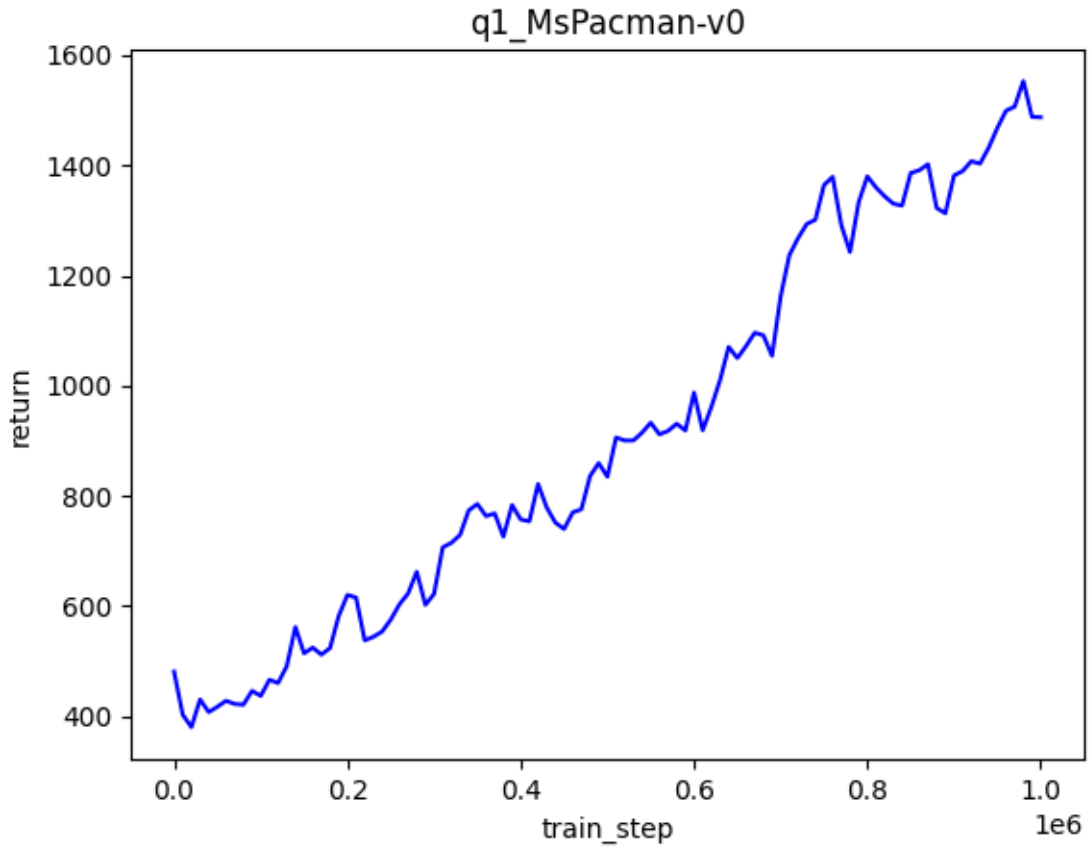
**CS285 Homework 3**

Cuiqianhe Du

**Question 1**

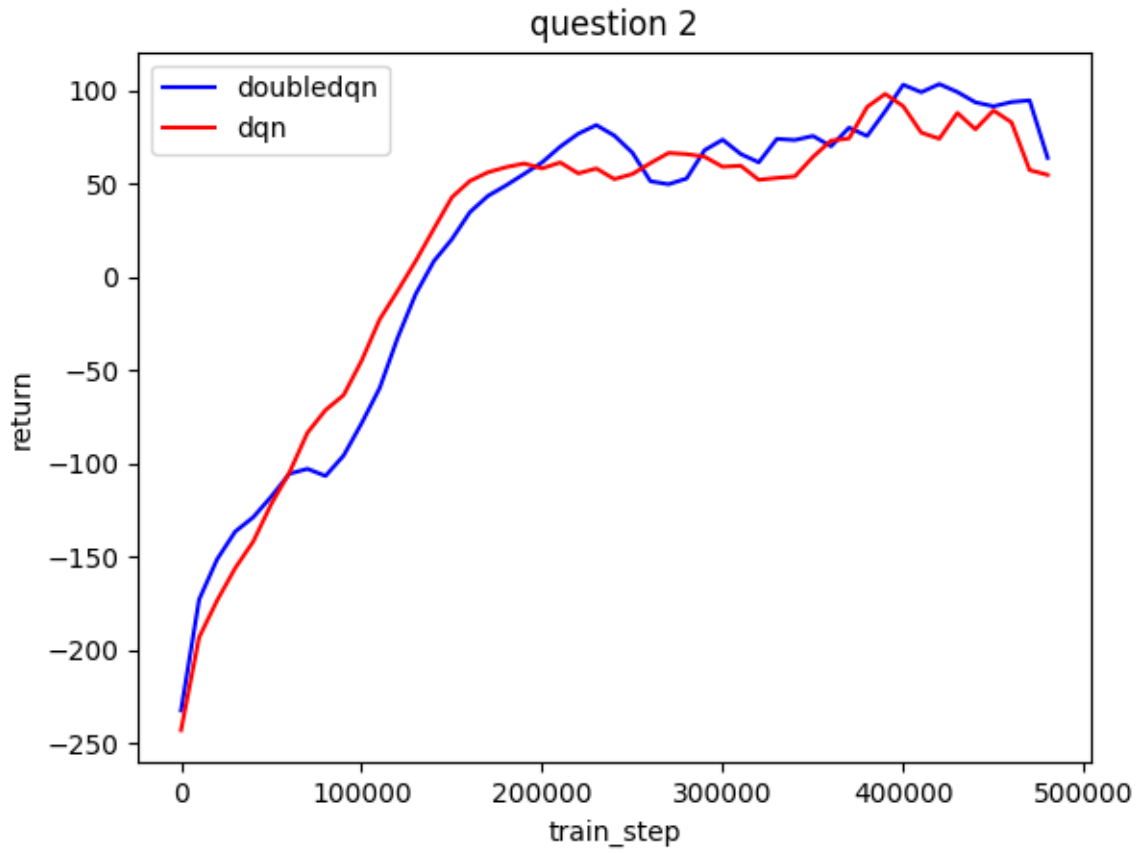
python cs285/scripts/run_hw3_dqn.py --env_name MsPacman-v0 --exp_name q1



**Question 2**

python cs285/scripts/run_hw3_dqn.py --env_name LunarLander-v3 --exp_name q2_dqn_1 --seed 1
python cs285/scripts/run_hw3_dqn.py --env_name LunarLander-v3 --exp_name q2_dqn_2 --seed 2
python cs285/scripts/run_hw3_dqn.py --env_name LunarLander-v3 --exp_name q2_dqn_3 --seed 3

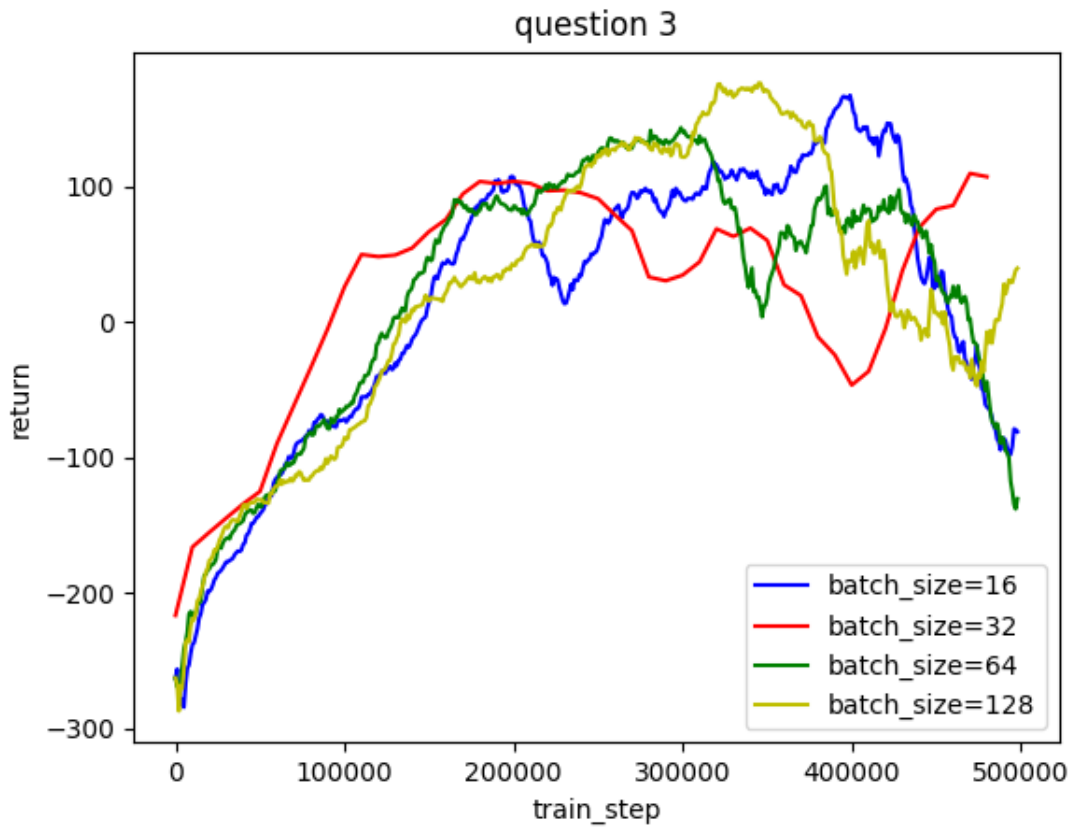
python cs285/scripts/run_hw3_dqn.py --env_name LunarLander-v3 --exp_name q2_doubledqn_1 --double_q --seed 1
python cs285/scripts/run_hw3_dqn.py --env_name LunarLander-v3 --exp_name q2_doubledqn_2 --double_q --seed 2
python cs285/scripts/run_hw3_dqn.py --env_name LunarLander-v3 --exp_name q2_doubledqn_3 --double_q --seed 3



**Question 3**

Command from question 1 with batch size 32:
python cs285/scripts/run_hw3_dqn.py --env_name LunarLander-v3 --exp_name q1

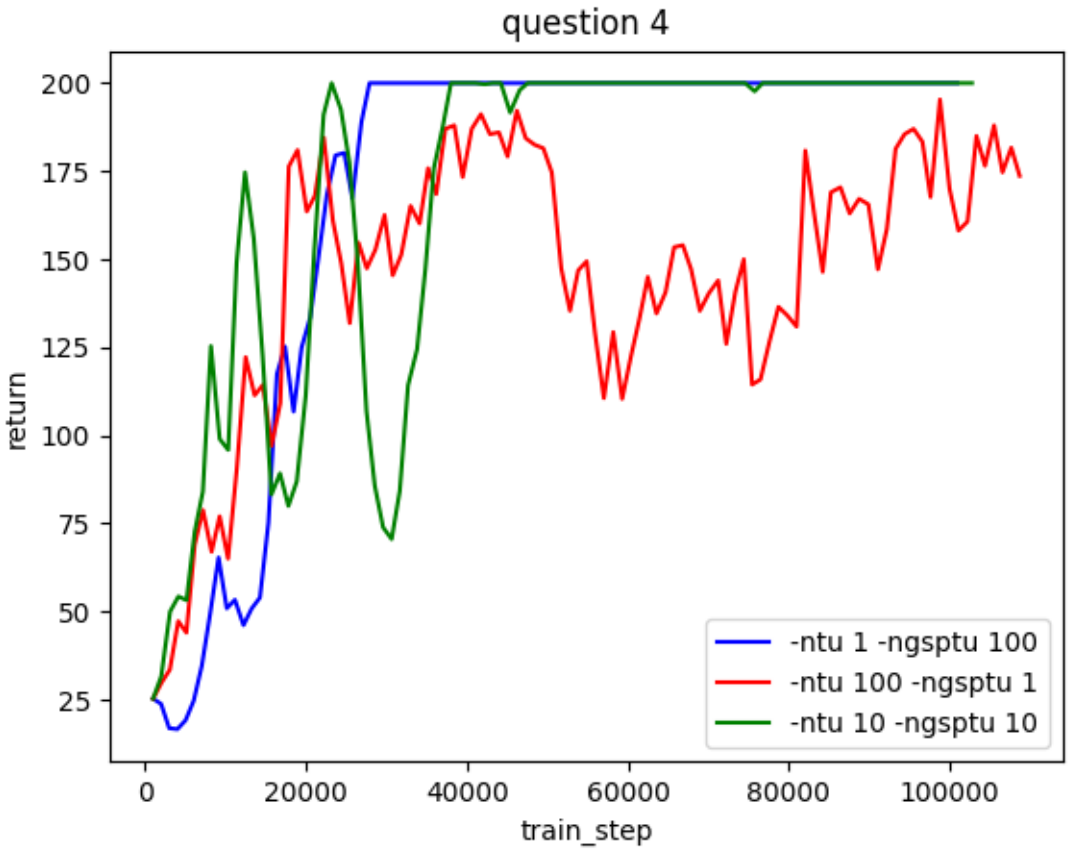
Command for question 3:
python cs285/scripts/run_hw3_dqn.py --env_name LunarLander-v3 --exp_name q3_batch_size_16 --batch_size 16
python cs285/scripts/run_hw3_dqn.py --env_name LunarLander-v3 --exp_name q3_batch_size_64 --batch_size 64
python cs285/scripts/run_hw3_dqn.py --env_name LunarLander-v3 --exp_name q3_batch_size_128 --batch_size 128




I choose to change batch size; I find out the performance is generally better with larger batch size.

**Question 4**

python cs285/scripts/run_hw3_actor_critic.py --env_name CartPole-v0 -n 100 -b 1000 --exp_name q4_100_1 -ntu 100 -ngsptu 1
python cs285/scripts/run_hw3_actor_critic.py --env_name CartPole-v0 -n 100 -b 1000 --exp_name q4_1_100 -ntu 1 -ngsptu 100
python cs285/scripts/run_hw3_actor_critic.py --env_name CartPole-v0 -n 100 -b 1000 --exp_name q4_10_10 -ntu 10 -ngsptu 10




From the graph we could learn that with ntu 100, ngsptu 1, the performance is the worst. Even though both (ntu 1, ngsptu100), (ntu 10, ngsptu10) could reach 200 eventually, (ntu 10, ngsptu10) is a quicker learner and could reach better performance in the starting stage compare with the other at the same step.

**Question 5:**

python cs285/scripts/run_hw3_actor_critic.py --env_name InvertedPendulum-v2 --ep_len 1000 --discount 0.95 -n 100 -l 2 -s 64 -b 5000 -lr 0.01 --exp_name q5_10_10 -ntu 10 -ngsptu 10
python cs285/scripts/run_hw3_actor_critic.py --env_name HalfCheetah-v2 --ep_len 150 --discount 0.90 --scalar_log_freq 1 -n 150 -l 2 -s 32 -b 30000 -eb 1500 -lr 0.02 --exp_name q5_10_10 -ntu 10 -ngsptu 10

I'm using -ntu 10 -ngsptu 10 combination because this has the best performance in question 4