
Project - Q-Learning on Atari

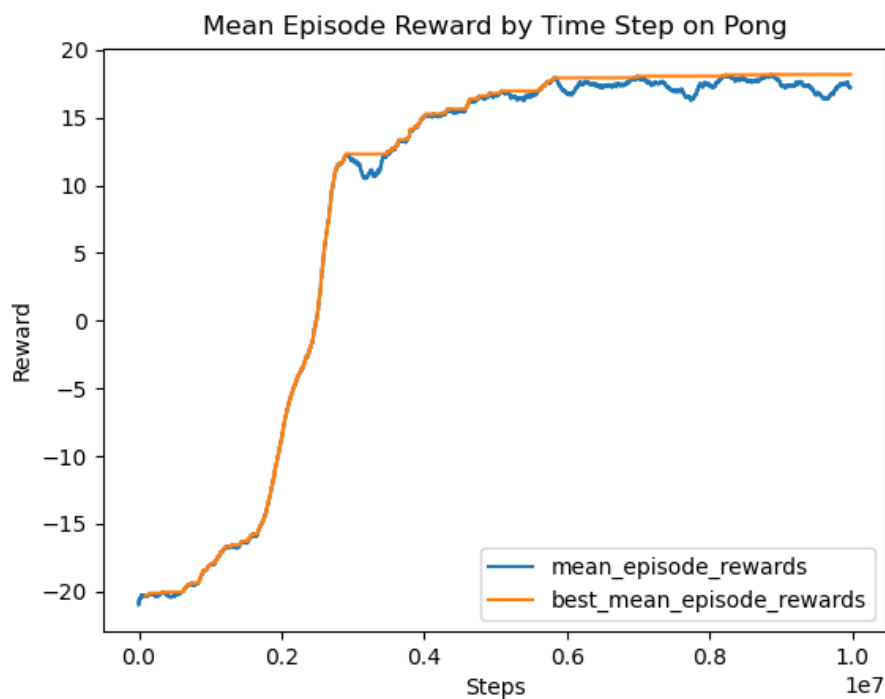
Students names: *Roe Hendel and Karina Buchmann*

Course: *Reinforcement Learning* – Professor: *Yishay Mansour*

Due date: *May 25th, 2022*

Question 1

We ran the training mostly with the default hyper-parameters. We only changed the `schedule_timesteps` parameter of the linear exploration schedule, from 1 million to 2 million steps, to allow for more time to explore. The results are shown in the figure below, we can see that we are able to achieve good results.



Question 2

We chose to experiment with the final exploration probability of the linear exploration schedule. We thought that having a high exploration probability towards the end of training can damage the performance, and therefore we wanted to see if lowering the final probability can improve the performance. The following table lists the values we tested:

Experiment #	final_p
0	0
1	0.05
2	0.1
3	0.2

The following plot shows the results of these experiments

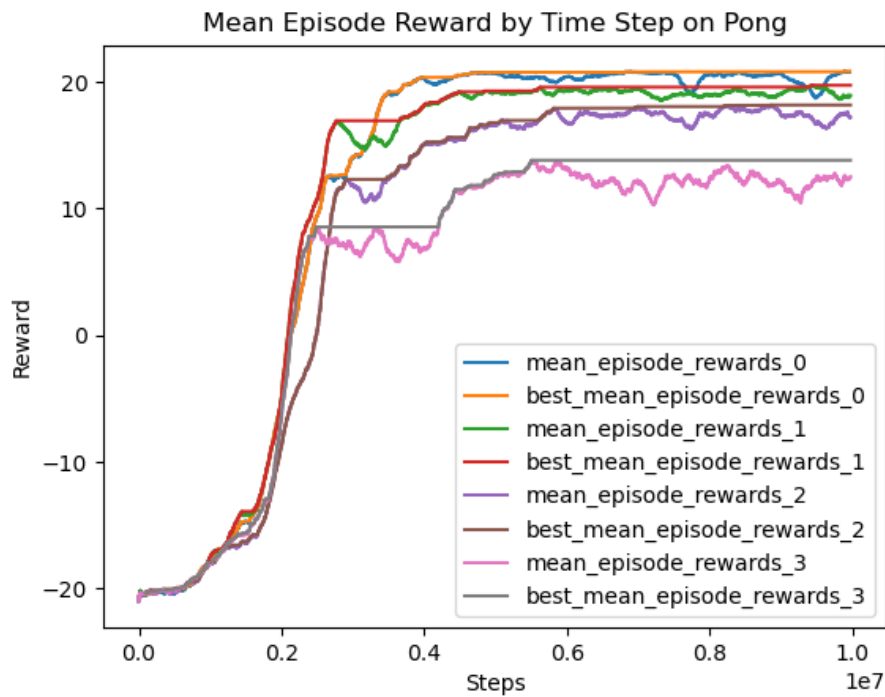


Figure 1: DQN training on Pong, with different values of the final exploration probability. We can see that lower final exploration probability leads to better performance on the task, as we expected.

Additional Experiments

We performed additional experiments to test the generality and robustness of the implementation and hyper-parameter choice to different games. We ran the same training procedure on the following games:

Experiment #	Game
0	Beam Rider
1	Breakout
2	Enduro
3	Space Invaders

Table 1: Game tested in each experiment

The following graph shows the learning curve obtained in these experiments (each graph is normalized by the maximal value since different games have different reward scales)

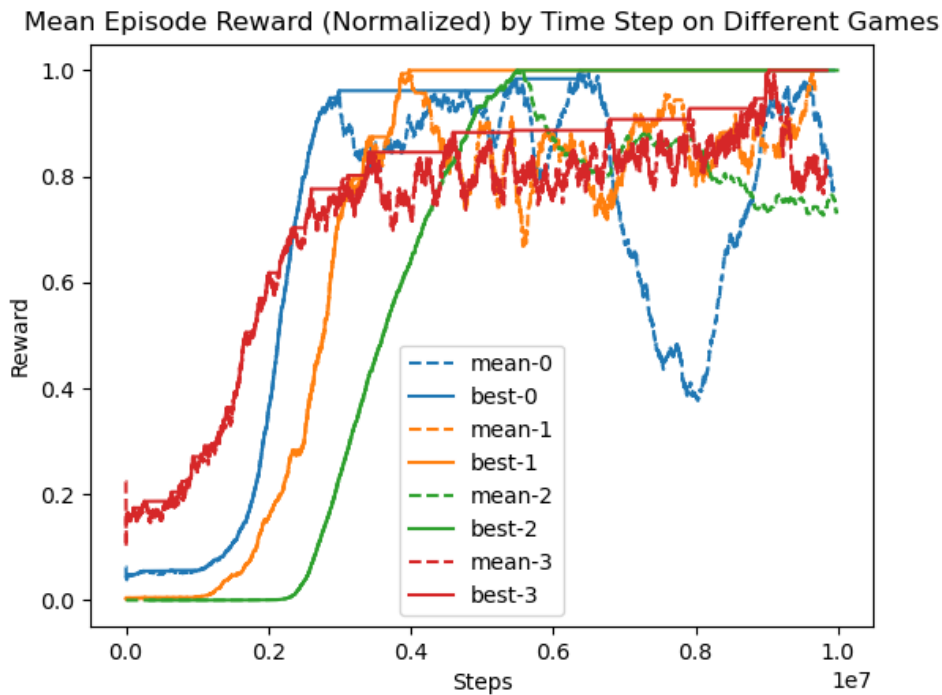


Figure 2: DQN training on the games described in Table 1. We can see that the implementation and hyper-parameters are robust to multiple games.