

DDPG Method for Solving Udacity's Continuous Control Problem-20 Agents

1) DDPG Algorithm

We implemented an *off-policy method* called Deep Deterministic Policy Gradient (DDPG) described in the paper *Continuous control with deep reinforcement learning* (<https://arxiv.org/abs/1509.02971>).

DDPG learns a Q-function and a policy in training. It uses off-policy data and the Bellman equation to learn the Q-function, and uses the Q-function to learn the policy.

The DDP folder holds 2 '.py' files that are then called in the [Continuous Control-20Agents-Final.ipynb](#) notebook:

-model.py : Implement the Actor and the Critic class:

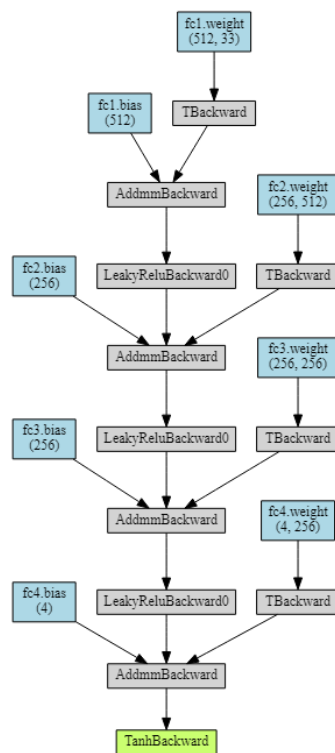
Both Actor and Critic class implement a Target and a Local Neural Network for training.

-ddpg_agent.py : Implement the DDPG agent, a Noise and a Replay Buffer class:

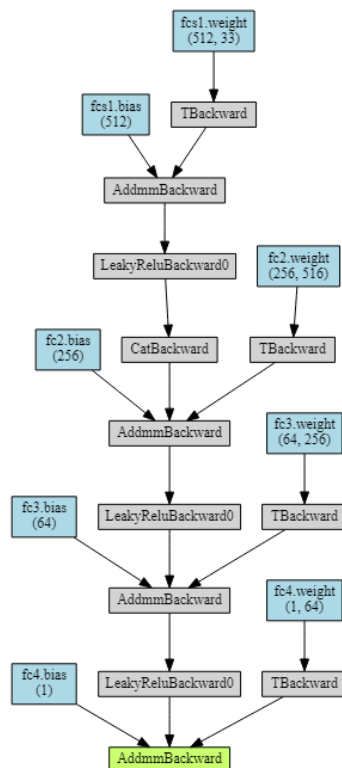
The Actor's Local and Target neural networks, and the Critic's Local and Target neural networks, the Noise process and the Replay Buffer are inside by the Agent's constructor and reset independently for each agent. The Noise uses OU process. The Replay Buffer is a fixed-size buffer to store experience tuples.

2) Model Used

The Actor Neural Network



The Critic Neural Network



A few notes:

-There was a need to go deeper in both networks. Learning was not happening without the additional layers from the original proposed networks.

-The use of Leaky_ReLU helps in the continuous control that can take negative values.

2) Agent Used

Here are the Hyper parameters used.

```
BUFFER_SIZE=int(1e5)
# replay buffer size

BATCH_SIZE = 128      # minibatch size
GAMMA = 0.99          # discount factor
TAU = 0.001           # for soft update of target parameters
LR_ACTOR = 1e-4        # learning rate of the actor
LR_CRITIC = 1e-3       # learning rate of the critic
WEIGHT_DECAY = 1e-5    # L2 weight decay
```

Two conditions were added on the loop:

-A maximum number of steps per episode

-A break on the number of 'dones' was added.

3) Training and Results

Episode 5	Average Score: 0.44	Score: 0.456
Episode 10	Average Score: 0.44	Score: 0.390
Episode 15	Average Score: 0.45	Score: 0.415
Episode 20	Average Score: 0.43	Score: 0.189
Episode 25	Average Score: 0.44	Score: 0.492
Episode 30	Average Score: 0.49	Score: 0.712
Episode 35	Average Score: 0.52	Score: 0.656
Episode 40	Average Score: 0.56	Score: 1.117
Episode 45	Average Score: 0.62	Score: 1.202
Episode 50	Average Score: 0.81	Score: 3.008
Episode 55	Average Score: 1.07	Score: 2.947
Episode 60	Average Score: 1.38	Score: 6.191
Episode 65	Average Score: 1.85	Score: 10.471
Episode 70	Average Score: 2.85	Score: 20.054
Episode 75	Average Score: 4.53	Score: 29.184
Episode 80	Average Score: 6.35	Score: 33.382
Episode 85	Average Score: 8.03	Score: 35.137
Episode 90	Average Score: 9.48	Score: 32.967
Episode 95	Average Score: 10.78	Score: 34.932
Episode 100	Average Score: 11.98	Score: 35.046
Episode 105	Average Score: 13.64	Score: 31.652
Episode 110	Average Score: 15.34	Score: 34.253
Episode 115	Average Score: 16.97	Score: 32.760
Episode 120	Average Score: 18.62	Score: 35.675
Episode 125	Average Score: 20.35	Score: 35.723
Episode 130	Average Score: 22.00	Score: 31.389
Episode 135	Average Score: 23.65	Score: 34.645
Episode 140	Average Score: 25.35	Score: 36.556
Episode 145	Average Score: 26.87	Score: 32.700
Episode 150	Average Score: 28.51	Score: 35.100
Episode 155	Average Score: 29.86	Score: 30.488
Episode 156	Average Score: 30.16	Score: 33.480

Environment solved in 156 episodes!

