1 DQN

1.1 Question 1

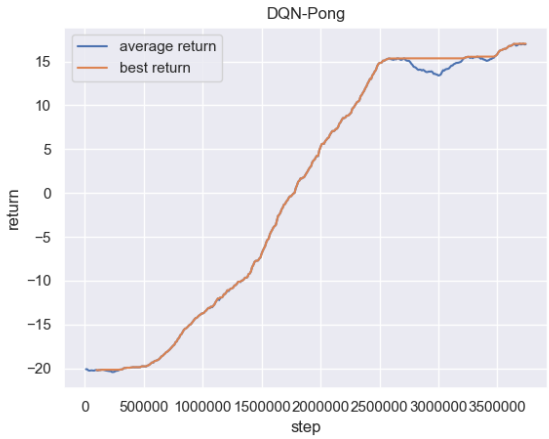

Figure 1: DQN on Pong.

Please see Figure 1 for performance of DQN on Pong.

1.2 Question 2

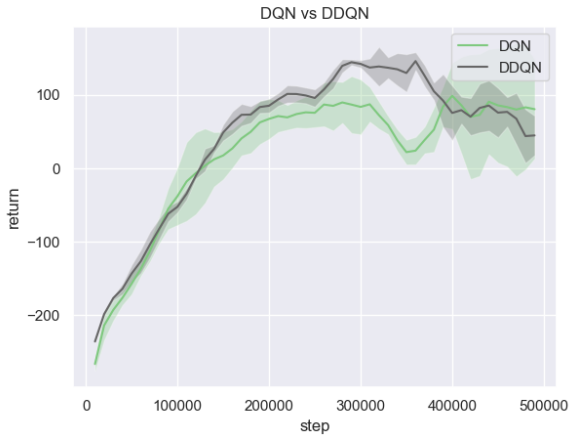

Figure 2: DQN vs. DDQN on LunarLander. Results are averaged over 3 seeds.

Please see Figure 2. DDQN performance is much better than DQN.

1.3 Question 3

I tried modifying `target_update_freq`. The results are shown in Figure 3.

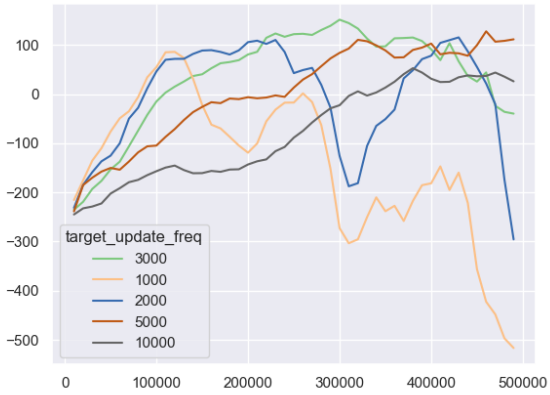

Figure 3: Effects of target update frequency. Large frequency makes learning more stable but slower, but small frequency makes learning very unstable.

2 Actor-Critic

2.1 Question 4

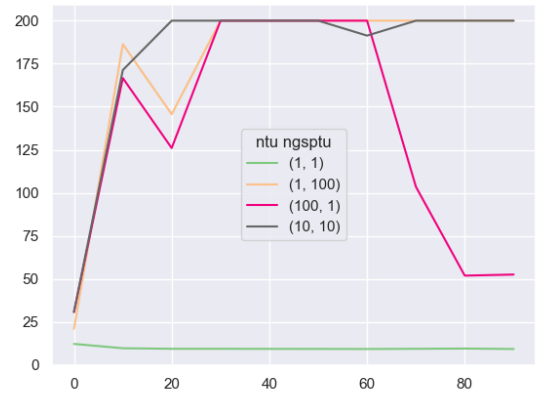

Figure 4: Actor-Critic on CartPole. This indicates that using slow target update is important for stable learning, but it doesn't have to be too large. Intuitively, more gradient update is always helpful.

Please see Figure 4. Takeaway: using slow target update is important for stable learning, but it doesn't have to be too large. Intuitively, more gradient update is always helpful.

2.2 Question 5

Please see Figure 5.

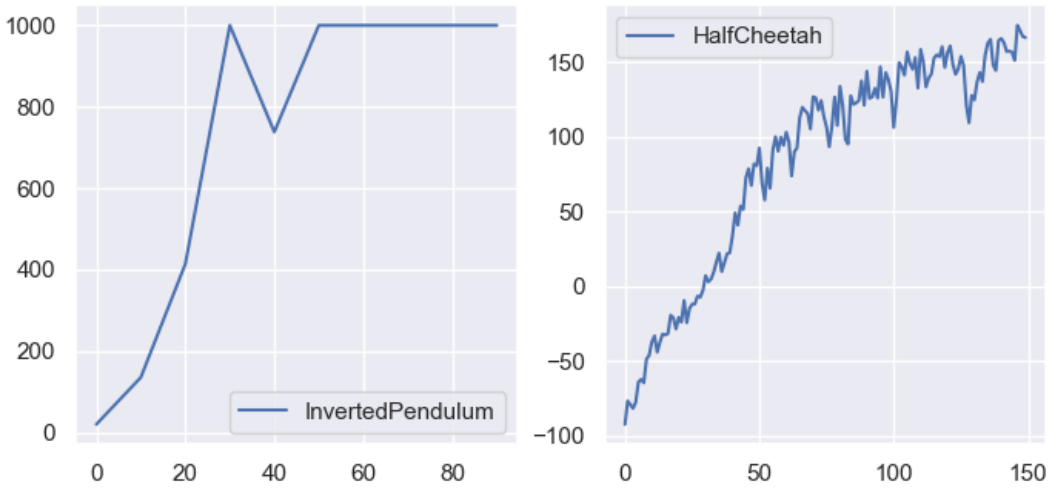

Figure 5: Results on InvertedPendulum and HalfCheetah.