

## Project 4. Reinforcement Learning for Maze Problem

Contact: Yuzhuo Li [liyuzhuo23@mails.tsinghua.edu.cn](mailto:liyuzhuo23@mails.tsinghua.edu.cn)

### Background

Reinforcement Learning (RL) is a subset of machine learning where an agent learns how to behave in an environment to achieve a certain goal. The agent makes observations about the state of the environment, decides on an action based on its current policy, and then observes the outcome (reward or penalty) of that action. Over time, the agent improves its policy to maximize the cumulative reward and achieve the goal in the best way.

Reinforcement Learning has been widely applied across various domains. A few notable applications including game playing (like AlphaGo by DeepMind), robot control, and energy optimization have been well developed. In this homework, we will try to train an agent that can solve a maze problem.

### Task Description

OpenAI Gym is a toolkit for developing and benchmarking reinforcement learning algorithms (please find a detailed introduction at <https://github.com/openai/gym>). It provides a wide variety of environments to test algorithms in a standardized way.

In this project, we will use the Frozen Lake environment. The player (the agent) needs to cross a frozen lake from start to goal without falling into any holes by walking over the frozen lake. The player may not always move in the intended direction due to the slippery nature of the frozen lake. See more details in the project page:

[https://gymnasium.farama.org/environments/toy\\_text/frozen\\_lake/](https://gymnasium.farama.org/environments/toy_text/frozen_lake/).

### Task

1. Install gym, and try the Frozen Lake game.
2. Implement any two of three reinforcement learning algorithms listed below, to train an agent to play the Frozen Lake game with a self-designed 16\*16 map. Your objective is to find the optimal path over a specified number of episodes.

Algorithms: *Policy Gradient*, *Q-learning*, *PPO*

3. Use what you have learned to solve a tourism path planning problem on the map of Tsinghua University (see attached files). Read the following requirements, define the RL problem and train an agent that accomplish them.

Requirement:

- a) Consider two different types of visitors in the campus: those who prefer to walk and those who choose to drive. For those who choose to drive, there are specific

restrictions, which are clearly marked on the map. For those who walk, there are no extra restrictions.

- b) Visitors should enter and exit through any gate marked on the map, but note that certain gates are open for pedestrians only. At least 10 scenic spots should be intermediate points, and the tourists should pass a restaurant or cafeteria. (Disregard the issue of whether tourists can dine in the cafeteria)

Scenic spots example: West Gate - Lotus Pond - Gong Zi Ting - Sundial - Tsinghua Xue Tang - Auditorium - Tsinghua University Library(old) - East Sports Field - Yu Shu Yuan Canteen - Tsinghua University Art Museum - Main Building – New Tsinghua Auditorium - South Gate

- c) Valid roads, school gates, possible spots and restaurants, and vehicle forbidden marks are recorded in the attached file ‘Tsinghua map net grid.jpg’. It should be your guidance to construct your map, while overlapped version may help to understand each spot.
- d) (option)Spots will have variant values (self-decided attractions) and due to the time limit the tourists may only visit few of them. Train an agent to have a tour in Tsinghua. Your objective now is to find the optimal path with highest values over a specified number of episodes.

NOTE: The grid map aims to specify the maze, you don’t have to visualize the process.

Attached files:

- A. Tsinghua map net background.jpg
- B. Tsinghua map net grid.jpg
- C. Tsinghua map net coords.jpg



Tsinghua map net background.jpg



& Tsinghua map net coords.jpg