Project 1: Capturing banana using

1 Introduction

In this assignment, we are asked to train an agent that can pick yellow banana and avoide picking blue banana. The state space is represented with 37 dimensions and action space with 4 discrete values: 0 - move forward.
1 - move backward.
2 - turn left.
3 - turn right.

The goal of an agent is to predict the best action $\in R^4$ given the state vector $\in R^37$ as input.

2 Method

In order to find best action given the state, we used dueling deep q-network (dueling DQN), with experience replay and using two networks for action computing and evaluation.

2.1 Network Architecture

Two neural networks with two hidden layers (64 and 128 neurons) and rectified linear unit has been used as an activation. Out layer is sum of value layer and normalized advantage layer.

Hyper-parameters used for dueling DQN:

- Maximum steps per episode: 1000
- Starting ϵ: 1.0
- Ending ϵ: 0.01
- ϵ decay rate: 0.999

3 Results

The agent is able to achieve thirteen an average score in 467 number of episodes. Figure 1 shows the mean score achieved by the agent with number of episodes.

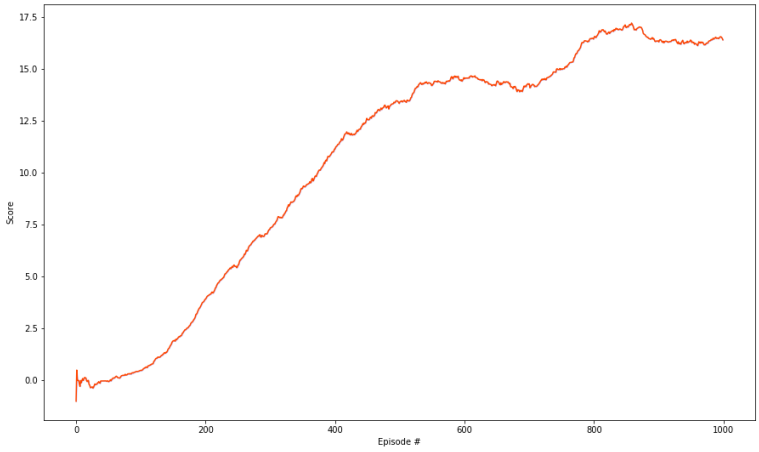

Figure 1: Progress of learning agent

4 Future work

It would be better to try other algorithms such as RAINBOW and actor-critic. Moreover, importance sampling with experience replay might train an agent faster.