

deep_reinforcement_learning_project3

Overview

For this project, you will work with the Tennis environment.

In this environment, two agents control rackets to bounce a ball over a net. If an agent hits the ball over the net, it receives a reward of $+0.1$. If an agent lets a ball hit the ground or hits the ball out of bounds, it receives a reward of -0.01 . Thus, the goal of each agent is to keep the ball in play.

The observation space consists of 8 variables corresponding to the position and velocity of the ball and racket. Each agent receives its own, local observation. Two continuous actions are available, corresponding to movement toward (or away from) the net, and jumping.

The task is episodic, and in order to solve the environment, your agents must get an average score of $+0.5$ (over 100 consecutive episodes, after taking the maximum over both 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 2 (potentially different) scores. We then take the maximum of these 2 scores. This yields a single score

for each episode. The environment is considered solved, when the average (over 100 episodes) of those scores is at least +0.5.

Code Architecture

Tennis.ipynb: jupyter notebook based solution

ddpg_agent.py: DDPG agent code

model.py: Actor/Critic model

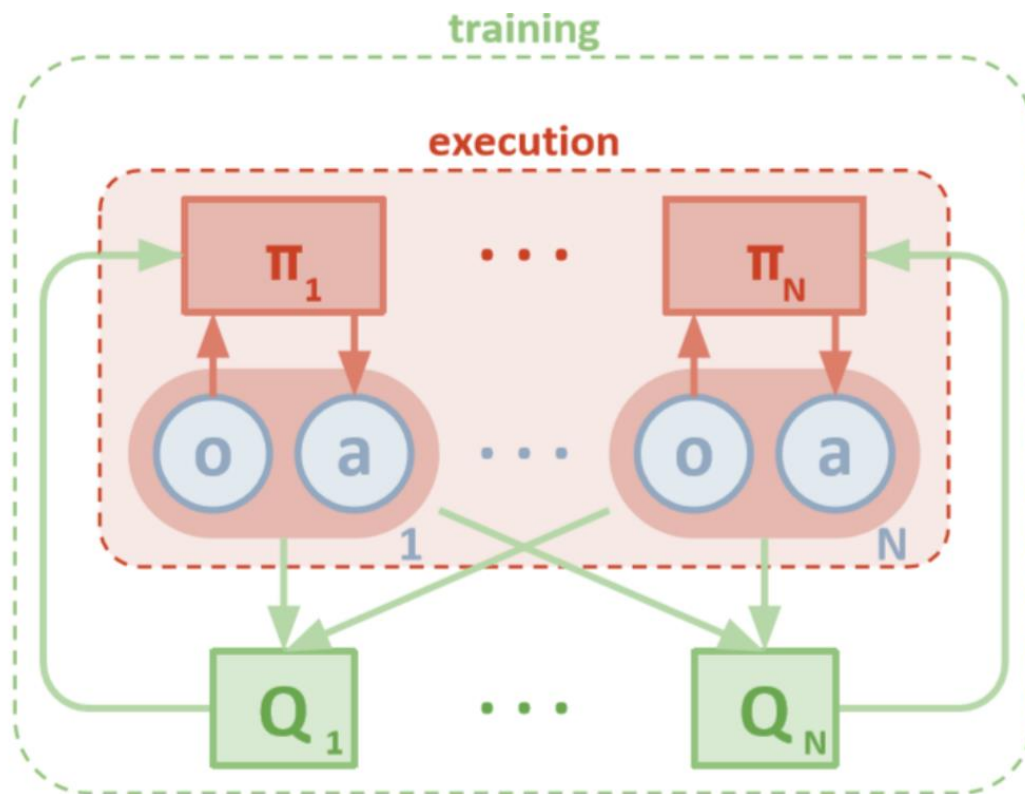
actor0.pth: weights of the agent0 actor model

critic0.pth: weights of the agent0 critic model

actor1.pth: weights of the agent1 actor model

critic1.pth: weights of the agent1 critic model

Multi-Agent Architecture



MADDPG Architecture (Lowe, 2018)

Every agent has an observation space and continuous action space. Also, each agent has three components:

- An actor-network that uses local observations for deterministic actions
- A target actor-network with identical functionality for training stability
- A critic-network that uses joint states action pairs to estimate Q-values

As the critic learns the joint Q-value function over time, it sends appropriate Q-value approximations to the actor to help training.

1) Refers to: [OpenAI's MADDPG Algorithm | Towards Data Science](#)

2) The code implementation referenced the code provided by Udacity

Multi-Agent DDPG Algorithm

Algorithm 1: Multi-Agent Deep Deterministic Policy Gradient for N agents

```
for episode = 1 to  $M$  do
  Initialize a random process  $\mathcal{N}$  for action exploration
  Receive initial state  $\mathbf{x}$ 
  for  $t = 1$  to max-episode-length do
    for each agent  $i$ , select action  $a_i = \mu_{\theta_i}(o_i) + \mathcal{N}_t$  w.r.t. the current policy and exploration
    Execute actions  $a = (a_1, \dots, a_N)$  and observe reward  $r$  and new state  $\mathbf{x}'$ 
    Store  $(\mathbf{x}, a, r, \mathbf{x}')$  in replay buffer  $\mathcal{D}$ 
     $\mathbf{x} \leftarrow \mathbf{x}'$ 
    for agent  $i = 1$  to  $N$  do
      Sample a random minibatch of  $S$  samples  $(\mathbf{x}^j, a^j, r^j, \mathbf{x}'^j)$  from  $\mathcal{D}$ 
      Set  $y^j = r_i^j + \gamma Q_i^{\mu'}(\mathbf{x}^j, a_1^j, \dots, a_N^j)|_{a_k^j = \mu_k^{\mu'}(o_k^j)}$ 
      Update critic by minimizing the loss  $\mathcal{L}(\theta_i) = \frac{1}{S} \sum_j \left( y^j - Q_i^{\mu}(\mathbf{x}^j, a_1^j, \dots, a_N^j) \right)^2$ 
      Update actor using the sampled policy gradient:
      
$$\nabla_{\theta_i} J \approx \frac{1}{S} \sum_j \nabla_{\theta_i} \mu_i(o_i^j) \nabla_{a_i} Q_i^{\mu}(\mathbf{x}^j, a_1^j, \dots, a_i, \dots, a_N^j)|_{a_i = \mu_i(o_i^j)}$$

    end for
    Update target network parameters for each agent  $i$ :
    
$$\theta_i' \leftarrow \tau \theta_i + (1 - \tau) \theta_i'$$

  end for
end for
```

*) Refers to: [A Pseudo Code for Multi-Agent DDPG algorithm. | Download Scientific Diagram \(researchgate.net\)](#)

Plot the score

It takes 3678 episodes to get an average score of 0.51

Episode 10 Average Score: 0.00

Episode 20 Average Score: 0.00

Episode 30 Average Score: 0.00

Episode 40 Average Score: 0.00

Episode 50 Average Score: 0.00

Episode 60 Average Score: 0.00

Episode 70 Average Score: 0.00

Episode 80 Average Score: 0.00

Episode 90 Average Score: 0.00

Episode 100 Average Score: 0.00

Episode 110 Average Score: 0.00

Episode 120 Average Score: 0.00

Episode 130 Average Score: 0.00

Episode 140 Average Score: 0.00

Episode 150 Average Score: 0.00

Episode 160 Average Score: 0.00

Episode 170 Average Score: 0.00

Episode 180 Average Score: 0.00

Episode 190 Average Score: 0.00

Episode 200 Average Score: 0.00

Episode 210 Average Score: 0.00

Episode 220 Average Score: 0.00

Episode 230 Average Score: 0.00

Episode 240 Average Score: 0.00

Episode 250 Average Score: 0.00

Episode 260 Average Score: 0.00

Episode 270 Average Score: 0.00

Episode 280 Average Score: 0.00

Episode 290 Average Score: 0.00

Episode 300	Average Score: 0.00
Episode 310	Average Score: 0.00
Episode 320	Average Score: 0.00
Episode 330	Average Score: 0.00
Episode 340	Average Score: 0.00
Episode 350	Average Score: 0.00
Episode 360	Average Score: 0.00
Episode 370	Average Score: 0.00
Episode 380	Average Score: 0.00
Episode 390	Average Score: 0.00
Episode 400	Average Score: 0.00
Episode 410	Average Score: 0.00
Episode 420	Average Score: 0.00
Episode 430	Average Score: 0.00
Episode 440	Average Score: 0.00
Episode 450	Average Score: 0.00
Episode 460	Average Score: 0.00
Episode 470	Average Score: 0.00
Episode 480	Average Score: 0.00
Episode 490	Average Score: 0.00
Episode 500	Average Score: 0.00
Episode 510	Average Score: 0.00

Episode 520	Average Score: 0.00
Episode 530	Average Score: 0.00
Episode 540	Average Score: 0.00
Episode 550	Average Score: 0.00
Episode 560	Average Score: 0.00
Episode 570	Average Score: 0.00
Episode 580	Average Score: 0.00
Episode 590	Average Score: 0.00
Episode 600	Average Score: 0.00
Episode 610	Average Score: 0.00
Episode 620	Average Score: 0.00
Episode 630	Average Score: 0.00
Episode 640	Average Score: 0.00
Episode 650	Average Score: 0.00
Episode 660	Average Score: 0.00
Episode 670	Average Score: 0.00
Episode 680	Average Score: 0.01
Episode 690	Average Score: 0.01
Episode 700	Average Score: 0.01
Episode 710	Average Score: 0.01
Episode 720	Average Score: 0.01
Episode 730	Average Score: 0.01

Episode 740	Average Score: 0.01
Episode 750	Average Score: 0.01
Episode 760	Average Score: 0.01
Episode 770	Average Score: 0.01
Episode 780	Average Score: 0.00
Episode 790	Average Score: 0.00
Episode 800	Average Score: 0.00
Episode 810	Average Score: 0.00
Episode 820	Average Score: 0.00
Episode 830	Average Score: 0.00
Episode 840	Average Score: 0.00
Episode 850	Average Score: 0.00
Episode 860	Average Score: 0.00
Episode 870	Average Score: 0.00
Episode 880	Average Score: 0.00
Episode 890	Average Score: 0.00
Episode 900	Average Score: 0.00
Episode 910	Average Score: 0.00
Episode 920	Average Score: 0.00
Episode 930	Average Score: 0.00
Episode 940	Average Score: 0.00
Episode 950	Average Score: 0.00

Episode 960	Average Score: 0.00
Episode 970	Average Score: 0.00
Episode 980	Average Score: 0.00
Episode 990	Average Score: 0.00
Episode 1000	Average Score: 0.00
Episode 1010	Average Score: 0.00
Episode 1020	Average Score: 0.00
Episode 1030	Average Score: 0.00
Episode 1040	Average Score: 0.00
Episode 1050	Average Score: 0.00
Episode 1060	Average Score: 0.00
Episode 1070	Average Score: 0.00
Episode 1080	Average Score: 0.00
Episode 1090	Average Score: 0.00
Episode 1100	Average Score: 0.00
Episode 1110	Average Score: 0.00
Episode 1120	Average Score: 0.00
Episode 1130	Average Score: 0.00
Episode 1140	Average Score: 0.01
Episode 1150	Average Score: 0.01
Episode 1160	Average Score: 0.01
Episode 1170	Average Score: 0.01

Episode 1180	Average Score: 0.01
Episode 1190	Average Score: 0.01
Episode 1200	Average Score: 0.01
Episode 1210	Average Score: 0.01
Episode 1220	Average Score: 0.01
Episode 1230	Average Score: 0.01
Episode 1240	Average Score: 0.00
Episode 1250	Average Score: 0.00
Episode 1260	Average Score: 0.00
Episode 1270	Average Score: 0.01
Episode 1280	Average Score: 0.01
Episode 1290	Average Score: 0.01
Episode 1300	Average Score: 0.01
Episode 1310	Average Score: 0.01
Episode 1320	Average Score: 0.01
Episode 1330	Average Score: 0.01
Episode 1340	Average Score: 0.02
Episode 1350	Average Score: 0.02
Episode 1360	Average Score: 0.02
Episode 1370	Average Score: 0.02
Episode 1380	Average Score: 0.02
Episode 1390	Average Score: 0.02

Episode 1400	Average Score: 0.02
Episode 1410	Average Score: 0.02
Episode 1420	Average Score: 0.02
Episode 1430	Average Score: 0.01
Episode 1440	Average Score: 0.01
Episode 1450	Average Score: 0.02
Episode 1460	Average Score: 0.02
Episode 1470	Average Score: 0.02
Episode 1480	Average Score: 0.02
Episode 1490	Average Score: 0.02
Episode 1500	Average Score: 0.03
Episode 1510	Average Score: 0.03
Episode 1520	Average Score: 0.03
Episode 1530	Average Score: 0.04
Episode 1540	Average Score: 0.04
Episode 1550	Average Score: 0.05
Episode 1560	Average Score: 0.04
Episode 1570	Average Score: 0.04
Episode 1580	Average Score: 0.04
Episode 1590	Average Score: 0.04
Episode 1600	Average Score: 0.04
Episode 1610	Average Score: 0.05

Episode 1620	Average Score: 0.04
Episode 1630	Average Score: 0.04
Episode 1640	Average Score: 0.04
Episode 1650	Average Score: 0.04
Episode 1660	Average Score: 0.05
Episode 1670	Average Score: 0.05
Episode 1680	Average Score: 0.05
Episode 1690	Average Score: 0.05
Episode 1700	Average Score: 0.05
Episode 1710	Average Score: 0.05
Episode 1720	Average Score: 0.06
Episode 1730	Average Score: 0.06
Episode 1740	Average Score: 0.06
Episode 1750	Average Score: 0.06
Episode 1760	Average Score: 0.06
Episode 1770	Average Score: 0.06
Episode 1780	Average Score: 0.05
Episode 1790	Average Score: 0.06
Episode 1800	Average Score: 0.05
Episode 1810	Average Score: 0.05
Episode 1820	Average Score: 0.06
Episode 1830	Average Score: 0.05

Episode 1840	Average Score: 0.05
Episode 1850	Average Score: 0.05
Episode 1860	Average Score: 0.06
Episode 1870	Average Score: 0.06
Episode 1880	Average Score: 0.06
Episode 1890	Average Score: 0.07
Episode 1900	Average Score: 0.07
Episode 1910	Average Score: 0.07
Episode 1920	Average Score: 0.07
Episode 1930	Average Score: 0.08
Episode 1940	Average Score: 0.08
Episode 1950	Average Score: 0.09
Episode 1960	Average Score: 0.08
Episode 1970	Average Score: 0.08
Episode 1980	Average Score: 0.08
Episode 1990	Average Score: 0.08
Episode 2000	Average Score: 0.08
Episode 2010	Average Score: 0.08
Episode 2020	Average Score: 0.08
Episode 2030	Average Score: 0.08
Episode 2040	Average Score: 0.07
Episode 2050	Average Score: 0.07

Episode 2060	Average Score: 0.08
Episode 2070	Average Score: 0.07
Episode 2080	Average Score: 0.08
Episode 2090	Average Score: 0.08
Episode 2100	Average Score: 0.07
Episode 2110	Average Score: 0.07
Episode 2120	Average Score: 0.07
Episode 2130	Average Score: 0.07
Episode 2140	Average Score: 0.08
Episode 2150	Average Score: 0.07
Episode 2160	Average Score: 0.07
Episode 2170	Average Score: 0.08
Episode 2180	Average Score: 0.08
Episode 2190	Average Score: 0.08
Episode 2200	Average Score: 0.08
Episode 2210	Average Score: 0.08
Episode 2220	Average Score: 0.08
Episode 2230	Average Score: 0.08
Episode 2240	Average Score: 0.08
Episode 2250	Average Score: 0.08
Episode 2260	Average Score: 0.08
Episode 2270	Average Score: 0.08

Episode 2280	Average Score: 0.08
Episode 2290	Average Score: 0.08
Episode 2300	Average Score: 0.08
Episode 2310	Average Score: 0.08
Episode 2320	Average Score: 0.08
Episode 2330	Average Score: 0.08
Episode 2340	Average Score: 0.08
Episode 2350	Average Score: 0.08
Episode 2360	Average Score: 0.08
Episode 2370	Average Score: 0.08
Episode 2380	Average Score: 0.09
Episode 2390	Average Score: 0.09
Episode 2400	Average Score: 0.09
Episode 2410	Average Score: 0.09
Episode 2420	Average Score: 0.09
Episode 2430	Average Score: 0.09
Episode 2440	Average Score: 0.09
Episode 2450	Average Score: 0.09
Episode 2460	Average Score: 0.09
Episode 2470	Average Score: 0.09
Episode 2480	Average Score: 0.09
Episode 2490	Average Score: 0.08

Episode 2500	Average Score: 0.08
Episode 2510	Average Score: 0.08
Episode 2520	Average Score: 0.07
Episode 2530	Average Score: 0.08
Episode 2540	Average Score: 0.08
Episode 2550	Average Score: 0.08
Episode 2560	Average Score: 0.08
Episode 2570	Average Score: 0.08
Episode 2580	Average Score: 0.08
Episode 2590	Average Score: 0.08
Episode 2600	Average Score: 0.08
Episode 2610	Average Score: 0.09
Episode 2620	Average Score: 0.09
Episode 2630	Average Score: 0.09
Episode 2640	Average Score: 0.09
Episode 2650	Average Score: 0.09
Episode 2660	Average Score: 0.09
Episode 2670	Average Score: 0.10
Episode 2680	Average Score: 0.10
Episode 2690	Average Score: 0.10
Episode 2700	Average Score: 0.10
Episode 2710	Average Score: 0.10

Episode 2720	Average Score: 0.09
Episode 2730	Average Score: 0.09
Episode 2740	Average Score: 0.09
Episode 2750	Average Score: 0.09
Episode 2760	Average Score: 0.09
Episode 2770	Average Score: 0.09
Episode 2780	Average Score: 0.09
Episode 2790	Average Score: 0.09
Episode 2800	Average Score: 0.09
Episode 2810	Average Score: 0.09
Episode 2820	Average Score: 0.10
Episode 2830	Average Score: 0.10
Episode 2840	Average Score: 0.10
Episode 2850	Average Score: 0.11
Episode 2860	Average Score: 0.11
Episode 2870	Average Score: 0.11
Episode 2880	Average Score: 0.11
Episode 2890	Average Score: 0.11
Episode 2900	Average Score: 0.11
Episode 2910	Average Score: 0.11
Episode 2920	Average Score: 0.11
Episode 2930	Average Score: 0.11

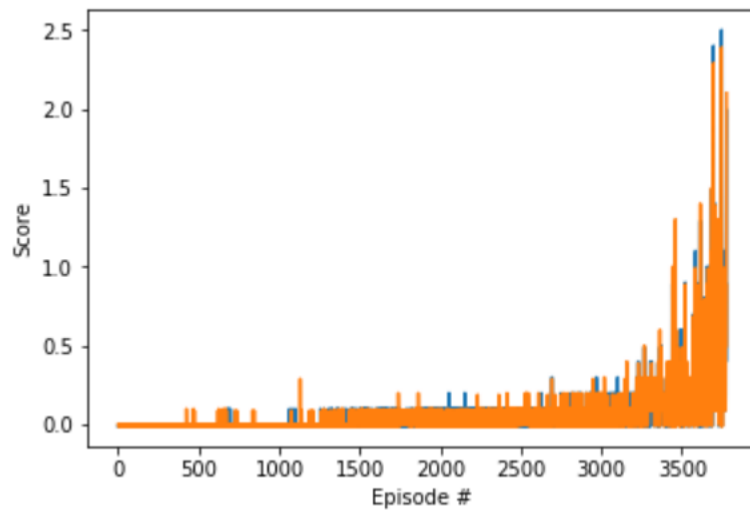
Episode 2940	Average Score: 0.11
Episode 2950	Average Score: 0.12
Episode 2960	Average Score: 0.11
Episode 2970	Average Score: 0.12
Episode 2980	Average Score: 0.12
Episode 2990	Average Score: 0.12
Episode 3000	Average Score: 0.12
Episode 3010	Average Score: 0.12
Episode 3020	Average Score: 0.12
Episode 3030	Average Score: 0.12
Episode 3040	Average Score: 0.12
Episode 3050	Average Score: 0.12
Episode 3060	Average Score: 0.12
Episode 3070	Average Score: 0.12
Episode 3080	Average Score: 0.12
Episode 3090	Average Score: 0.11
Episode 3100	Average Score: 0.11
Episode 3110	Average Score: 0.12
Episode 3120	Average Score: 0.11
Episode 3130	Average Score: 0.11
Episode 3140	Average Score: 0.11
Episode 3150	Average Score: 0.12

Episode 3160	Average Score: 0.12
Episode 3170	Average Score: 0.12
Episode 3180	Average Score: 0.12
Episode 3190	Average Score: 0.12
Episode 3200	Average Score: 0.12
Episode 3210	Average Score: 0.12
Episode 3220	Average Score: 0.12
Episode 3230	Average Score: 0.12
Episode 3240	Average Score: 0.12
Episode 3250	Average Score: 0.12
Episode 3260	Average Score: 0.12
Episode 3270	Average Score: 0.13
Episode 3280	Average Score: 0.13
Episode 3290	Average Score: 0.13
Episode 3300	Average Score: 0.13
Episode 3310	Average Score: 0.13
Episode 3320	Average Score: 0.14
Episode 3330	Average Score: 0.13
Episode 3340	Average Score: 0.13
Episode 3350	Average Score: 0.14
Episode 3360	Average Score: 0.14
Episode 3370	Average Score: 0.14

Episode 3380	Average Score: 0.15
Episode 3390	Average Score: 0.15
Episode 3400	Average Score: 0.15
Episode 3410	Average Score: 0.15
Episode 3420	Average Score: 0.15
Episode 3430	Average Score: 0.15
Episode 3440	Average Score: 0.15
Episode 3450	Average Score: 0.16
Episode 3460	Average Score: 0.17
Episode 3470	Average Score: 0.18
Episode 3480	Average Score: 0.18
Episode 3490	Average Score: 0.18
Episode 3500	Average Score: 0.19
Episode 3510	Average Score: 0.20
Episode 3520	Average Score: 0.20
Episode 3530	Average Score: 0.21
Episode 3540	Average Score: 0.21
Episode 3550	Average Score: 0.21
Episode 3560	Average Score: 0.19
Episode 3570	Average Score: 0.19
Episode 3580	Average Score: 0.20
Episode 3590	Average Score: 0.22

Episode 3600	Average Score: 0.23
Episode 3610	Average Score: 0.24
Episode 3620	Average Score: 0.25
Episode 3630	Average Score: 0.24
Episode 3640	Average Score: 0.25
Episode 3650	Average Score: 0.27
Episode 3660	Average Score: 0.29
Episode 3670	Average Score: 0.30
Episode 3680	Average Score: 0.30
Episode 3690	Average Score: 0.31
Episode 3700	Average Score: 0.32
Episode 3710	Average Score: 0.36
Episode 3720	Average Score: 0.37
Episode 3730	Average Score: 0.39
Episode 3740	Average Score: 0.41
Episode 3750	Average Score: 0.44
Episode 3760	Average Score: 0.45
Episode 3770	Average Score: 0.47

Environment solved in 3678 episodes! Average Score: 0.51



Ideas for future work

DD4G could be further investigated if it has better results.