xuwd11 / cs294-112_hws

Branch: master ▾ cs294-112_hws / hw3 / Create new file Upload files Find file History

xuwd11 complete hw3 Latest commit e2170bf on Feb 1

..		
data	complete hw3	5 months ago
old	update note	5 months ago
results	complete hw3	5 months ago
README.md	complete hw3	5 months ago
atari_wrappers.py	init	6 months ago
dqn.py	fix double q	5 months ago
dqn_utils.py	init	6 months ago
hw3_instructions.pdf	init	6 months ago
logz.py	init	6 months ago
lunar_lander.py	init	6 months ago
plot.py	complete hw3	5 months ago
plot_part1.py	add hw3 part 1	5 months ago
requirements.txt	init	6 months ago
run_11.sh	update hw3 part 1	5 months ago
run_12.sh	update hw3 part 1	5 months ago
run_13.sh	update hw3 part 1	5 months ago
run_14.sh	update hw3 part 1	5 months ago
run_21.sh	complete hw3	5 months ago
run_22.sh	complete hw3	5 months ago
run_dqn_atari.py	fix double q	5 months ago
train_ac_f18.py	complete hw3	5 months ago

README.md

CS294-112 HW 3: Q-Learning

Usage

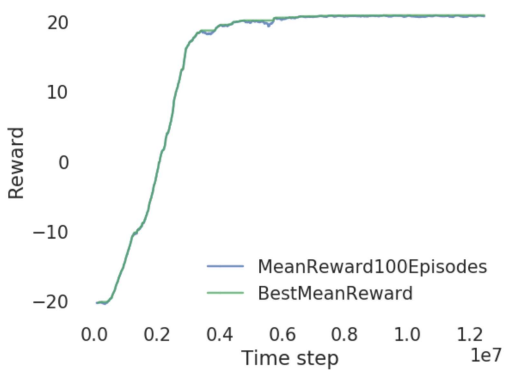
To run all experiments and plot figures for the report, run

```
bash run_11.sh
bash run_12.sh
bash run_13.sh
bash run_14.sh
python plot_part1.py
bash run_21.sh
bash run_22.sh
```

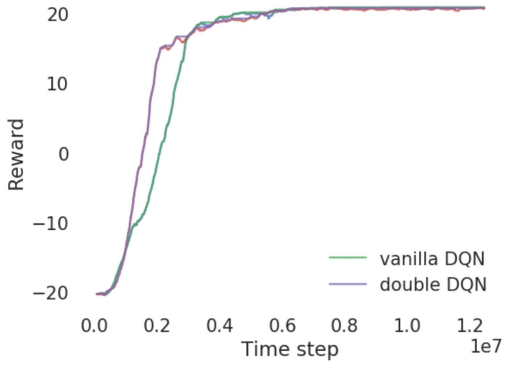
Results

Part 1

Question 1

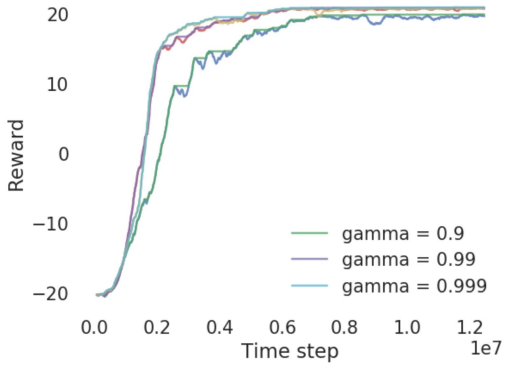


Question 2



Question 3

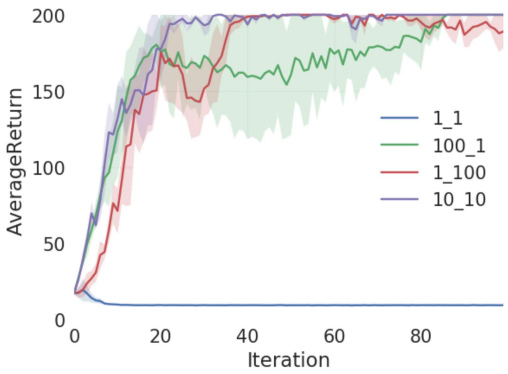
I experimented the effect of discount factor on performance.



As we can see, it takes longer to converge for small discount factor.

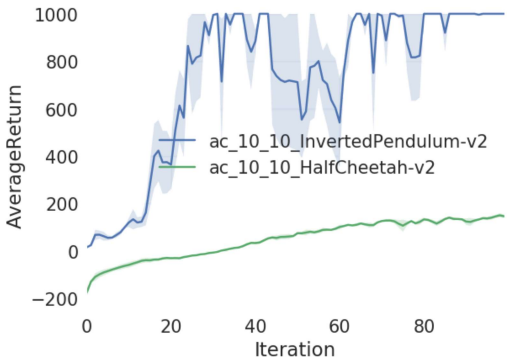
Part 2

Question 1



Setting both `num_grad_steps_per_target_update` and `num_target_updates` to 10 works best.

Question 2



Original README

Dependencies:

- Python **3.5**
- Numpy version **1.14.5**
- TensorFlow version **1.10.5**
- MuJoCo version **1.50** and mujoco-py **1.50.1.56**
- OpenAI Gym version **0.10.5**
- seaborn
- Box2D==**2.3.2**
- OpenCV
- ffmpeg

Before doing anything, first replace `gym/envs/box2d/lunar_lander.py` with the provided `lunar_lander.py` file.

The only files that you need to look at are `dqn.py` and `train_ac_f18.py`, which you will implement.

See the HW3 PDF for further instructions.

The starter code was based on an implementation of Q-learning for Atari generously provided by Szymon Sidor from OpenAI.