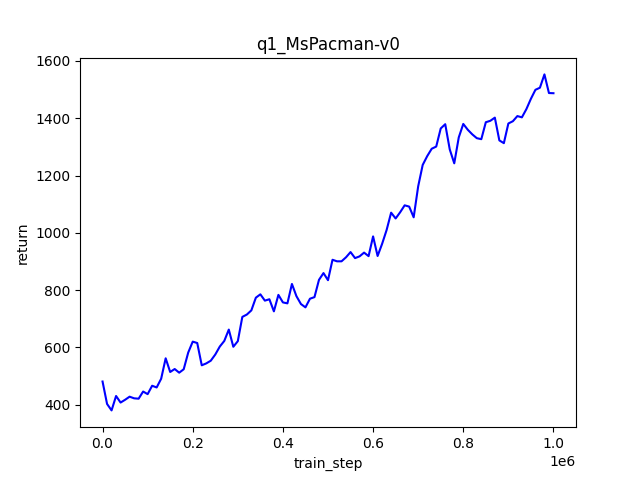
**CS285 Homework 3**

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

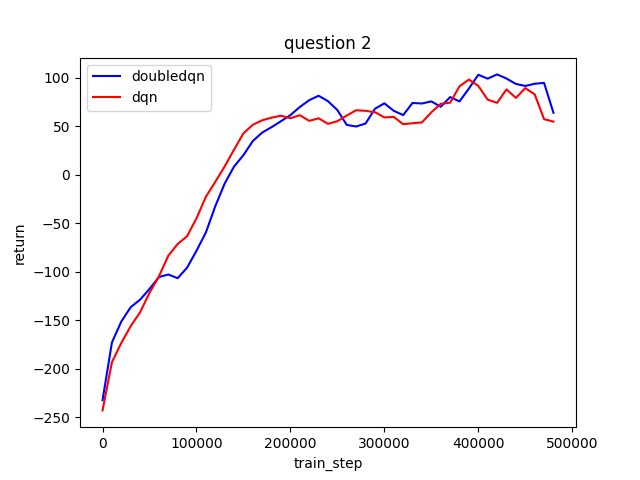
python cs285/scripts/run\_hw3\_dqn.py --env\_name LunarLander-v3 --exp\_name q2\_dqn\_2 --seed 2

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:

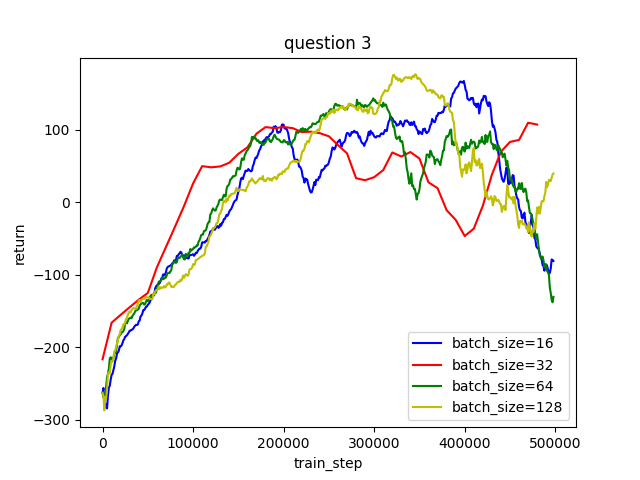
python cs285/scripts/run\_hw3\_dqn.py --env\_name LunarLander-v3 --exp\_name q1

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



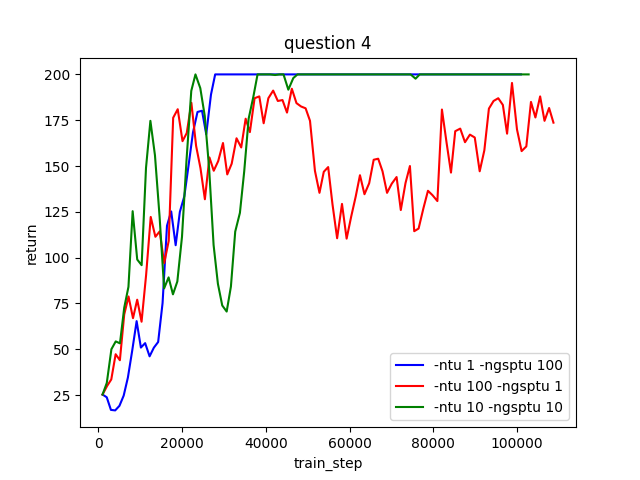
I choose to change batch size; I find out the performance is generally better with larger batch size.

**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

