REPORT

Model

A simple default model was used:
Actor and critics both use 3 linear layer
The first hidden layer has 256 units and the second hidden layer has 128 units (this architecture played an important role on how well the model can perform)

LR_ACTOR = 1e-4 # learning rate of the actor
LR_CRITIC = 2e-4 # learning rate of the critic
WEIGHT_DECAY = 0. # L2 weight decay
GAMMA = 0.99 # discount factor
TAU = 1e-3 # for soft update of target parameters
UPDATE_EVERY = 1
UPDATE_COUNT = 1
BATCH_SIZE = 128
BUFFER_SIZE = 1e5

Applying:

UPDATE_EVERY 1:
It seems that Using update_every cause some inconsistency or worsen perforamcne this can be found in the following thread

noise decay during agent.act on top of OUNoise

Learning algorithm:
DDPG: Deep deterministic Policy gradient was used:kind of actor critic method, an approximate Dqn (allow to train in continuous space)
, however in this case I used 2 agents to train.
Actor - critic architecture were used to perform this task

Additional highlight:

Future for improvement
Applying MADDPG rather than getting away with training loop
Applying PER: (Prioritized Experience Replay) to utilize the efficiency and effectiveness of experience replay

Improving Neural network and hyper tuning hyperparameter:

Applying TAU = 1e-2
Source:
https://knowledge.udacity.com/questions/888642
https://knowledge.udacity.com/questions/65068

Plot of reward

```
Episode 100     Average Score: 0.001000000      Max Score: 0.100000001  Max Score added: 0.000000000
Episode 200     Average Score: 0.017000000      Max Score: 0.100000001  Max Score added: 0.000000000
Episode 300     Average Score: 0.021400000      Max Score: 0.100000001  Max Score added: 0.090000002
Episode 400     Average Score: 0.047500001      Max Score: 0.200000003  Max Score added: 0.090000002
Episode 500     Average Score: 0.052800001      Max Score: 0.100000001  Max Score added: 0.090000002
Episode 600     Average Score: 0.078400001      Max Score: 0.100000001  Max Score added: 0.100000001
Episode 700     Average Score: 0.082400001      Max Score: 0.200000003  Max Score added: 0.100000001
Episode 800     Average Score: 0.121700002      Max Score: 0.400000006  Max Score added: 0.100000001
Episode 900     Average Score: 0.164400002      Max Score: 0.500000007  Max Score added: 0.100000001
Episode 979     Average Score: 0.501900008      Max Score: 2.700000040  Max Score added: 1.100000016
Environment solved in 879 episodes!     Average Score: 0.50
```