Homework - 2 : DQN

KRITH SANVITH MANTHRIPRAGADA
KM59

I. Introduction

THIS document contains results from four cases corresponding to the DQN Algorithm. Due to very high computational time, all the cases were not taken any average value, I could just run each case once. To decrease the computational time even more, I had to run the replay memory after storing all the states in the memory rather than incorporating replay memory at every step. These two things might cause the plots to be different from the actual plots

II. RESULTS FOR DQN

A. Standard Algorithm

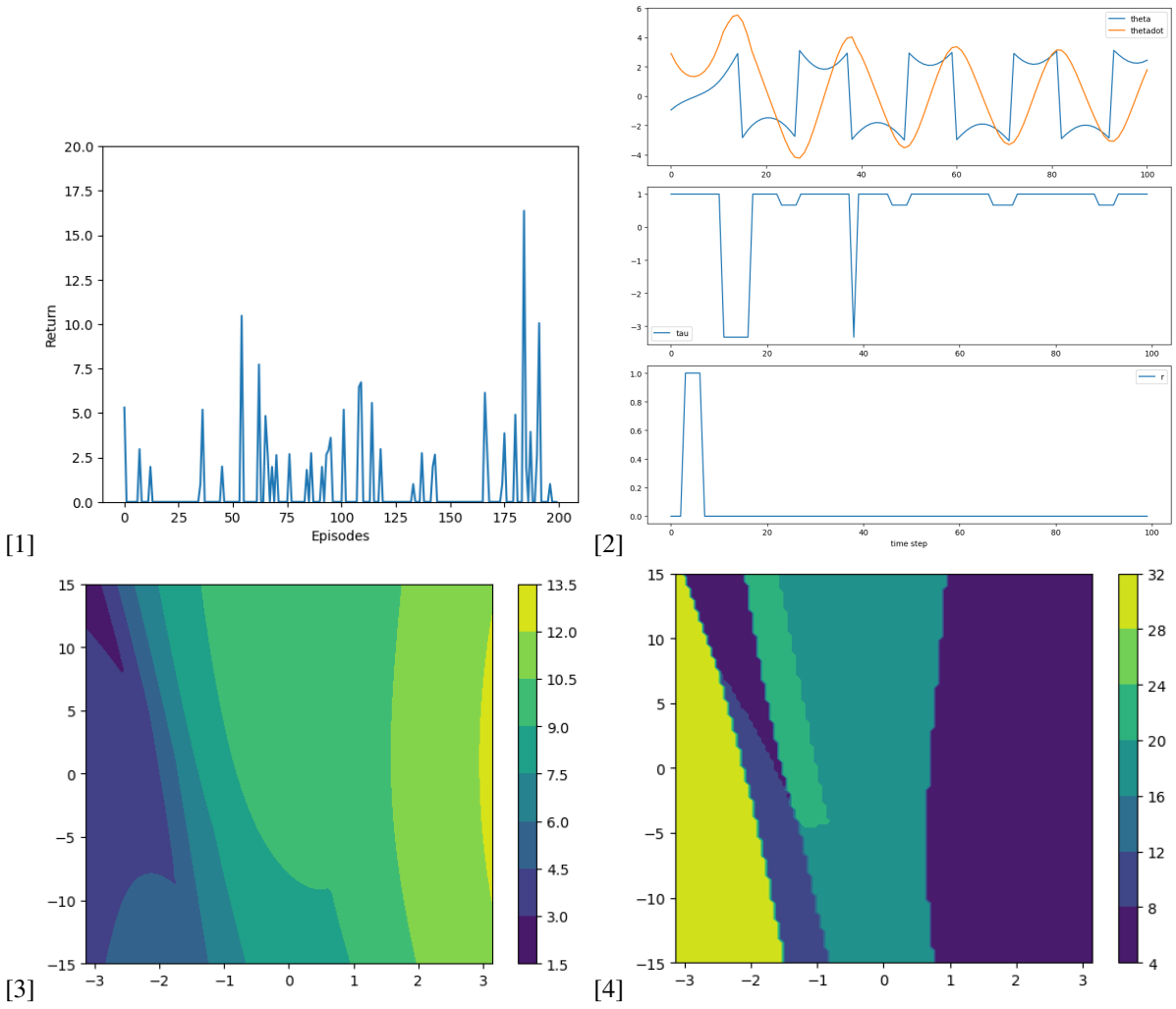
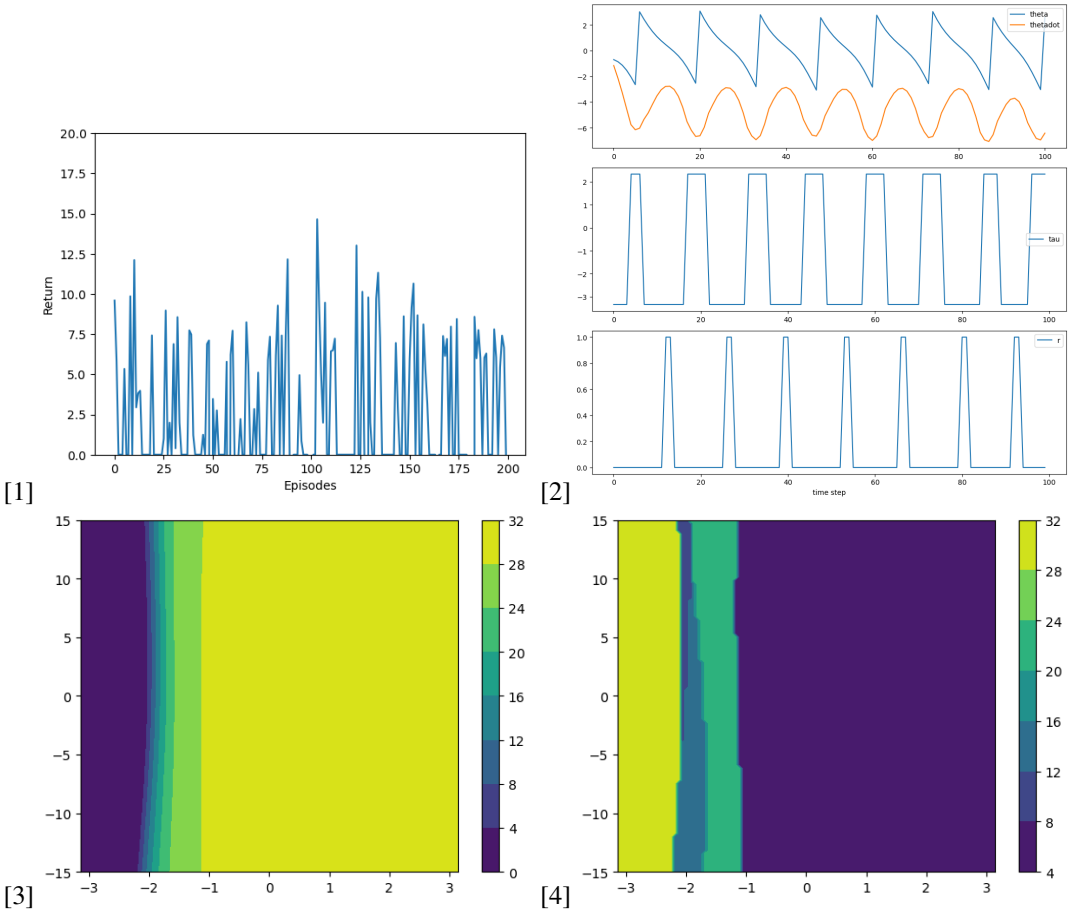


Fig. 1 (a) Learning Curve (b) Trajectory (c) State Value function (d) Policy

Initially, incorporating replaymemory at every step and looping the algorithm four times took around 93mins to run, in order to decrease it, I had to remove the loop and incorporate the replay memory outside the while not done loop. My laptop could barely handle this computation, almost 4 times my laptop crashed and had to restart to run even this case. The standard algorithm was run for 200 episodes with both replay buffer with batch size of 32 and target-Q network getting updated every 33 episodes.

B. With Replay and Without Target-Q



[1] [2] [3] [4]

Fig. 2 (a) Learning Curve (b) Trajectory (c) State Value function (d) Policy

Here we have the target-Q network getting updated after every episode, ideal scenario would we updating the target-Q every step but due to the computational inefficiency, I had to update it every episode.

C. Without Replay and With Target-Q

This case is without replay ie the size of the replay buffer is same as the size of the minibatch. The target-q network is getting updated after every one-third of episodes.

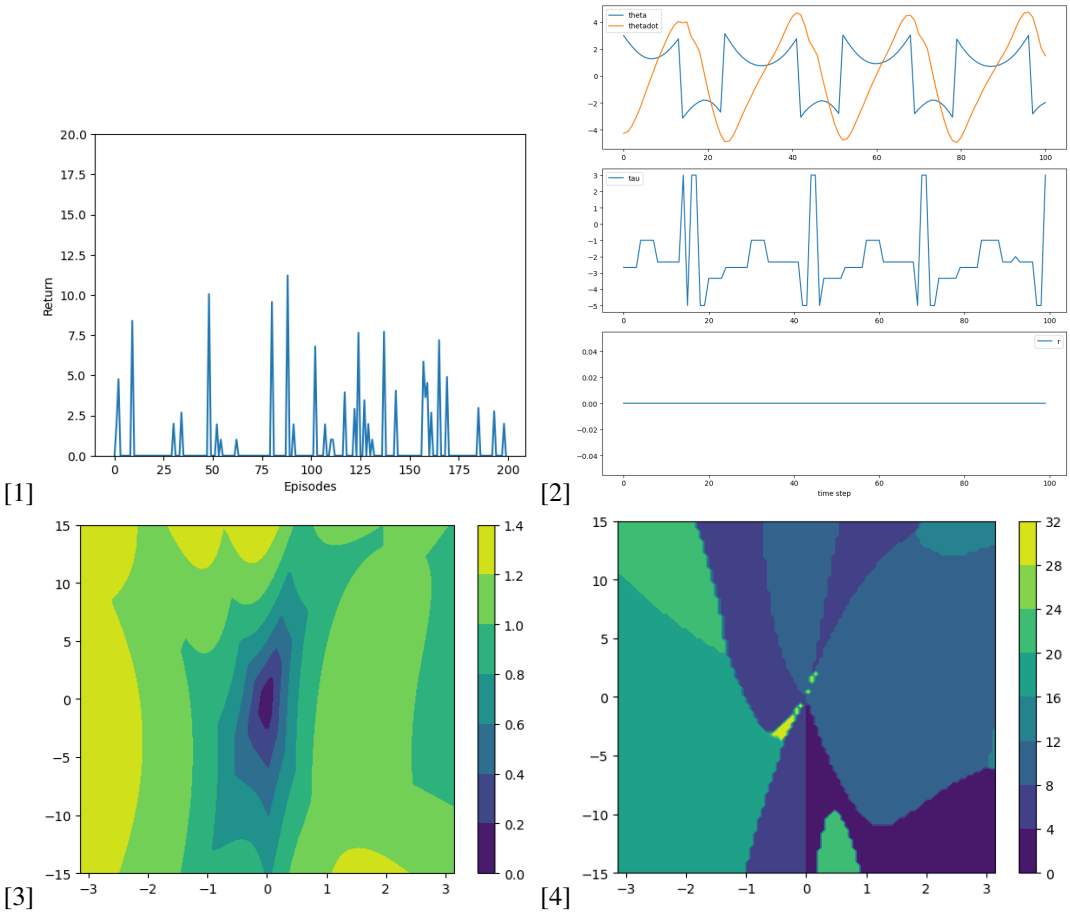
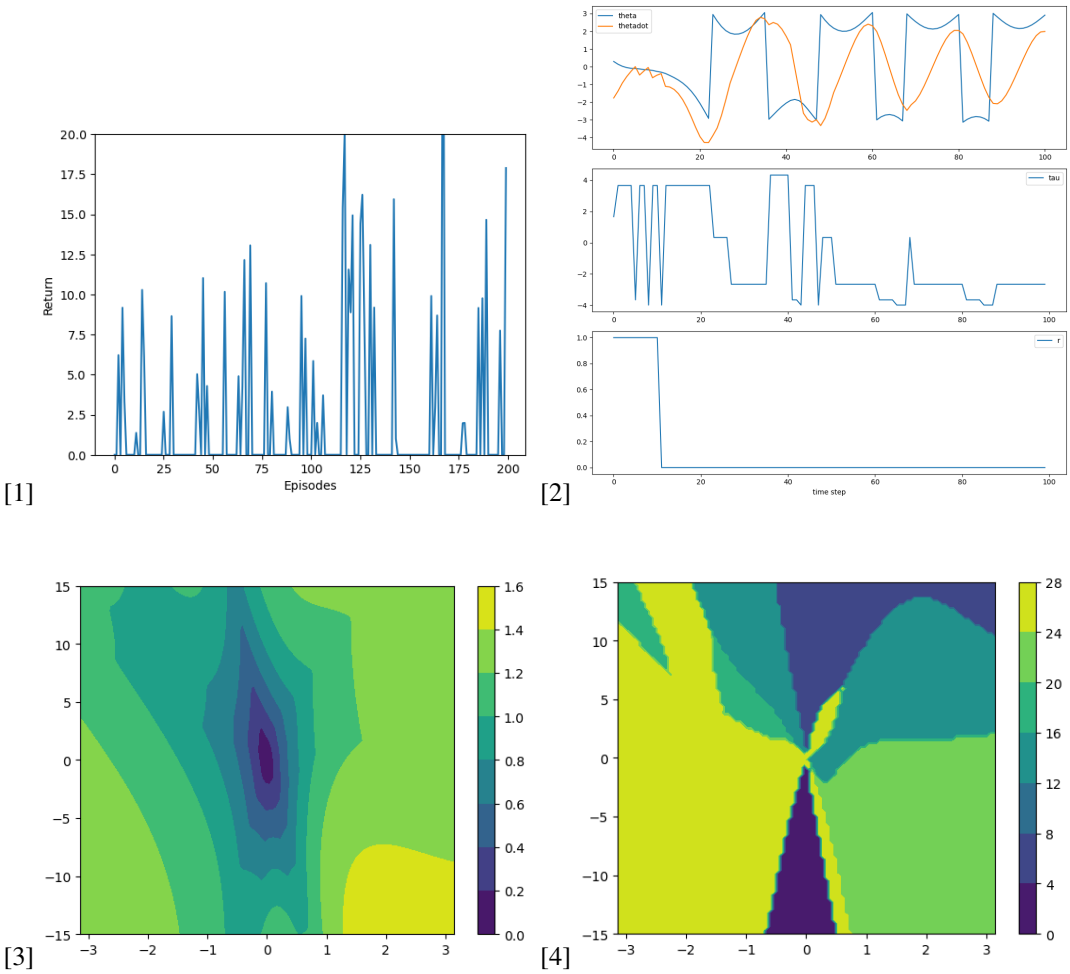


Fig. 3 (a) Learning Curve (b) Trajectory (c) State Value function (d) Policy

D. Without Replay and Without Target-Q



[1] [2] [3] [4]

Fig. 4 (a) Learning Curve (b) Trajectory (c) State Value function (d) Policy

Here, in this case the replay buffer memory has the same size that of the mini batch and also the target-q network gets updated every episode.