# DRLND Project 3: Collab-Compete

michaelfitzke

August 2019

## 1 Overview

This project utilised the architecture for solving the DDPG-Bipedal Udacity project repo. coding exercise as outlined in the solution provided by udacity.

## 2 State and Action Space

The State Space consists of 8 Variables for each Agent. The actions space consits of 2 continous variables.

## 3 Model

We use the DDPG algorithm following [1].

### 3.1 Hyperparameters

in a complete competitive manner. I.e. both agents don't share infromation.

The environment was solved with the following hyperparameters

- $BUFFER_{SIZE} = int(1e5)$
- $BATCH_{SIZE} = 256$
- $\gamma = 0.99$
- $\tau = 1e-2$
- $LR_{ACTOR} = 1e-3$
- $LR_{CRITIC} = 1e-3$
- $WEIGHT_{DECAY} = 0.0$

## 4 Result

An average score over 0.5 was reached in 887 episodes.

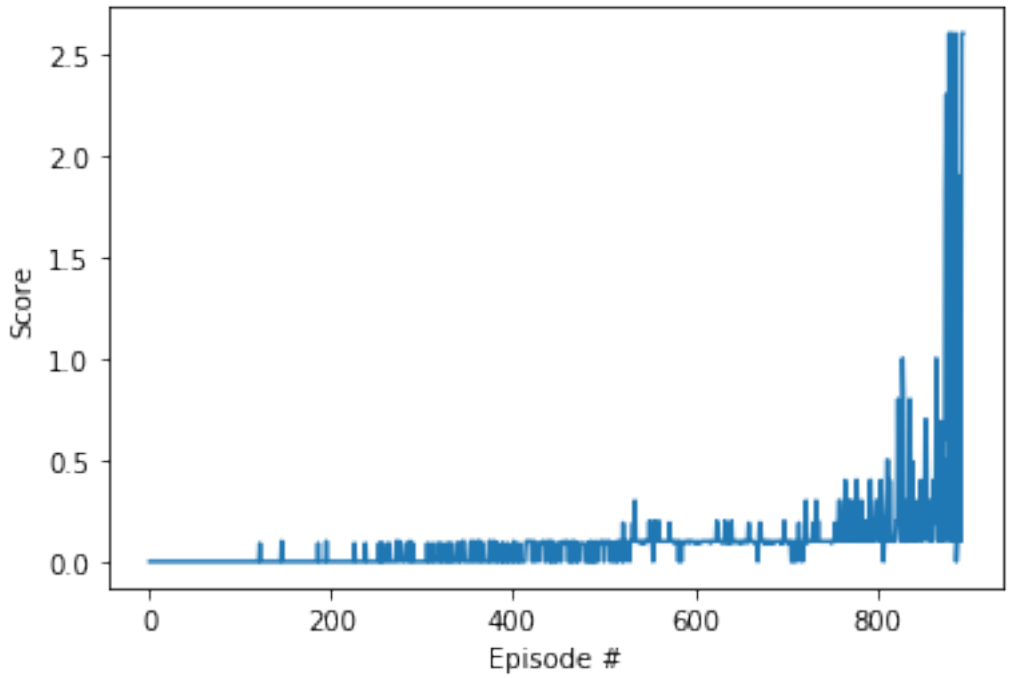


Figure 1: Learning

## 5 Ideas for making this better

Colaborative strategies should be explored, such as the ones discussed in the MARL literature.

## References

[1] Timothy P Lillicrap, Jonathan J Hunt, Alexander Pritzel, Nicolas Heess, Tom Erez, Yuval Tassa, David Silver, and Daan Wierstra. Continuous control with deep reinforcement learning. *arXiv preprint arXiv:1509.02971*, 2015.