Ehsan Arabnezhad

Implementation description

Learning Algorithm

Multi DDPG agent algorithm with replay buffer is used in this project.

Configurations:
* 2 hidden layers with 64 and 64 hidden units for both actor and critic
* Replay batch size 128
* Buffer size 1e7
* Replay without prioritization
* Update frequency 1
* TAU from 1e-2
* Learning rate 1e-4 for actor and 1e-3 for critic
* Ornstein-Uhlenbeck noise

Plot of Rewards
Plot of rewards can be seen after the environment has been solved.
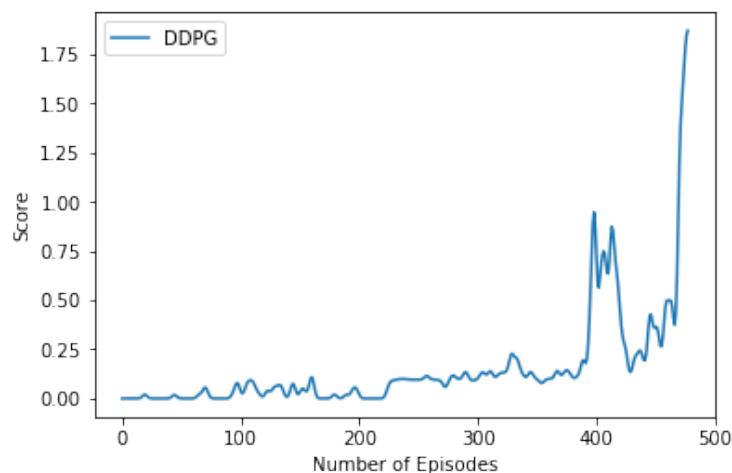
Environment solved in 378 episodes.

```
Episode 100      Average Score: 0.01      Current Score: 0.00
Episode 200      Average Score: 0.04      Current Score: 0.00
Episode 300      Average Score: 0.08      Current Score: 0.10
Episode 400      Average Score: 0.18      Current Score: 1.40
Episode 478      Average Score: 0.50      Current Score: 2.60
Environment solved in 378 episodes!      Average Score: 0.50
```

Ideas for Future Work

1. Build an agent that finds the best hyperparameters for an agent

2. Prioritization for replay buffer
3. Parameter space noise for better exploration
4. Test dropout for Critic