Collaboration and Competition Project Borja González
Deep Reinforcement Learning Nanodegree January 14th, 2018

Overview

The goal of this report I briefly summarize the learnings and final modeling decisions taken as part of the Collaboration and Competition project.

Neural Network Architecture

This project was solved using the same architecture as the one for Continuous Control project. Both, actor and critic used two hidden layers with 256 neurons and batch normalization after the first layer. The actor used ReLU as activation function while the Critic used Leaky ReLU.

Further Improvements

From my experience with the previous project I reduced the number of episodes for learning and changed the learning rates. I also increased the batch size.

Choice of hyperparameters

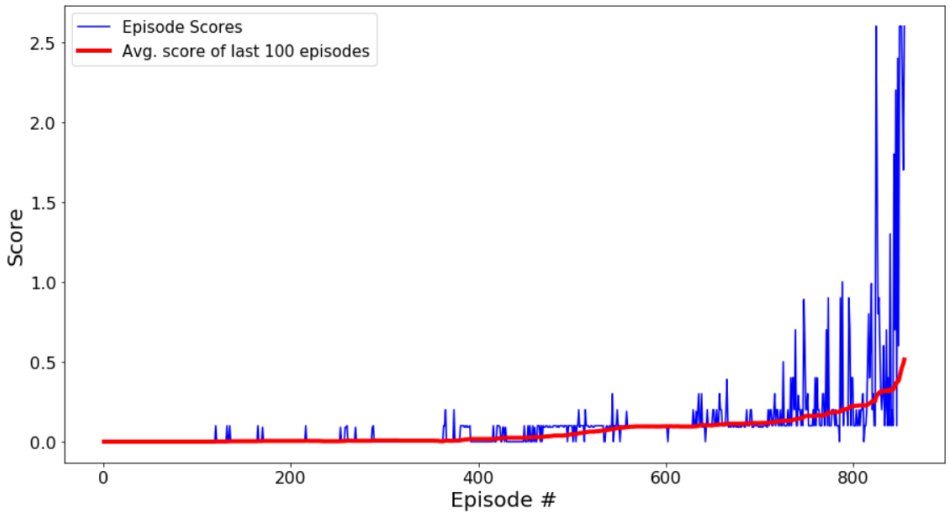
Final choice of hyperparameters was:

```
BUFFER_SIZE = int(1e6)          # replay buffer size
BATCH_SIZE = 256                # minibatch size
GAMMA = 0.99                    # discount factor
TAU = 1e-3                      # for soft update of target parameters
LR_ACTOR = 1e-4                 # learning rate of the actor
LR_CRITIC = 3e-4                # learning rate of the critic
WEIGHT_DECAY = 0                # L2 weight decay
N_LEARN_UPDATES = 10            # number of learning updates
N_TIME_STEPS = 5        # every n time step do update
```

Results

The agent was able to solve the environment after 854 episodes, which means after 170 learning-step updates. This was a great performance compared with Udacity's reference which needed more than 1500 episodes.

```
Episode 200     Average Score: 0.005    Score: 0.00
Episode 400     Average Score: 0.017    Score: 0.00
Episode 600     Average Score: 0.096    Score: 0.09
Episode 800     Average Score: 0.224    Score: 0.10
Episode 854     Average Score: 0.514    Score: 2.60
Environment solved in 854 episodes!     Average Score: 0.514
```



Further Work

First, I will work with the soccer environment. Then, I will try with different algorithms such as A2C or PPO.