Project Report

Learning Algorithm

The following algorithms was used as part of the solution:

1 - Deep Q Learning: https://storage.googleapis.com/deepmind-media/dqn/DQNNaturePaper.pdf

2 - Double Q Learning: https://arxiv.org/pdf/1509.06461.pdf

3 - Dueling Network Architectures: https://arxiv.org/pdf/1511.06581.pdf

The architecture used was:

1 - Fully connected layer: (input: 37, output: 64)

2 - Fully connected layer: (input: 64, output: 64)

3 - Value layer: (input: 64, output: 1) - Dueling Network

4 - Advantage Layer: (input: 64, output: 4) - Dueling Network

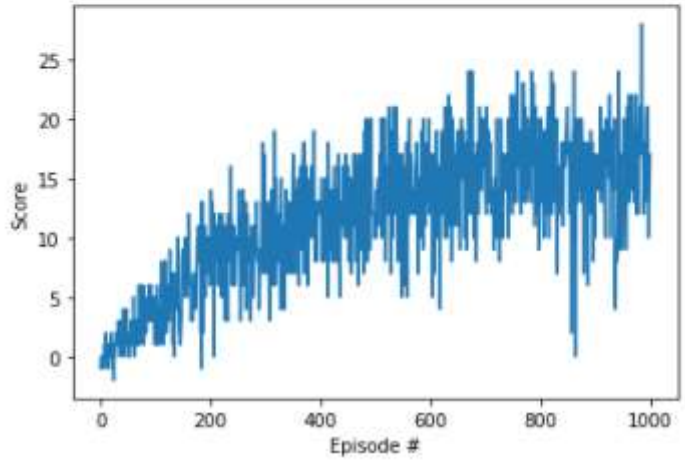
Hyperparameters

Number of episodes: 1000
Episilon Start: 1.0
Min Episilon: 0.01
Episilon Decay: 0.995
Replay Buffer Size: 10000
Bach Size for Learn: 64
Gamma (discount factor): 0.99
Learning rate: 5e-4

Plot of Rewards

```
Episode 100     Average Score: 1.91
Episode 200     Average Score: 5.86
Episode 300     Average Score: 8.95
Episode 400     Average Score: 11.20
Episode 500     Average Score: 12.30
Episode 600     Average Score: 14.09
Episode 700     Average Score: 15.44
Episode 800     Average Score: 16.08
Episode 900     Average Score: 15.60
Episode 1000    Average Score: 16.08

Environment solved in 424 episodes!
```



Ideas for future work

1 – Parameterize the options to use “Dueling Network” and “Double Q Learning”

2 – Implement “Prioritize Experience Replay”