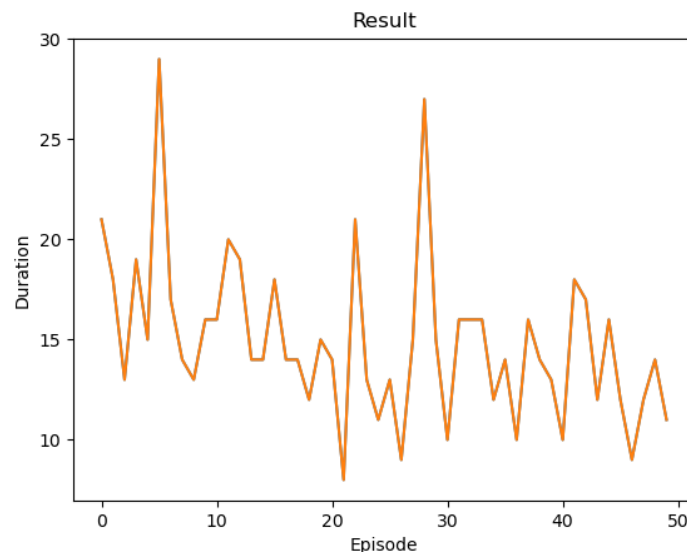


EE5885 HW2

A. Implement the Deep Q Network algorithm

Implementing deep q learning on a Cart-pole problem. The code can be found on https://pytorch.org/tutorials/intermediate/reinforcement_q_learning.html

1. (20 points) Build the tool chain. You can either use the tool chain recommended in the class lecture, i.e., Anaconda + Pytorch +Pycharm, or use your own favorite tool chain. The goal is to implement the given code and obtain a duration-episode plot similar to below.



2. (10 points) Based on the original code, change the episode number to 1000. Plot the duration-episode plot. Make observations and try to explain your observations.
3. (10 points) Based on the question 2 code, change the value function discount rate to 0.89. Plot the duration-episode plot. Make observations and try to explain your observations.
4. (10 points) Based on the question 2 code, change the mini-batch size of the experience pool to 1500. Plot the duration-episode plot. Make observations and try to explain your observations.
5. (10 points) Based on the question 2 code, change the neural network learning rate to 1e-2. Plot the duration-episode plot. Make observations and try to explain your observations.

B. Cliff Walk

Read the cliff walk example from the lecture slides or in Richard Sutton's "Reinforcement Learning: An Introduction" book. Write SARSA and Q-learning code to compare the episodic sum of rewards.

1. (20 pts) Try $\gamma = 0.01, 0.1, 0.5, 0.99, 1$ and plot the episodic sum of rewards. Try to explain.
2. (20 pts) Under what circumstances Q-learning and SARSA will have the same episodic sum of rewards?

Deliverable:

The results should be submitted electronically via WyoCourses. Please create a ZIP file include your four code and a report including the following items.

- Notice of what lines of code were modified from the original code.
- The answers to questions 2-5.
- Plots for all questions.