

Banana Collector

Learning Algorithm

It is based on Deep Q-Learning(DQN) with the following improvements

- Experience Replay
- Fixed Q-Target
- Double Deep Q-Learning

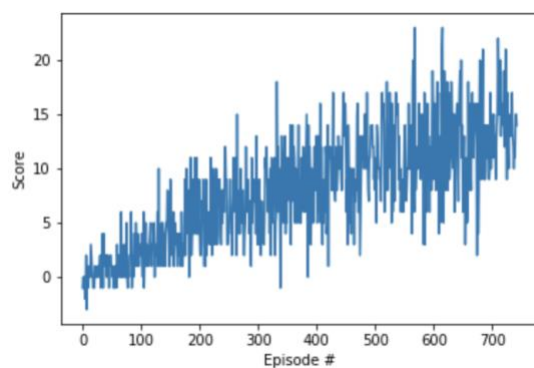
Hyperparameters:

- Size of memory = 100000
- Batch size = 64
- Discount Factor (GAMMA) = 0.99
- Rate of transfer from local to target network (TAU) = 0.1
- Frequency of update of local network = 4
- Frequency of update of target network = 4
- GLIE – Epsilon Initialization = 1
- GLIE – Epsilon Decay = 0.995
- GLIE – Minimum Epsilon = 0.01

Model Architecture:

Input(37)
Fully Connected Layer (128)
RELU
Fully Connected Layer (64)
RELU
Fully Connected Layer (32)
RELU
Output(4)

Plot of Rewards



Ideas for future work

I would like to try out Dueling Network and Prioritized Experience Replay as improvements over DQN.