, domain.discount_factor)
#### Experiment ####
experiment = rlpy.Experiments.Experiment(agent, domain, max_steps=100)
experiment.run()
experiment.save()
```

### Improved Granularity of the agent using OO Python



### Built-in Profiling



### Built-in Hyperparameter Optimization

### Improved Experimentation

### Reproducible

### Parallel Execution

### Optimized Implementation (Cython, C++)

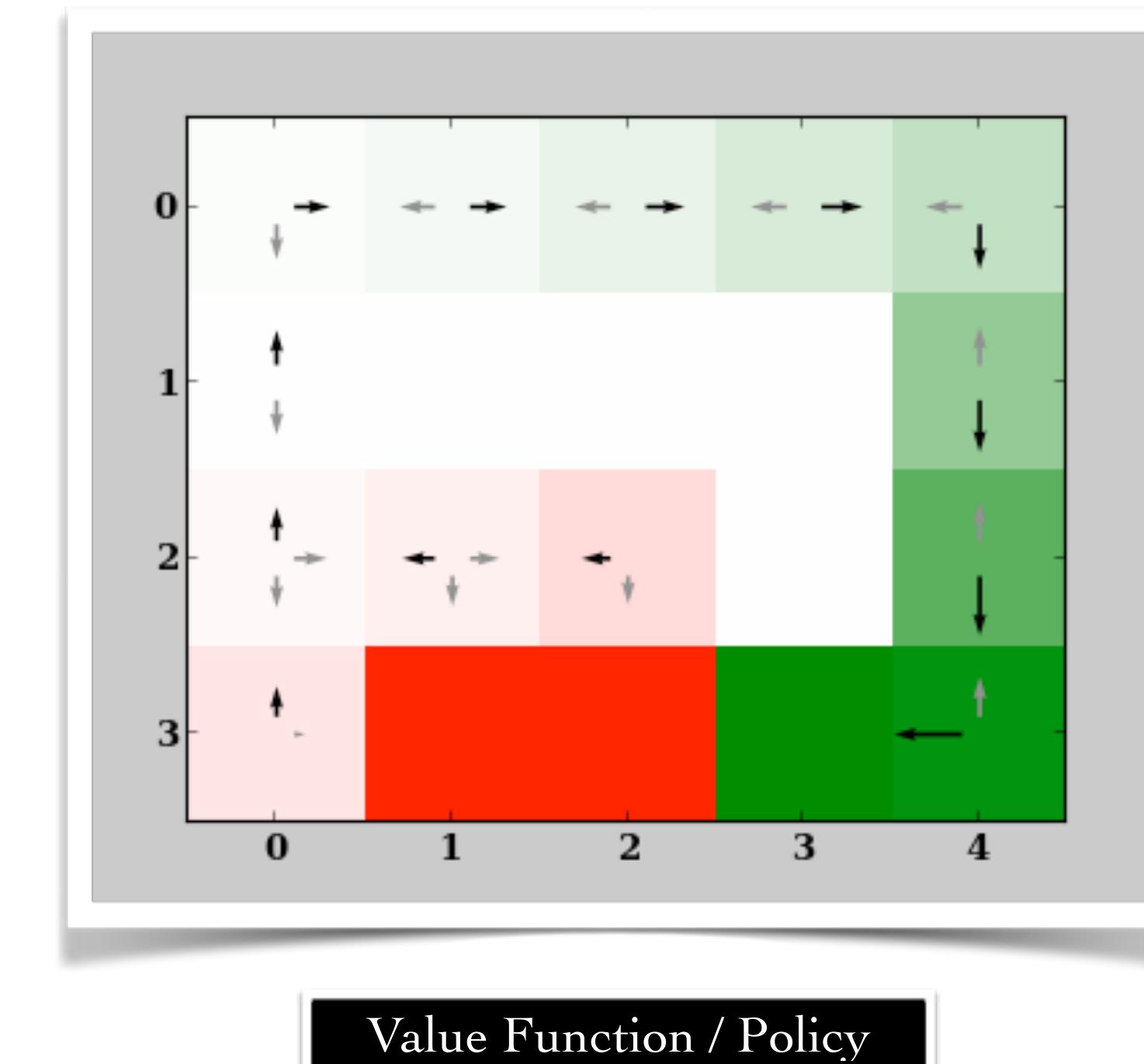
### Batteries Included!

### 20 Domains, 8 Learning Algorithms

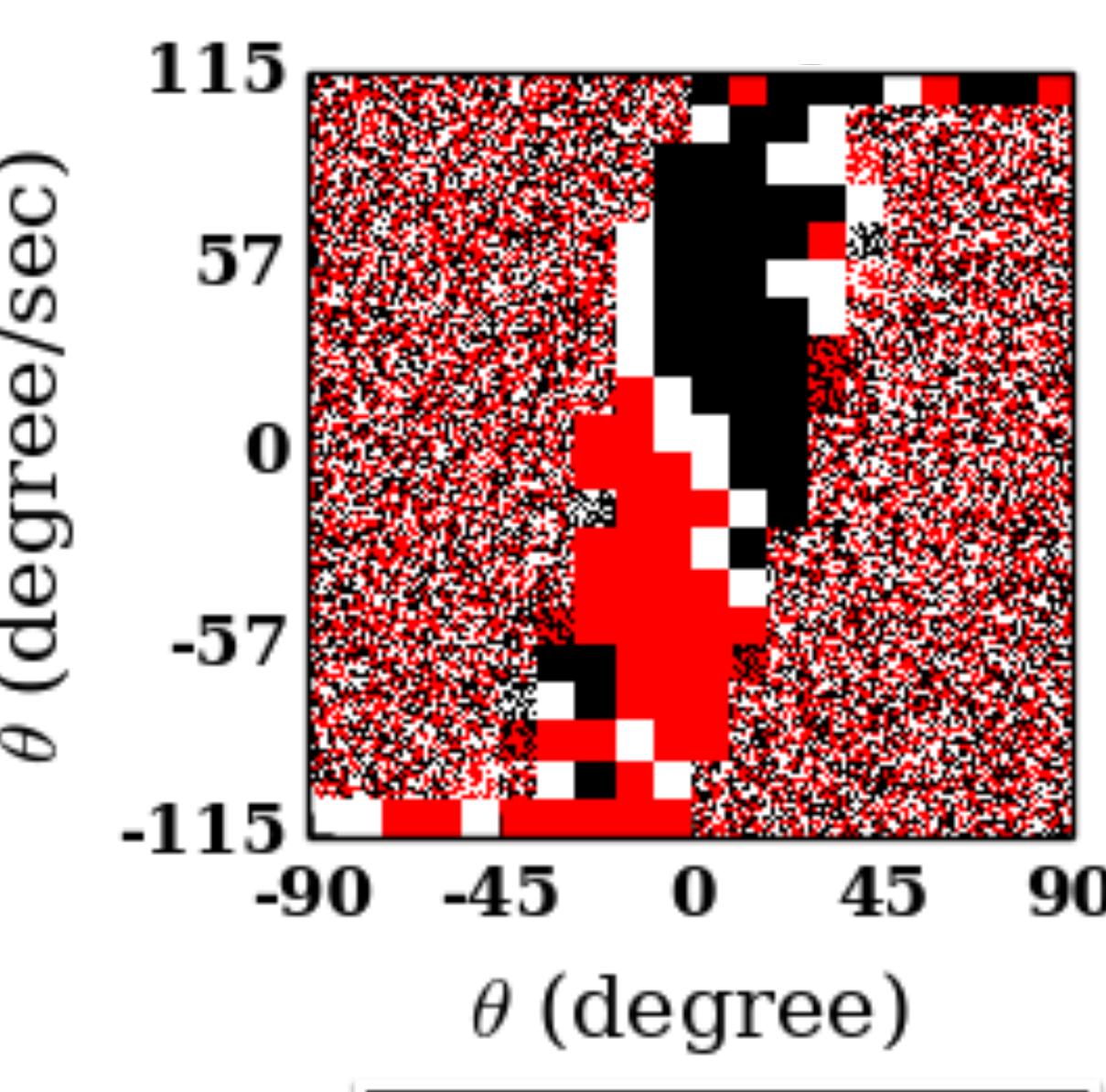
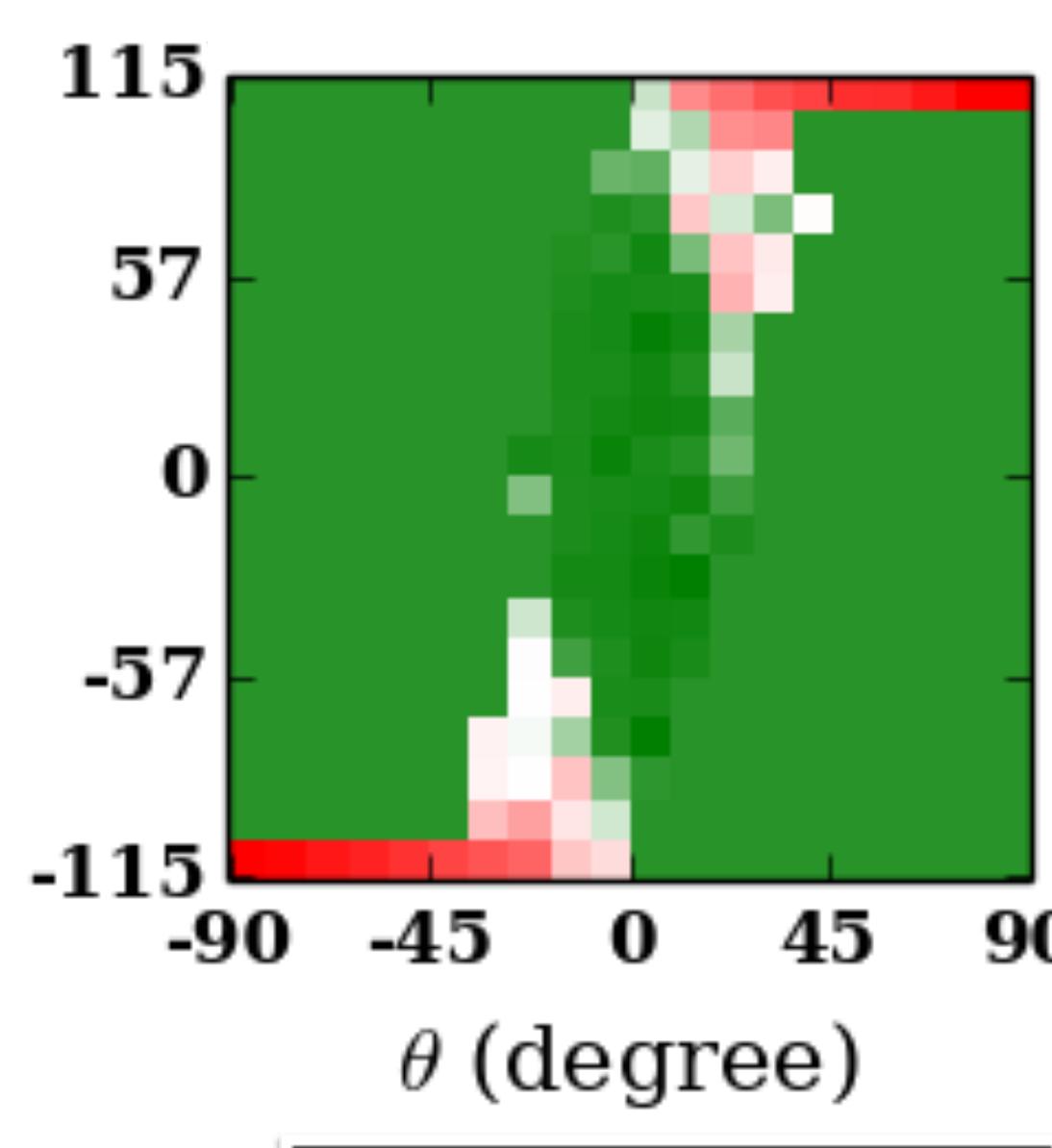
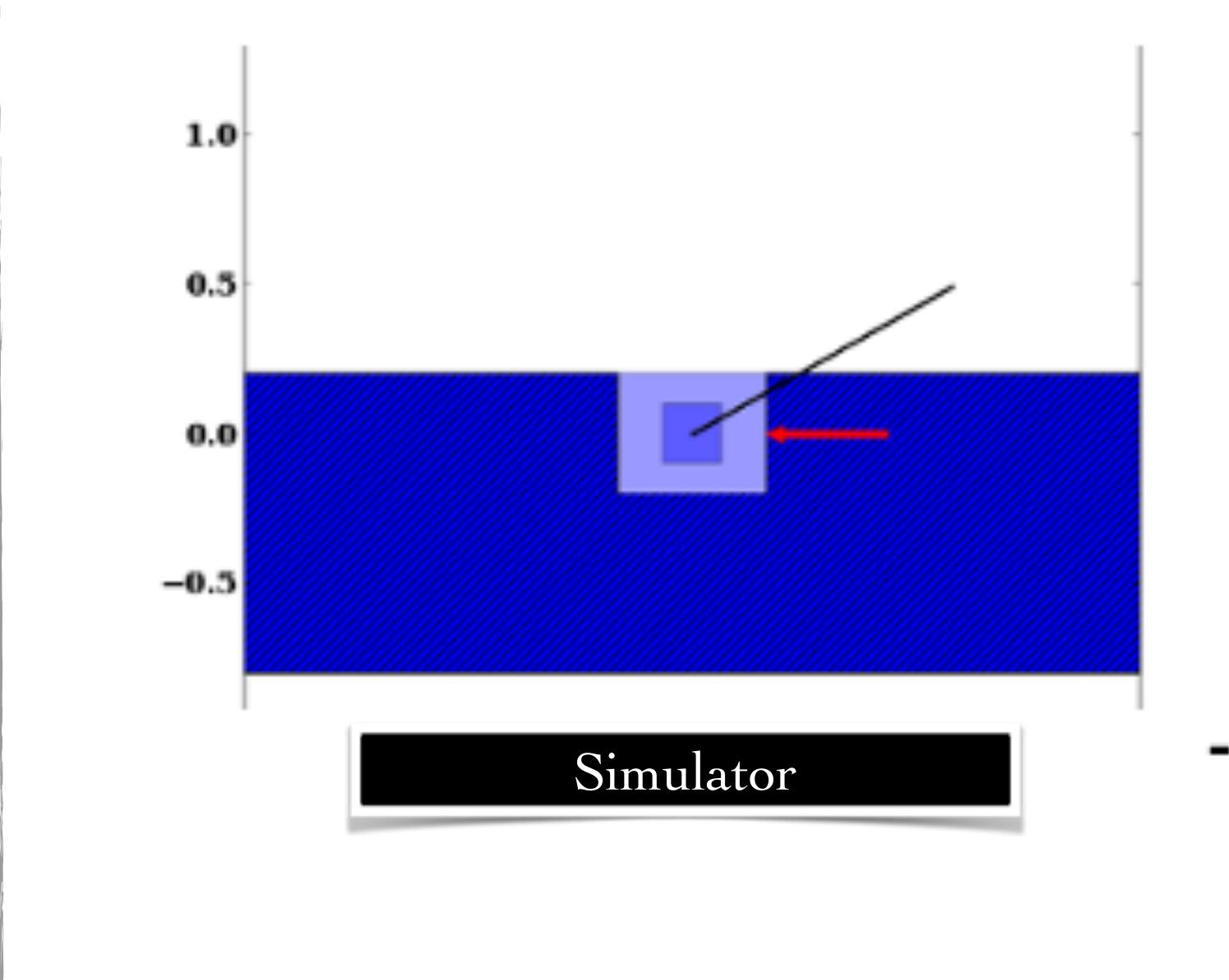
### 4 Policies, 7 Representations

## Examples

### Grid World

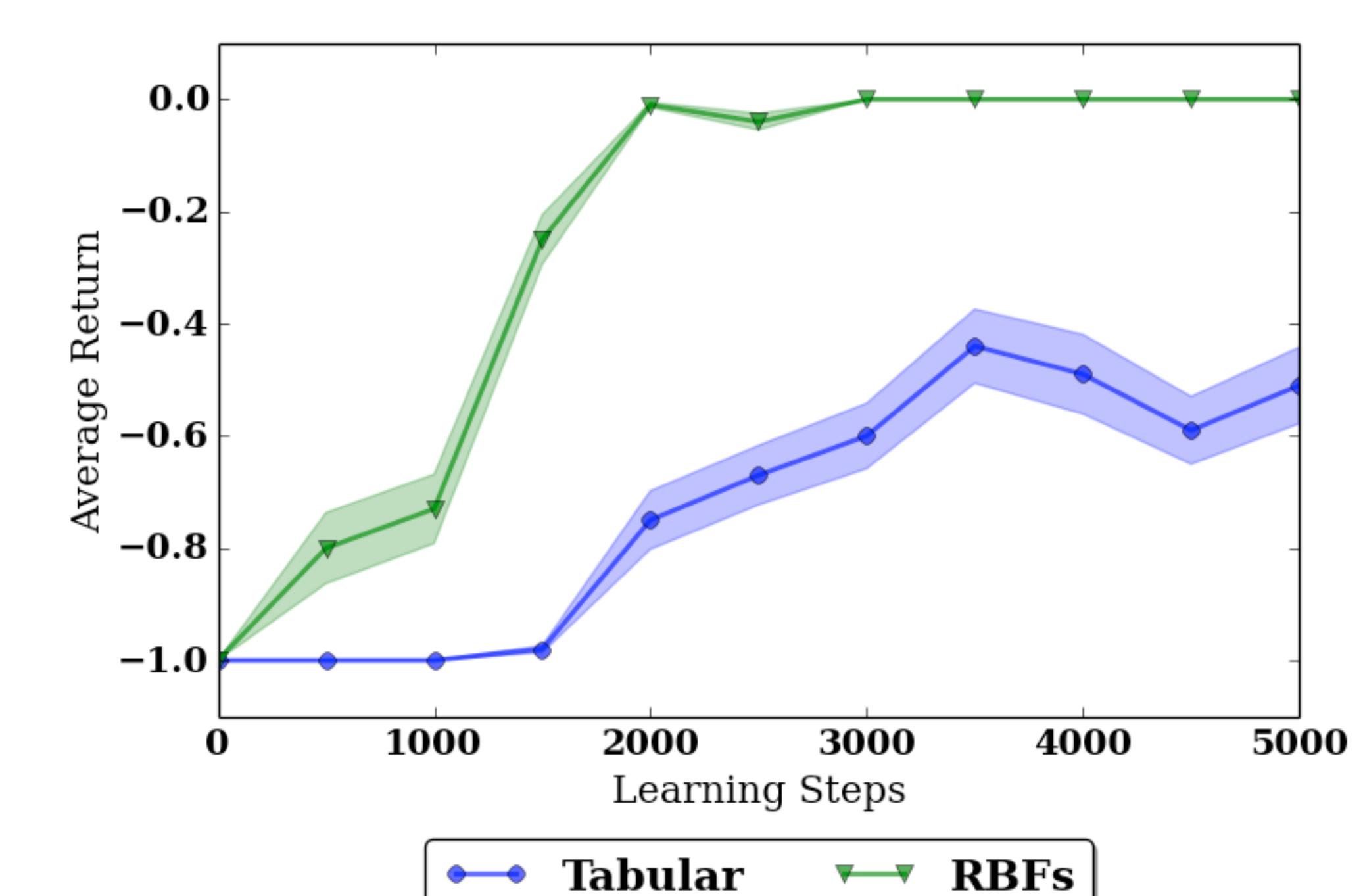


### Cart Pole



### Plotting

- Learning **Steps**
- Learning **Episodes**
- Learning **Time**
- Return**
- Number of **Features**
- Termination**



## Conclusion

- RLPy is a new **python based open-source** RL framework.
- Simplifies the construction of new RL ideas
- Accessible for both **novice** and **expert** users
- Realizes **reproducible** experiments

[1] A. Tamar, H. Xu, S. Mannor, "Scaling Up Approximate Value Iteration with Options: Better Policies with Fewer Iterations", In Proceedings of the 31st International Conference on Machine Learning (ICML), pages 127-135, 2014

[2] D. Calandriello, A. Lazaric, M. Restelli, "Sparse Multi-Task Reinforcement Learning", In Proceedings of Advances in Neural Information Processing Systems (NIPS), pages 819-827, Quebec, Canada, 2014

[3] J. Z. Kolter and A. Y. Ng, "Regularization and Feature Selection in Least-squares Temporal Difference Learning", In Proceedings of the 26th Annual International Conference on Machine Learning (ICML), pages 521-528, New York, NY, USA, 2009

[4] A. Geramifard, T. Walsh, N. Roy, and J. How "Batch iFDD: A Scalable Matching Pursuit Algorithm for Solving MDPs", The Conference on Uncertainty in Artificial Intelligence (UAI), 2013

[5] D. Aberdeen, "LibPGRL: A high performance reinforcement learning library in C++", 2007. URL <https://code.google.com/p/libpgrl/>.

[6] F. de Comite, "PIQUE: A Platform for Implementation of Q-Learning Experiments", 2006. URL <http://pique.sourceforge.net>.

[7] G. Neumann, "The Reinforcement Learning Toolbox , Reinforcement Learning for Optimal Control Tasks". MSc thesis, TU Graz, 2005

\* Author is currently employed at Amazon.