Temporal Difference Learning

Introduction

This repository consists of implementations of Temporal Difference (TD) Learning methods in the deep learning setting. Algorithms are implemented in `PyTorch` on a range of problems. Custom toy problems are provided in the `MDPs` folder.

TD Algorithms

Following are the TD algorithms implemented-

Algorithm	Link	Implementation	Status
Sarsa	Sutton & Barto	`sarsa.py`	Works well
Q-Learning	Sutton & Barto	`qlearning.py`	✓
Expected Sarsa	Sutton & Barto	`expectedsarsa.py`	✓
Double Sarsa	Sutton & Barto	`doublesarsa.py`	✓
Double Q-Learning	Sutton & Barto	`doubleqlearning.py`	✓
Double Expected Sarsa	Sutton & Barto	`doubleexpectedsarsa.py`	✓

Custom Environments

Following is the list of custom toy environments-

Environment Name	Link	Implementation
Cyclic MDP	ESAC	link
One-state MDP	Sutton & Barto	link
One-state Gaussian MDP	Sutton & Barto	link

Usage

To run an implementation, use the following command-

```
python main.py --alg <ALGORITHM> --env <ENV> --num_steps <STEPS>
```

For example, to run Q-Learning on the CartPole-v0 environment for 10000 steps, use the following-

```
python main.py --alg QLearning --env CartPole-v0 --num_steps 10000
```

To view the full list of arguments run the following-

```
python main.py --help
```

Citation

If you find these implementations helpful then please cite the following-

```
@misc{karush17tdlearning,
  author = {Karush Suri},
  title = {Temporal Difference Learning},
  year = {2021},
  howpublished = {\url{https://github.com/karush17/Temporal-Difference-Learning}},
  note = {commit xxxxxxx}
}
```