Navigation

I used Dueling Q-Network to train the agent so agent can select best action against each state.

Why Dueling Q Network:

Dueling network represents two separate estimators: one for the state value function and one for the state-dependent action advantage function. The main benefit of this factoring is to generalize learning across actions without imposing any change to the underlying reinforcement learning algorithm.

Reference Used:

1. https://arxiv.org/abs/1511.06581
2. https://arxiv.org/pdf/1511.06581.pdf

Summary Of Dueling Q Network:

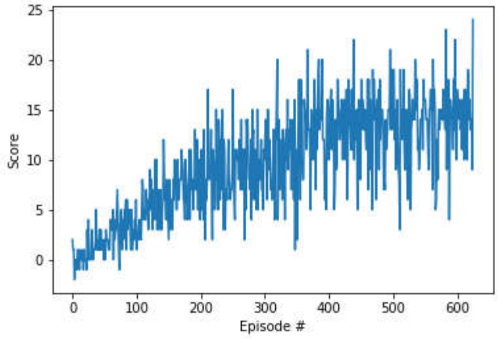
- **Fully-connected layer :**
    a. input: 37 (state size)
    b. output: 64
- **Fully-connected layer :**
    a. input: 128 output 64 for approximation state-dependent action advantage function.
    b. input: 128 output 64 for approximation state value function.
- **Fully-connected layer :**
    a. input: 64 output: (action size) for approximation state-dependent action advantage function.
    b. input: 64 output: 1 for approximation state value function.
- **Maximum steps per episode:** 1000
- **Starting epsilon:** 1.0
- **Ending epsilon:** 0.01
- **Epsilon decay rate:** 0.995

Rewards Function Performance During Training:

- **Average Score after every 100 episodes:**

```
Episode 100	Average Score: 1.93
Episode 200	Average Score: 6.26
Episode 300	Average Score: 8.99
Episode 400	Average Score: 11.50
Episode 500	Average Score: 13.18
Episode 600	Average Score: 13.54
Episode 625	Average Score: 14.07
Environment solved in 525 episodes!	Average Score: 14.07
```

- **Plot shows average rewards against each episode:**



Rewards Function Performance During Prediction:

```
Episode: 98     Score: 12.755102040816327
Episode: 99     Score: 12.757575757575758
Episode: 100     Score: 12.8
Average Score: 12.8
```

Networks want to try in future:

1. Double Deep Q Networks with Prioritized Experience Replay
2. Rainbow