

# Udacity

## *Deep Reinforcement Learning Nanodegree*

### *Project 2- Continuous Control Report*

#### Algorithm

The algorithm used to train our agent is DDPG (Deep Deterministic Policy Gradient). DDPG is an algorithm which concurrently learns a Q-function and a policy. It uses off-policy data and the Bellman equation to learn the Q-function, and uses the Q-function to learn the policy. This dual mechanism is the actor-critic method. The DDPG algorithm uses two additional mechanisms: Replay Buffer and Soft Updates.

#### Goal

Our agent must get an average score of +30 (over 100 consecutive episodes, and over all agents). Specifically,

- After each episode, we add up the rewards that each agent received (without discounting), to get a score for each agent. This yields 20 (potentially different) scores. We then take the average of these 20 scores.
- This yields an average score for each episode (where the average is over all 20 agents).

**The environment is considered solved, when the average (over 100 episodes) of those average scores is at least +30.**

#### Hyperparameters

```
BUFFER_SIZE = int(1e6)
BATCH_SIZE = 256
GAMMA = 0.99
TAU = 1e-3
LR_ACTOR = 1e-3
LR_CRITIC = 1e-3
WEIGHT_DECAY = 0
EPSILON = 1.0
EPSILON_DECAY = 1e-6
LEARNING_PERIOD = 20
UPDATE_FACTOR = 10
```

## DDPG Neural Networks

```
actor_local = Actor(state_size, action_size, random_seed).to(device)
actor_target = Actor(state_size, action_size, random_seed).to(device)
critic_local = Critic(state_size, action_size, random_seed).to(device)
critic_target = Critic(state_size, action_size, random_seed).to(device)
```

The behavior of classes *Actor* and *Critic* as per model.py is as follows:

```
actor_target(state) -> action
critic_target(state, action) -> Q-value
actor_local(states) -> actions_pred
-critic_local(states, actions_pred) -> actor_loss
```

## Updating critic Neural Network

```
Q_targets_next = critic_target(next_states, actions_next)
Q_targets = rewards + (gamma * Q_targets_next * (1 - dones))
Q_expected = critic_local(states, actions)
critic_loss = MSE_loss(Q_expected, Q_targets)
```

## Updating actor Neural Network

```
actions_pred = actor_local(states)
actor_loss = -critic_local(states, actions_pred).mean()
```

**Note:** In both neural networks we minimize the loss. By Gradient Descent and Backward Propagation, the weights of the network are updated.

## Results of Training

Using Udacity workspace with GPU enabled it took 01h:05m:59s and 200 Episodes to give an Average Score: 30.21. We got following results after training our agent

```
Episode: 1, Score: 0.31, Max: 0.82, Min: 0.00
Episode: 2, Score: 0.46, Max: 1.75, Min: 0.00
Episode: 3, Score: 0.77, Max: 3.90, Min: 0.03
```

Episode: 4, Score: 0.89, Max: 2.24, Min: 0.00  
Episode: 5, Score: 0.77, Max: 2.02, Min: 0.00  
Episode: 6, Score: 0.46, Max: 1.34, Min: 0.00  
Episode: 7, Score: 0.56, Max: 1.66, Min: 0.00  
Episode: 8, Score: 0.43, Max: 1.22, Min: 0.00  
Episode: 9, Score: 0.88, Max: 2.12, Min: 0.16  
Episode: 10, Score: 0.91, Max: 2.67, Min: 0.00  
\*\*\* Episode 10 Average Score: 0.64, Time: 00:01:54 \*\*\*  
Episode: 11, Score: 1.03, Max: 2.30, Min: 0.18  
Episode: 12, Score: 0.77, Max: 1.77, Min: 0.12  
Episode: 13, Score: 0.65, Max: 1.70, Min: 0.00  
Episode: 14, Score: 1.20, Max: 2.29, Min: 0.10  
Episode: 15, Score: 0.65, Max: 1.53, Min: 0.00  
Episode: 16, Score: 0.96, Max: 2.09, Min: 0.13  
Episode: 17, Score: 1.27, Max: 2.28, Min: 0.22  
Episode: 18, Score: 1.09, Max: 2.68, Min: 0.35  
Episode: 19, Score: 1.13, Max: 2.52, Min: 0.18  
Episode: 20, Score: 1.52, Max: 3.19, Min: 0.47  
\*\*\* Episode 20 Average Score: 0.84, Time: 00:04:00 \*\*\*  
Episode: 21, Score: 1.53, Max: 3.05, Min: 0.22  
Episode: 22, Score: 1.19, Max: 2.34, Min: 0.00  
Episode: 23, Score: 1.61, Max: 3.59, Min: 0.35  
Episode: 24, Score: 1.52, Max: 2.57, Min: 0.38  
Episode: 25, Score: 1.74, Max: 3.97, Min: 0.00  
Episode: 26, Score: 1.48, Max: 4.01, Min: 0.12  
Episode: 27, Score: 2.18, Max: 4.90, Min: 0.31  
Episode: 28, Score: 1.80, Max: 3.68, Min: 0.17  
Episode: 29, Score: 1.60, Max: 2.96, Min: 0.43  
Episode: 30, Score: 2.33, Max: 4.94, Min: 0.47  
\*\*\* Episode 30 Average Score: 1.12, Time: 00:06:28 \*\*\*  
Episode: 31, Score: 1.84, Max: 3.48, Min: 0.13  
Episode: 32, Score: 2.21, Max: 4.11, Min: 0.66  
Episode: 33, Score: 2.67, Max: 5.76, Min: 1.33  
Episode: 34, Score: 2.44, Max: 3.75, Min: 1.21  
Episode: 35, Score: 2.61, Max: 4.51, Min: 0.61  
Episode: 36, Score: 2.98, Max: 4.83, Min: 0.98  
Episode: 37, Score: 2.95, Max: 7.54, Min: 0.52  
Episode: 38, Score: 3.42, Max: 6.37, Min: 0.22  
Episode: 39, Score: 3.71, Max: 6.66, Min: 1.38  
Episode: 40, Score: 3.06, Max: 4.89, Min: 1.90  
\*\*\* Episode 40 Average Score: 1.54, Time: 00:09:22 \*\*\*  
Episode: 41, Score: 4.09, Max: 6.47, Min: 1.52  
Episode: 42, Score: 3.82, Max: 6.37, Min: 1.25  
Episode: 43, Score: 4.33, Max: 6.17, Min: 2.15  
Episode: 44, Score: 4.13, Max: 8.91, Min: 2.29  
Episode: 45, Score: 4.67, Max: 10.16, Min: 0.60  
Episode: 46, Score: 4.62, Max: 8.15, Min: 1.57  
Episode: 47, Score: 4.50, Max: 7.19, Min: 1.81  
Episode: 48, Score: 4.30, Max: 8.08, Min: 0.91  
Episode: 49, Score: 5.19, Max: 8.64, Min: 2.03  
Episode: 50, Score: 6.41, Max: 16.78, Min: 2.12  
\*\*\* Episode 50 Average Score: 2.15, Time: 00:12:41 \*\*\*  
Episode: 51, Score: 4.99, Max: 8.50, Min: 1.78  
Episode: 52, Score: 5.20, Max: 10.32, Min: 2.89

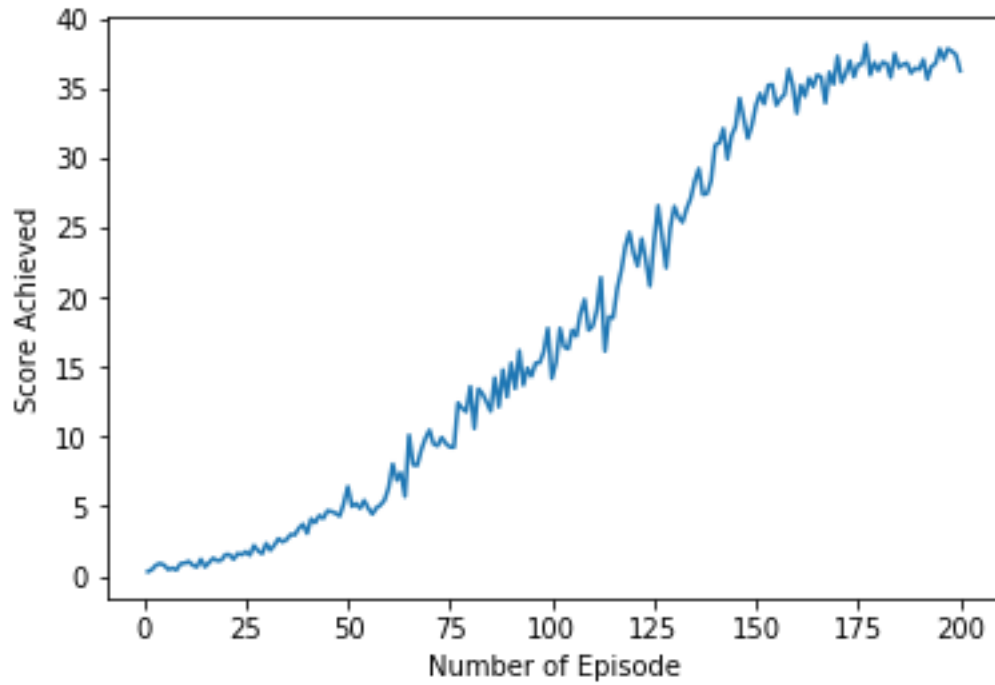
Episode: 53, Score: 4.85, Max: 8.94, Min: 2.22  
Episode: 54, Score: 5.42, Max: 7.43, Min: 3.58  
Episode: 55, Score: 4.86, Max: 7.67, Min: 1.58  
Episode: 56, Score: 4.44, Max: 8.08, Min: 0.92  
Episode: 57, Score: 4.89, Max: 7.82, Min: 1.05  
Episode: 58, Score: 5.09, Max: 9.14, Min: 1.83  
Episode: 59, Score: 5.45, Max: 8.71, Min: 0.75  
Episode: 60, Score: 6.33, Max: 12.91, Min: 3.99  
\*\*\* Episode 60 Average Score: 2.65, Time: 00:16:15 \*\*\*  
Episode: 61, Score: 8.02, Max: 30.58, Min: 3.80  
Episode: 62, Score: 6.87, Max: 26.65, Min: 2.99  
Episode: 63, Score: 7.41, Max: 19.09, Min: 3.70  
Episode: 64, Score: 5.71, Max: 13.16, Min: 1.17  
Episode: 65, Score: 10.10, Max: 29.29, Min: 4.60  
Episode: 66, Score: 7.98, Max: 15.51, Min: 3.25  
Episode: 67, Score: 7.94, Max: 12.32, Min: 3.26  
Episode: 68, Score: 9.05, Max: 18.56, Min: 3.38  
Episode: 69, Score: 9.87, Max: 16.22, Min: 3.36  
Episode: 70, Score: 10.47, Max: 16.70, Min: 3.03  
\*\*\* Episode 70 Average Score: 3.47, Time: 00:19:47 \*\*\*  
Episode: 71, Score: 9.47, Max: 14.10, Min: 5.37  
Episode: 72, Score: 9.34, Max: 15.61, Min: 4.20  
Episode: 73, Score: 9.94, Max: 15.83, Min: 5.27  
Episode: 74, Score: 9.51, Max: 15.84, Min: 3.23  
Episode: 75, Score: 9.25, Max: 19.36, Min: 3.18  
Episode: 76, Score: 9.24, Max: 12.16, Min: 5.64  
Episode: 77, Score: 12.41, Max: 30.06, Min: 6.31  
Episode: 78, Score: 12.03, Max: 19.00, Min: 4.25  
Episode: 79, Score: 11.79, Max: 16.88, Min: 7.00  
Episode: 80, Score: 13.59, Max: 22.56, Min: 7.82  
\*\*\* Episode 80 Average Score: 4.36, Time: 00:23:19 \*\*\*  
Episode: 81, Score: 10.58, Max: 16.80, Min: 3.17  
Episode: 82, Score: 13.42, Max: 30.82, Min: 4.87  
Episode: 83, Score: 13.07, Max: 17.99, Min: 8.92  
Episode: 84, Score: 12.48, Max: 17.30, Min: 5.49  
Episode: 85, Score: 11.83, Max: 18.89, Min: 5.16  
Episode: 86, Score: 14.22, Max: 20.19, Min: 8.52  
Episode: 87, Score: 12.14, Max: 18.28, Min: 2.46  
Episode: 88, Score: 14.79, Max: 21.32, Min: 9.35  
Episode: 89, Score: 12.86, Max: 17.61, Min: 6.12  
Episode: 90, Score: 15.27, Max: 23.78, Min: 7.09  
\*\*\* Episode 90 Average Score: 5.33, Time: 00:26:52 \*\*\*  
Episode: 91, Score: 13.45, Max: 17.91, Min: 3.52  
Episode: 92, Score: 16.17, Max: 31.16, Min: 8.55  
Episode: 93, Score: 13.70, Max: 18.53, Min: 3.68  
Episode: 94, Score: 14.92, Max: 21.80, Min: 5.82  
Episode: 95, Score: 14.34, Max: 18.32, Min: 9.54  
Episode: 96, Score: 15.27, Max: 25.33, Min: 6.36  
Episode: 97, Score: 15.30, Max: 21.03, Min: 6.29  
Episode: 98, Score: 16.08, Max: 22.37, Min: 8.26  
Episode: 99, Score: 17.77, Max: 39.47, Min: 10.96  
Episode: 100, Score: 14.18, Max: 21.90, Min: 1.59  
\*\*\* Episode 100 Average Score: 6.31, Time: 00:30:25 \*\*\*  
Episode: 101, Score: 15.34, Max: 20.73, Min: 9.14

Episode: 102, Score: 17.77, Max: 22.09, Min: 7.72  
Episode: 103, Score: 16.44, Max: 25.77, Min: 3.52  
Episode: 104, Score: 16.30, Max: 23.22, Min: 0.10  
Episode: 105, Score: 17.64, Max: 21.84, Min: 7.86  
Episode: 106, Score: 17.21, Max: 25.10, Min: 7.77  
Episode: 107, Score: 18.85, Max: 35.11, Min: 6.72  
Episode: 108, Score: 19.84, Max: 32.51, Min: 10.62  
Episode: 109, Score: 17.65, Max: 29.23, Min: 6.38  
Episode: 110, Score: 17.93, Max: 27.31, Min: 10.75  
\*\*\* Episode 110           Average Score: 8.00, Time: 00:33:58 \*\*\*  
Episode: 111, Score: 19.07, Max: 25.10, Min: 10.77  
Episode: 112, Score: 21.42, Max: 26.33, Min: 9.45  
Episode: 113, Score: 16.14, Max: 25.71, Min: 6.71  
Episode: 114, Score: 18.61, Max: 24.55, Min: 9.74  
Episode: 115, Score: 18.52, Max: 26.03, Min: 8.68  
Episode: 116, Score: 20.62, Max: 27.36, Min: 9.76  
Episode: 117, Score: 21.92, Max: 36.63, Min: 16.37  
Episode: 118, Score: 23.66, Max: 39.44, Min: 19.50  
Episode: 119, Score: 24.66, Max: 34.72, Min: 15.23  
Episode: 120, Score: 23.16, Max: 29.91, Min: 9.75  
\*\*\* Episode 120           Average Score: 9.97, Time: 00:37:32 \*\*\*  
Episode: 121, Score: 22.25, Max: 34.46, Min: 3.16  
Episode: 122, Score: 24.21, Max: 36.02, Min: 12.87  
Episode: 123, Score: 22.65, Max: 32.32, Min: 15.47  
Episode: 124, Score: 20.81, Max: 32.60, Min: 12.39  
Episode: 125, Score: 23.90, Max: 32.22, Min: 13.00  
Episode: 126, Score: 26.57, Max: 37.82, Min: 22.32  
Episode: 127, Score: 24.40, Max: 29.42, Min: 11.83  
Episode: 128, Score: 22.11, Max: 29.30, Min: 15.14  
Episode: 129, Score: 24.99, Max: 39.07, Min: 15.69  
Episode: 130, Score: 26.52, Max: 39.33, Min: 17.51  
\*\*\* Episode 130           Average Score: 12.18, Time: 00:41:05 \*\*\*  
Episode: 131, Score: 25.76, Max: 32.82, Min: 19.93  
Episode: 132, Score: 25.39, Max: 34.31, Min: 13.73  
Episode: 133, Score: 26.36, Max: 33.80, Min: 16.51  
Episode: 134, Score: 27.10, Max: 37.00, Min: 16.08  
Episode: 135, Score: 28.35, Max: 37.23, Min: 14.98  
Episode: 136, Score: 29.23, Max: 36.15, Min: 15.05  
Episode: 137, Score: 27.38, Max: 35.27, Min: 14.40  
Episode: 138, Score: 27.44, Max: 38.00, Min: 15.86  
Episode: 139, Score: 28.36, Max: 35.56, Min: 14.62  
Episode: 140, Score: 30.91, Max: 35.01, Min: 25.33  
\*\*\* Episode 140           Average Score: 14.67, Time: 00:44:38 \*\*\*  
Episode: 141, Score: 31.09, Max: 37.86, Min: 20.11  
Episode: 142, Score: 32.11, Max: 39.15, Min: 21.97  
Episode: 143, Score: 29.93, Max: 39.35, Min: 11.90  
Episode: 144, Score: 31.60, Max: 36.15, Min: 23.17  
Episode: 145, Score: 32.25, Max: 36.41, Min: 25.36  
Episode: 146, Score: 34.29, Max: 37.25, Min: 30.02  
Episode: 147, Score: 32.80, Max: 37.67, Min: 23.58  
Episode: 148, Score: 31.41, Max: 37.80, Min: 17.91  
Episode: 149, Score: 32.34, Max: 36.61, Min: 19.51  
Episode: 150, Score: 33.77, Max: 38.09, Min: 22.63  
\*\*\* Episode 150           Average Score: 17.42, Time: 00:48:12 \*\*\*

Episode: 151, Score: 34.62, Max: 38.90, Min: 25.92  
Episode: 152, Score: 33.88, Max: 39.17, Min: 23.90  
Episode: 153, Score: 35.21, Max: 39.16, Min: 26.61  
Episode: 154, Score: 35.27, Max: 39.52, Min: 31.15  
Episode: 155, Score: 33.79, Max: 39.50, Min: 20.23  
Episode: 156, Score: 34.26, Max: 39.45, Min: 26.17  
Episode: 157, Score: 34.55, Max: 38.49, Min: 24.37  
Episode: 158, Score: 36.35, Max: 39.19, Min: 30.45  
Episode: 159, Score: 35.22, Max: 38.54, Min: 28.45  
Episode: 160, Score: 33.22, Max: 38.38, Min: 27.30  
\*\*\* Episode 160      Average Score: 20.37, Time: 00:51:46 \*\*\*  
Episode: 161, Score: 35.21, Max: 38.26, Min: 31.03  
Episode: 162, Score: 34.42, Max: 38.44, Min: 26.88  
Episode: 163, Score: 35.72, Max: 39.30, Min: 30.08  
Episode: 164, Score: 35.11, Max: 38.84, Min: 26.74  
Episode: 165, Score: 35.95, Max: 38.97, Min: 25.23  
Episode: 166, Score: 35.81, Max: 39.49, Min: 30.40  
Episode: 167, Score: 33.96, Max: 38.78, Min: 20.60  
Episode: 168, Score: 36.15, Max: 39.46, Min: 32.27  
Episode: 169, Score: 35.30, Max: 39.47, Min: 30.27  
Episode: 170, Score: 37.29, Max: 39.47, Min: 32.17  
\*\*\* Episode 170      Average Score: 23.09, Time: 00:55:19 \*\*\*  
Episode: 171, Score: 35.42, Max: 39.03, Min: 26.74  
Episode: 172, Score: 36.07, Max: 39.24, Min: 31.35  
Episode: 173, Score: 36.95, Max: 39.48, Min: 29.55  
Episode: 174, Score: 35.83, Max: 39.50, Min: 24.39  
Episode: 175, Score: 36.70, Max: 39.10, Min: 30.66  
Episode: 176, Score: 36.73, Max: 39.40, Min: 28.42  
Episode: 177, Score: 38.18, Max: 39.58, Min: 34.24  
Episode: 178, Score: 35.96, Max: 39.13, Min: 30.84  
Episode: 179, Score: 36.85, Max: 39.14, Min: 28.81  
Episode: 180, Score: 36.28, Max: 39.46, Min: 30.43  
\*\*\* Episode 180      Average Score: 25.67, Time: 00:58:53 \*\*\*  
Episode: 181, Score: 36.87, Max: 39.30, Min: 30.64  
Episode: 182, Score: 36.78, Max: 39.03, Min: 29.87  
Episode: 183, Score: 35.78, Max: 38.46, Min: 28.22  
Episode: 184, Score: 37.48, Max: 39.29, Min: 35.77  
Episode: 185, Score: 36.50, Max: 39.39, Min: 32.53  
Episode: 186, Score: 36.71, Max: 39.14, Min: 31.65  
Episode: 187, Score: 36.81, Max: 38.97, Min: 33.91  
Episode: 188, Score: 36.09, Max: 39.38, Min: 29.28  
Episode: 189, Score: 36.45, Max: 39.17, Min: 30.47  
Episode: 190, Score: 36.35, Max: 38.42, Min: 31.58  
\*\*\* Episode 190      Average Score: 28.02, Time: 01:02:26 \*\*\*  
Episode: 191, Score: 37.05, Max: 38.82, Min: 35.09  
Episode: 192, Score: 35.66, Max: 39.49, Min: 27.43  
Episode: 193, Score: 36.57, Max: 39.56, Min: 33.43  
Episode: 194, Score: 36.77, Max: 38.62, Min: 31.92  
Episode: 195, Score: 37.84, Max: 39.08, Min: 35.12  
Episode: 196, Score: 37.08, Max: 39.45, Min: 33.46  
Episode: 197, Score: 37.84, Max: 39.22, Min: 36.32  
Episode: 198, Score: 37.65, Max: 39.57, Min: 33.53  
Episode: 199, Score: 37.41, Max: 39.21, Min: 29.68  
Episode: 200, Score: 36.26, Max: 39.28, Min: 33.08

\*\*\* Episode 200      Average Score: 30.21, Time: 01:05:59 \*\*\*  
Environment solved !

## Plot



## Ideas for Future Work

1. We can add additional layers in Actor and Critic Neural Network
2. We can use other algorithms than DDPG, like PPO or A3C.