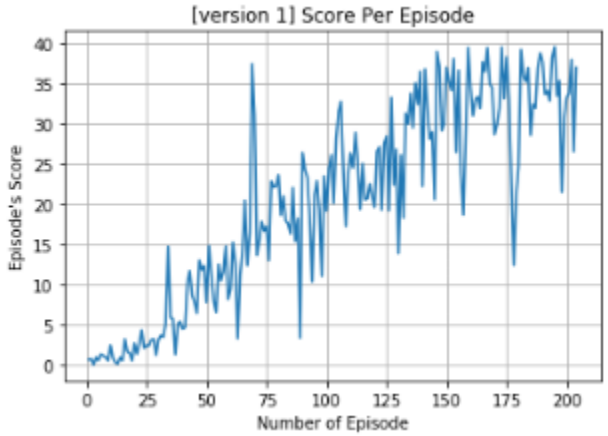
UDACITY DEEP REINFORCMENT LEARNING – PROJECT 2 REPORT "CONTINUOUS CONTROL RL"

Mahmoud Adel Nour

m.adel.nour@gmail.com

Report

1. Learning Algorithm
    - Building DDPG network using 3 hidden full connected layers for the actor and critic, with ReLu activations and the optimizer is ADAM
    - Using batch normalization for better accuracy
    - Using Experience replay and review every 128 random replays and check every 4 record if we reached the 64 replays to do the learning
    - Using different values of gamma
    - Creating target weights and current(local) weights knowing that the target weights changes with small steps from the current weights to keep soft update concept
    - I tried to use add on noise and it gives better values without noise
    - The current implementation achieved mean score over 100 episodes greater than 30 after 204 episodes of training
    - Hyper parameters:
        - BUFFER_SIZE = 100000 : replay buffer size
        - BATCH_SIZE = 128 : batch size to start learning over the stored replays
        - GAMMA = 0.99 : discount factor
        - TAU = 0.001 : soft update factor which control the change of the weights
        - LR_ACTOR = 0.0002 : learning rate for the optimizer “ADAM” of the Actor
        - LR_ACTOR = 0.0002 : learning rate for the optimizer “ADAM” of the Critic
2. Plot of Rewards



3. Ideas for Future Work
    - to use more complicated algorithm like A3C, A2C and GAE