# Homework 4:

# Reinforcement Learning

# Report Template

**Please keep the title of each section and delete examples. Note that please keep the questions listed in Part II.**

## Part I. Experiment Results (the score here is included in your implementation):

**Please paste taxi.png, cartpole.png, DQN.png and compare.png here.**

**1. taxi.png:**

**2. cartpole.png**

**3. DQN.png**

**4. compare.png**

## Part II. Question Answering (50%):

1. Calculate the optimal Q-value of a given state in Taxi-v3 (the state is assigned in google sheet), and compare with the Q-value you learned (Please screenshot the result of the "check_max_Q" function to show the Q-value you learned). **(4%)**

2. Calculate the max Q-value of the initial state in CartPole-v0, and compare with the Q-value you learned. (Please screenshot the result of the "check_max_Q" function to show the Q-value you learned) **(4%)**

3.

   a. Why do we need to discretize the observation in Part 2? **(2%)**

   b. How do you expect the performance will be if we increase “num_bins”? **(2%)**

   c. Is there any concern if we increase “num_bins”? **(2%)**

4. Which model (DQN, discretized Q learning) performs better in Cartpole-v0, and what are the reasons? **(3%)**

5.

   a. What is the purpose of using the epsilon greedy algorithm while choosing an action? **(2%)**

   b. What will happen, if we don’t use the epsilon greedy algorithm in the CartPole-v0 environment? **(3%)**

   c. Is it possible to achieve the same performance without the epsilon greedy algorithm in the CartPole-v0 environment? Why or Why not? **(3%)**

   d. Why don’t we need the epsilon greedy algorithm during the testing section? **(2%)**

6. Why is there “`with torch.no_grad():`“ in the “choose_action” function in DQN? **(3%)**

7.

   a. Is it necessary to have two networks when implementing DQN? **(1%)**

   b. What are the advantages of having two networks? **(3%)**

   c. What are the disadvantages? **(2%)**

8.

a. What is a replay buffer(memory)? Is it necessary to implement a replay buffer? What are the advantages of implementing a replay buffer? **(5%)**

b. Why do we need batch size? **(3%)**

c. Is there any effect if we adjust the size of the replay buffer(memory) or batch size? Please list some advantages and disadvantages. **(2%)**

9.

a. What is the condition that you save your neural network? **(1%)**

b. What are the reasons? **(2%)**

10. What have you learned in the homework? **(2%)**