

# Report

July 9, 2019

Deep Reinforcement Learning Project 2 - Continuous Control Problem

## 1 Learning Algorithm

The learning algorithm used to train this agent was DDPG. This is an ActorCritic Algorithm. ActorCritic algorithms takes the best of Value-Based and Policy-Based algorithms into consideration.

## 2 DDPG Agent with batch normalization

- Agents Hyper Parameter:

```
BUFFER_SIZE = int(1e6)      # replay buffersize
BATCH_SIZE = 2*10           # mini batchsize
GAMMA = 0.99                 # discount factor
TAU = 1e-3                  # for softupdate of target parameters
LR_ACTOR = 1e-3              # learningrate of the actor
LR_CRITIC = 1e-3             # learning rate of the critic
WEIGHT_DECAY = 0.000         # L2weightdecay
NET_UPDATE_EVERY = 20        # Update the network policy every
N_EPISODES = 1000            # Number of total episodes to train the agent
MAX_T = 1000                 # Max time steps in every episode
```

- ModelStructureActor:

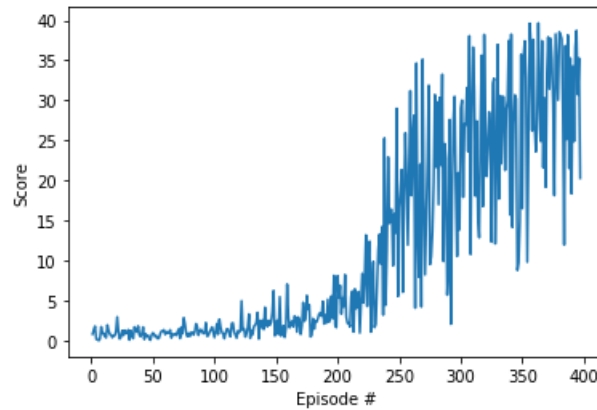
Actor:

```
state_size--> BN0--> fc1--> BN1--> ReLU--> fc2--> BN2--> ReLU--> fc3--> tanh
--> action_size

#fc1_units = 400
# fc2_units = 300
```

Critic:

```
state_size--> BN0--> fc1--> BN1--> ReLU--> fc2--> BN2--> ReLU--> fc3
--> state-valuefunction
```



ScoresPlot.png

### 3 Plot of Rewards:

Episode 50 Average Score: 0.87

Episode 100 Average Score: 1.03

Episode 150 Average Score: 1.71

Episode 200 Average Score: 2.91

Episode 250 Average Score: 8.49

Episode 300 Average Score: 19.43

Episode 350 Average Score: 25.16

Episode 397 Average Score: 30.08

### 4 Ideas for Future Work

It would be interesting to try the following modidifications:

Training the agent with PP02

Transfer the trained network on Multiple Agents

Training Multiple Agents Version