CS285 - HW3

Ho Nam Nguyen

October 18, 2022

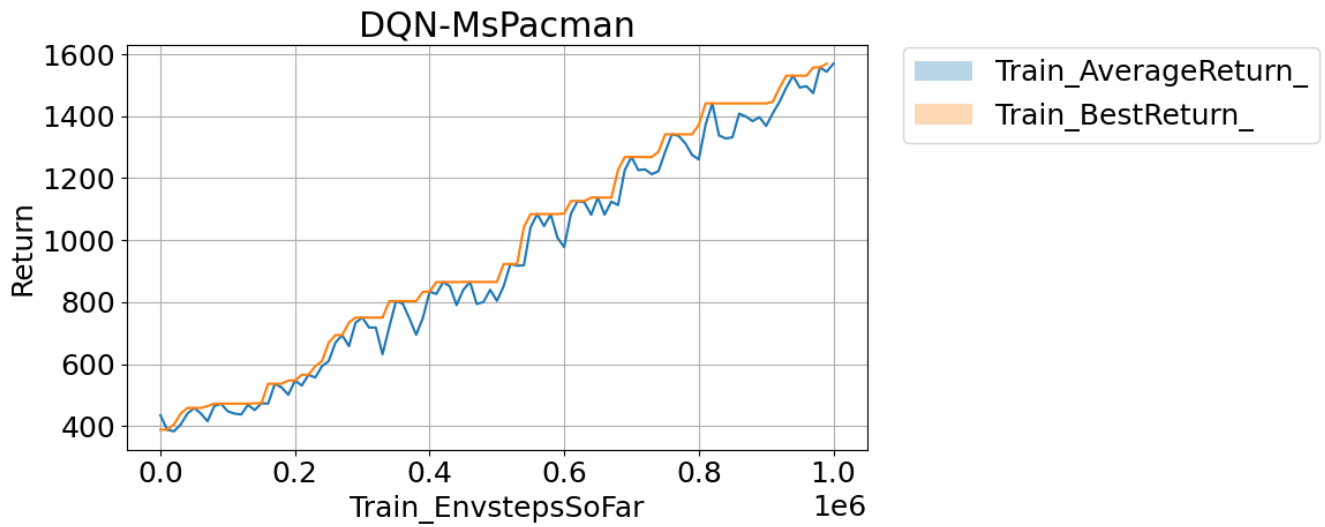## 1 Basic Q-learning performance (DQN)

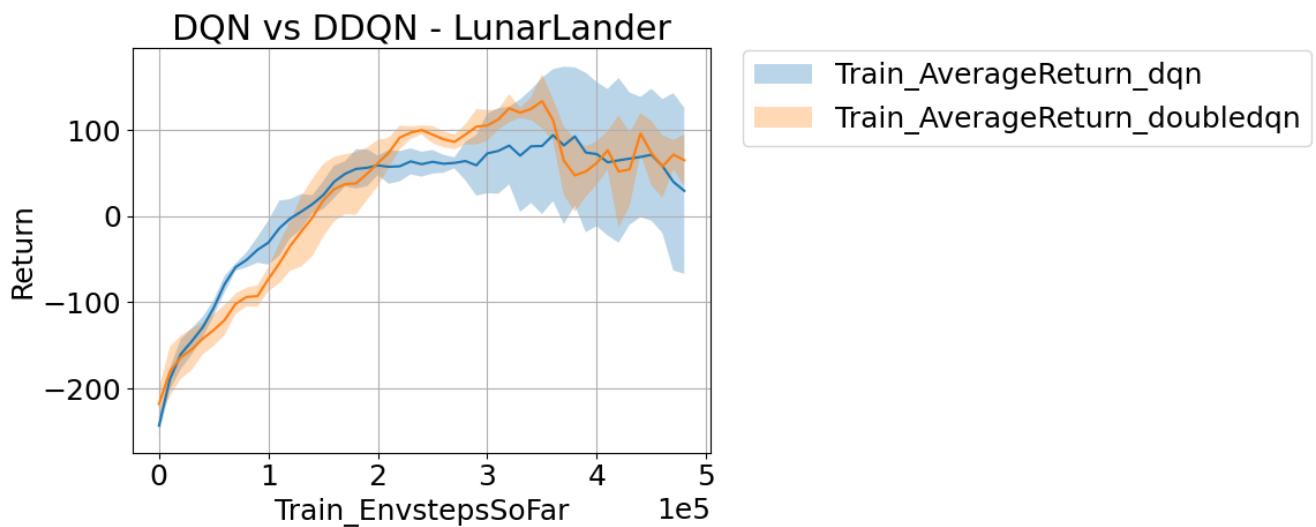Figure 1

2 Double Q-learning (DDQN)
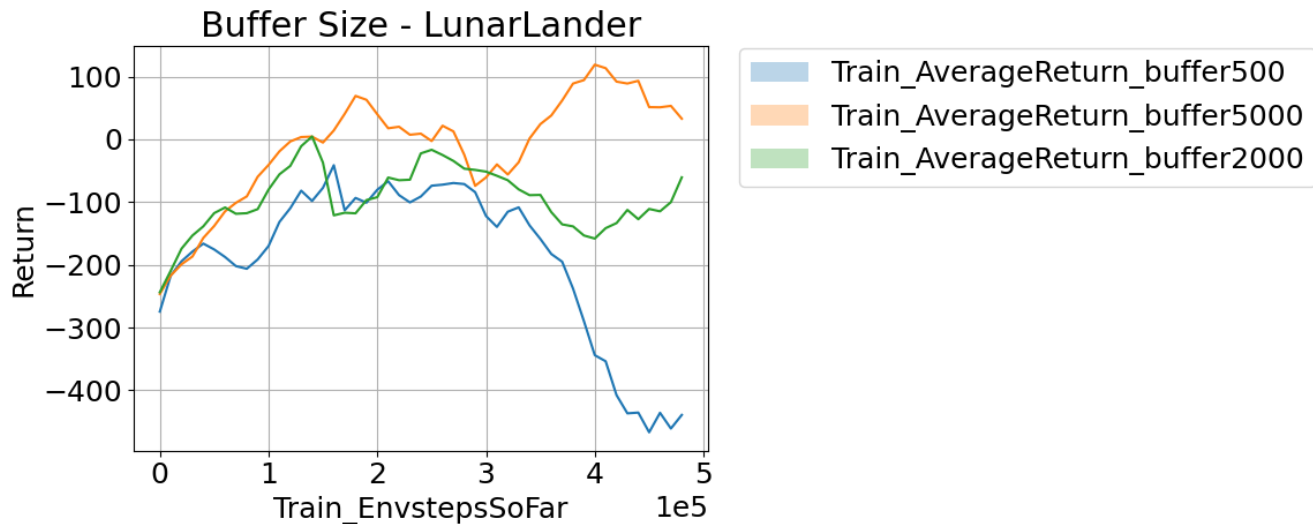

Figure 2

3 Hyperparameters - DQN

Figure 3: I vary the buffer size and observe that with too small of a buffer, transitions are getting thrown out to fast that little to no learning is possible. The default size is 50000 so decreasing this 10 times actually has little effect on the performance, i.e. orrange curve still achieves a score of 100.
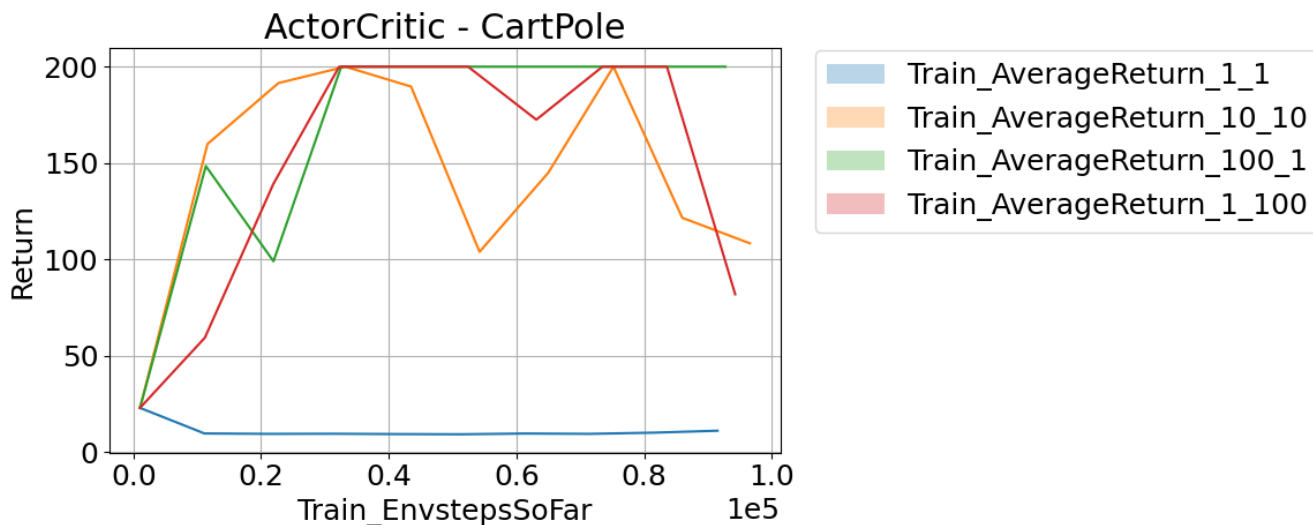
4 Actor-Critic sanity check
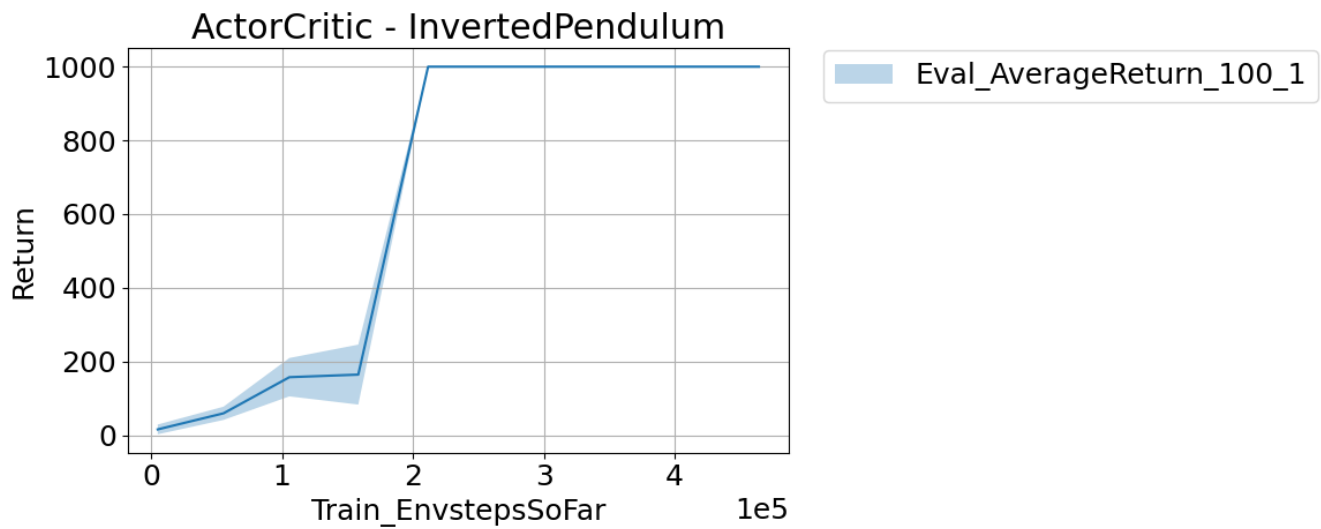

Figure 4

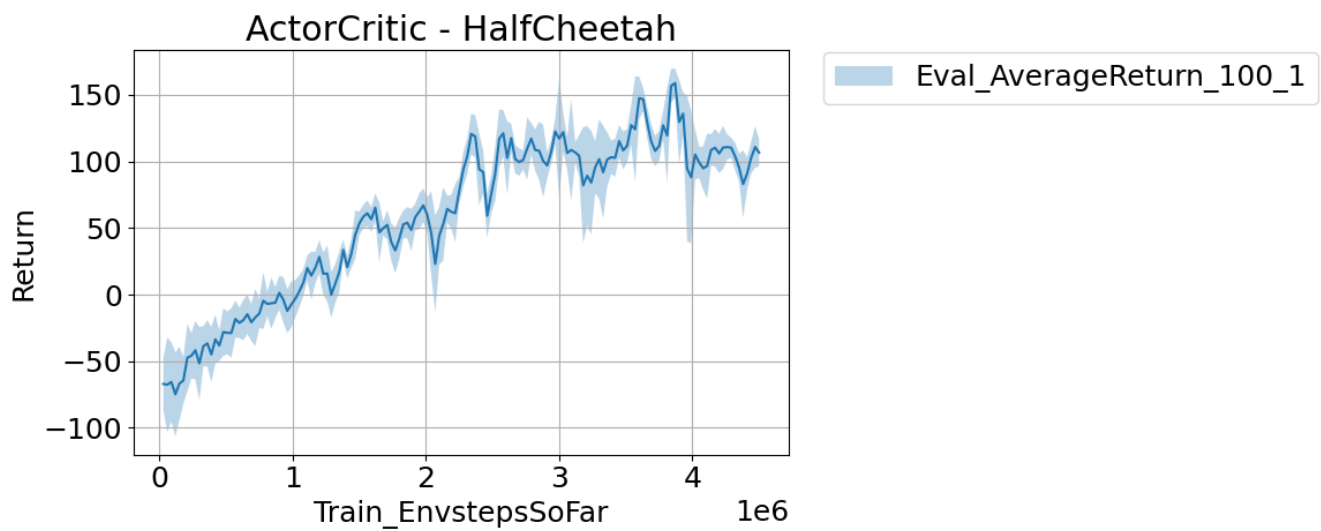# 5 Actor-Critic on difficut tasks
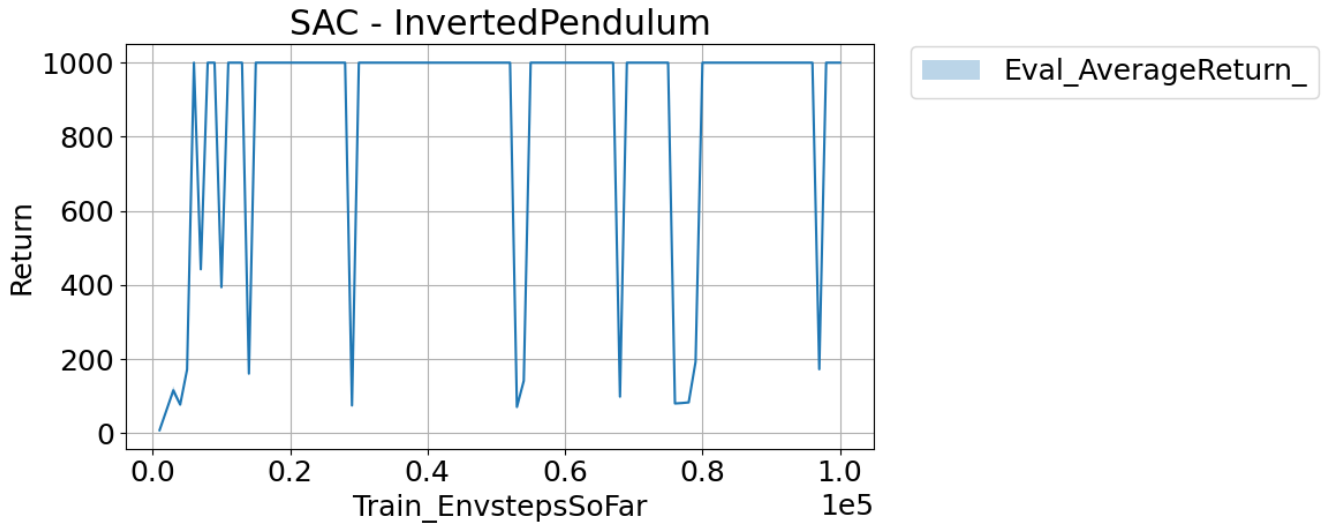
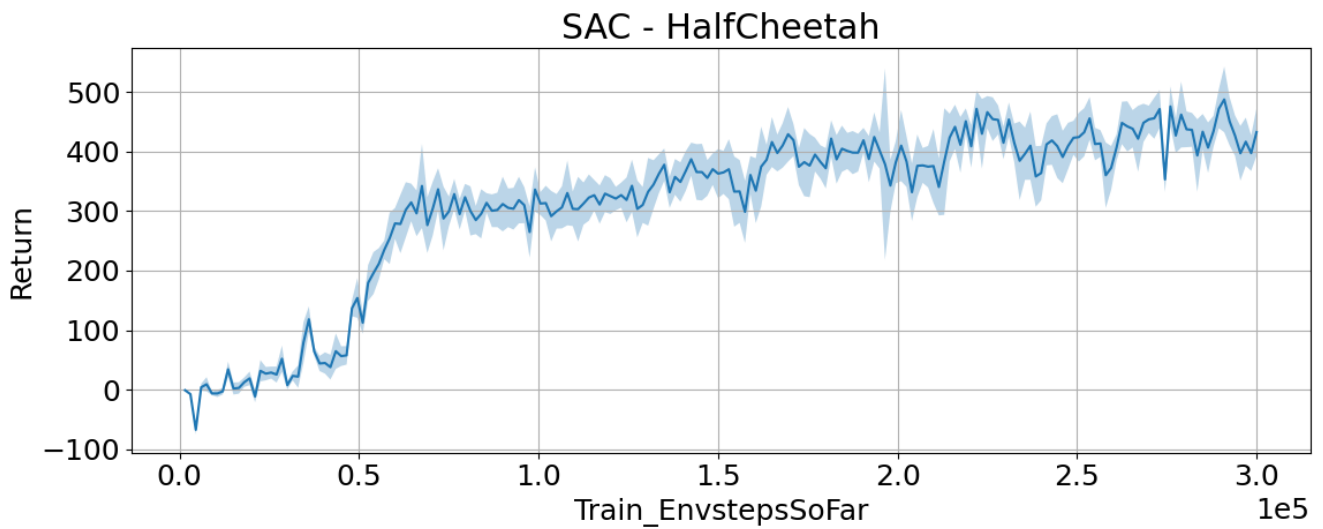

Figure 5

Figure 6

6   Soft Actor-Critic

Figure 7

Figure 8