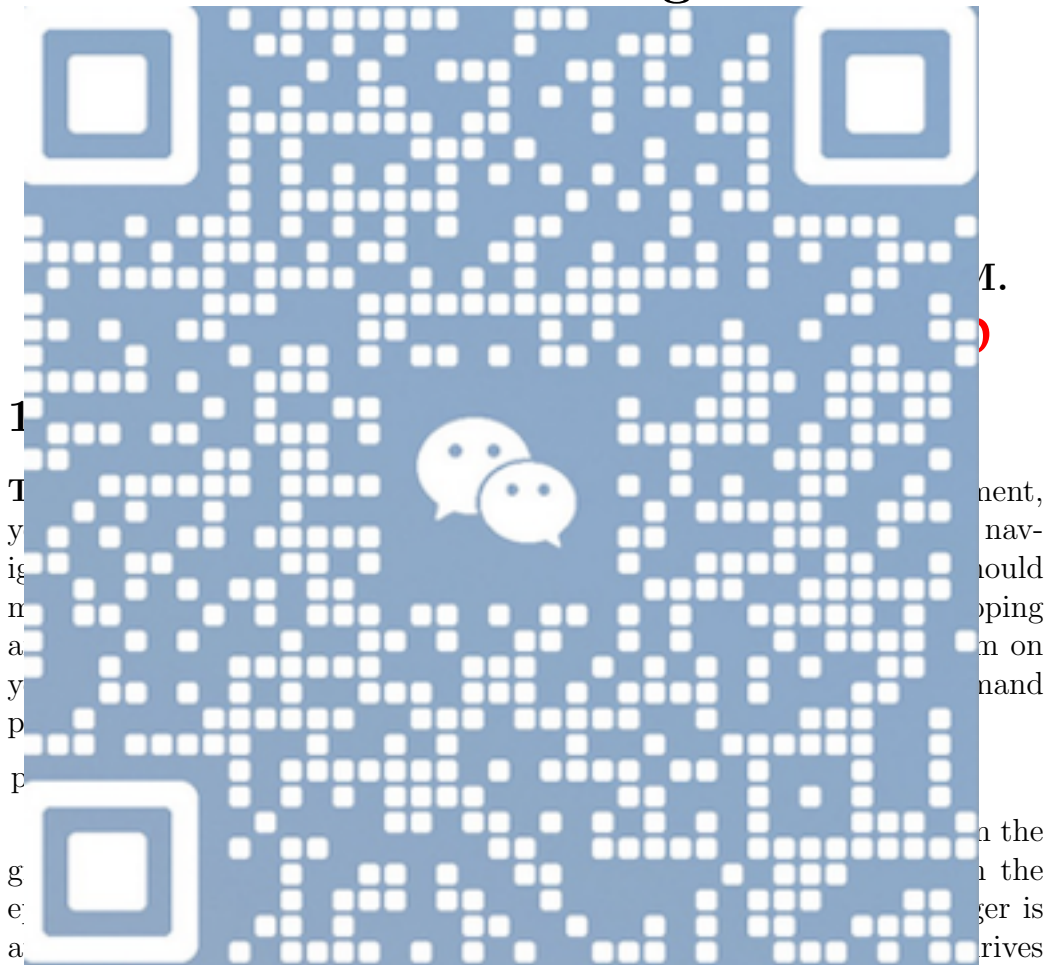


COMP9414 24T2

Artificial Intelligence

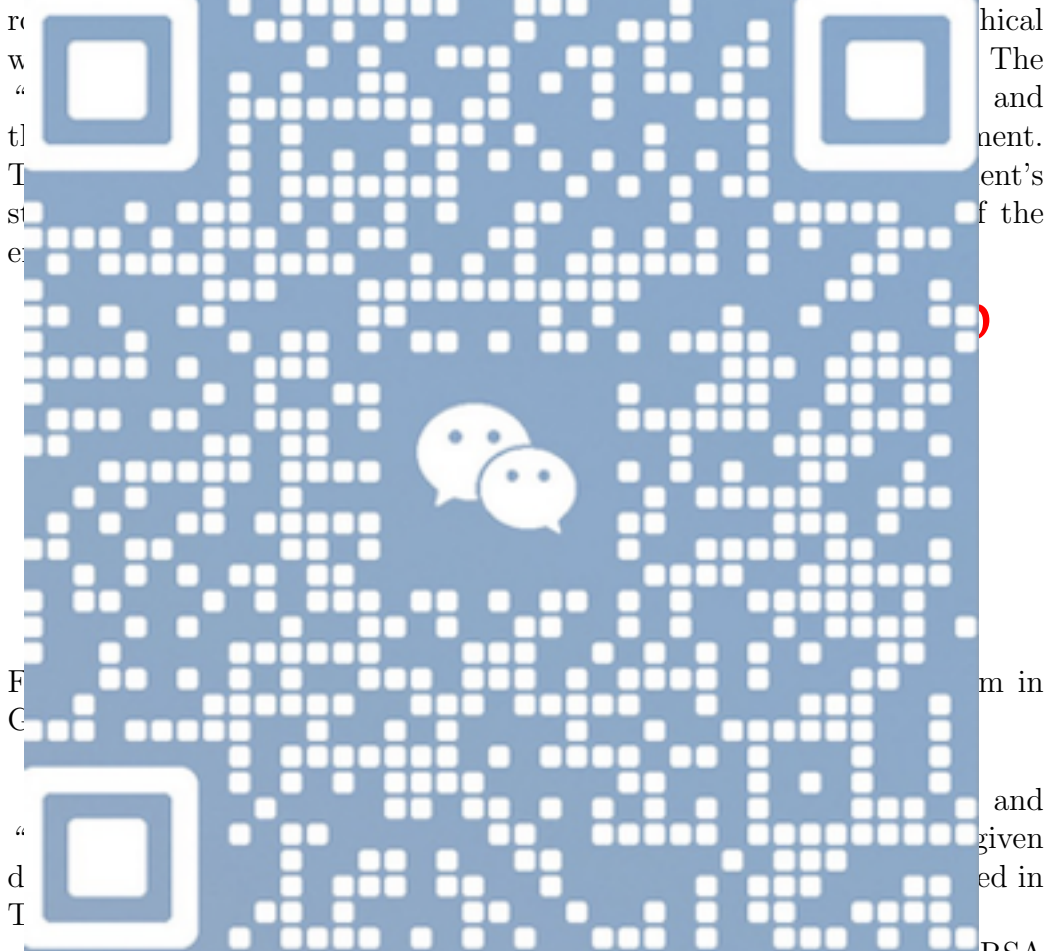


to the passenger's location, picks up the passenger, drives to the passenger's destination (another one of the four specified locations), and then drops off the passenger. Once the passenger is dropped off, the episode ends. To show the taxi grid world environment, you can use the following code:

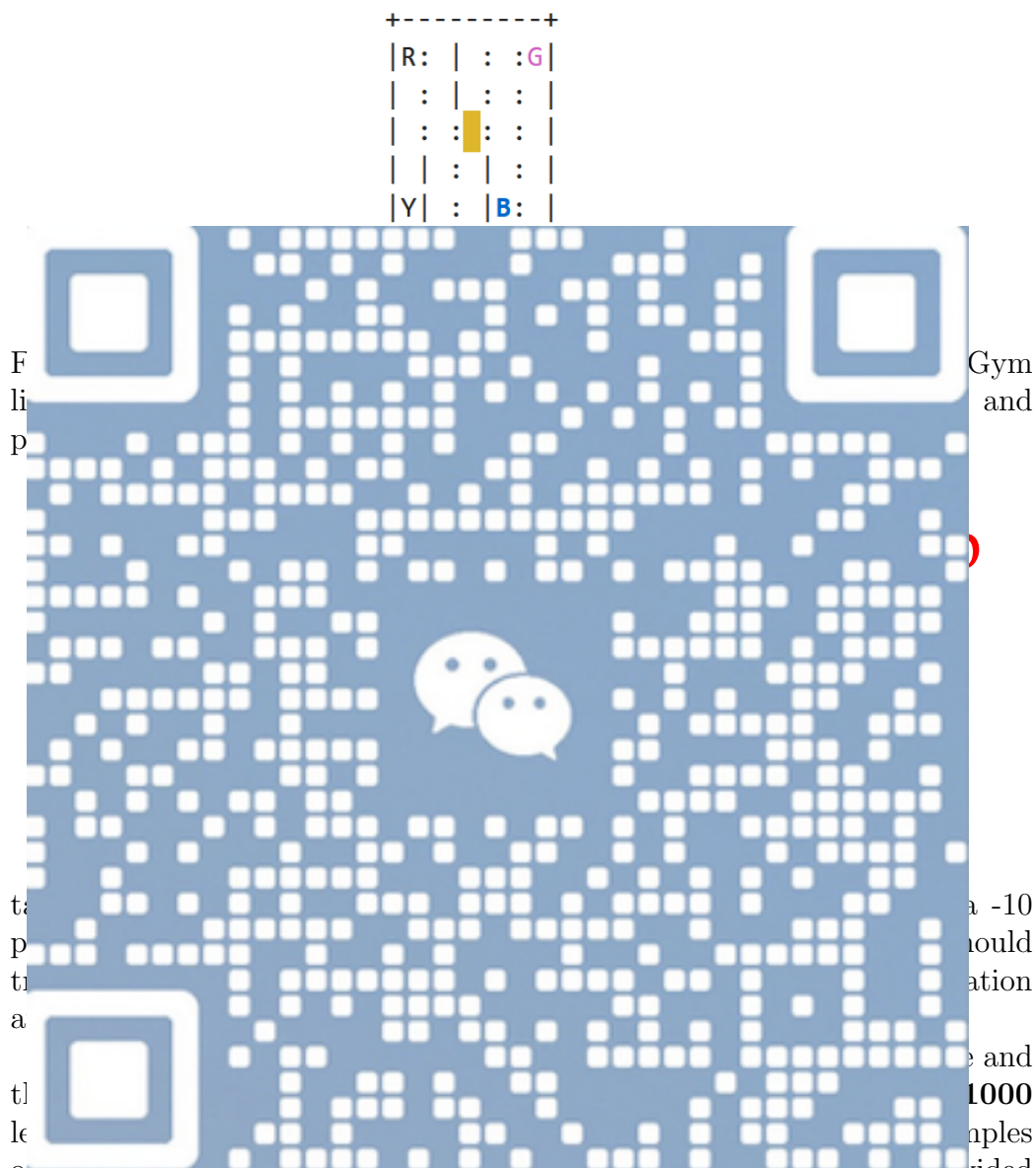
¹https://www.gymnasium.dev/environments/toy_text/taxi/

```
env = gym.make("Taxi-v3", render_mode="ansi").env
state = env.reset()
rendered_env = env.render()
print(rendered_env)
```

In order to render the environment, there are three modes known as “human”, “rgb_array”, and “ansi”. The “human” mode visualizes the envi-

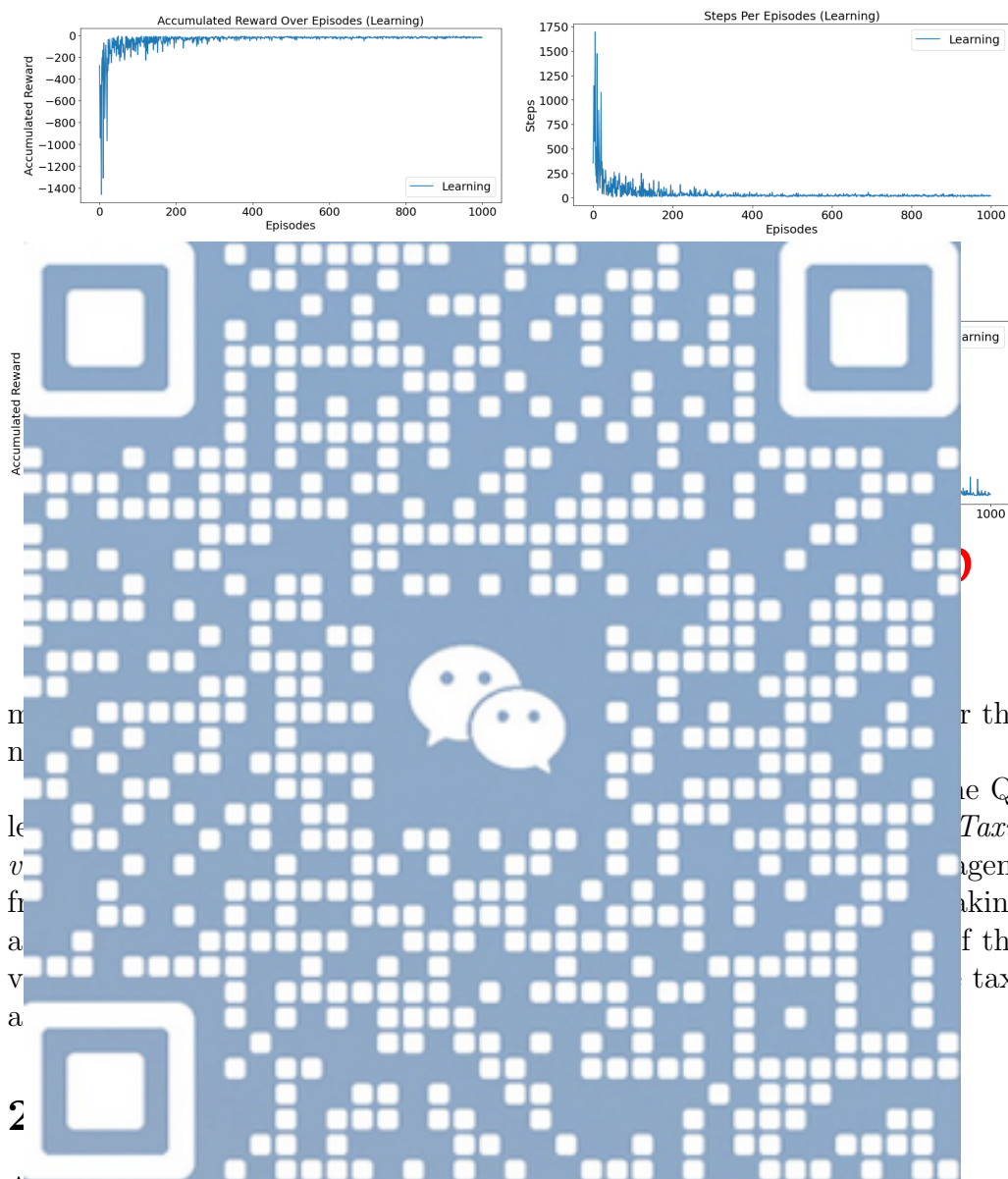


For this assignment, you need to implement the Q-learning and SARSA algorithms for the taxi navigation environment. The main objective for this assignment is for the agent (taxi) to learn how to navigate the grid-world and drive the passenger with the minimum possible steps. To accomplish the learning task, you should empirically determine hyperparameters, e.g., the learning rate α , exploration parameters (such as ϵ or T), and discount factor γ for your algorithm. Your agent should be penalized -1 per step it



of these two plots are shown in Figures 5-6. Please note that the provided plots are just examples and, therefore, your plots will not be exactly like the provided ones, as the learning parameters will differ for your algorithm.

After training your algorithm, you should save your Q-values. Based on your saved Q-table, your algorithms will be tested on at least 100 random grid-world scenarios with the same characteristics as the taxi environment for both the Q-learning and SARSA algorithms using the greedy action selection

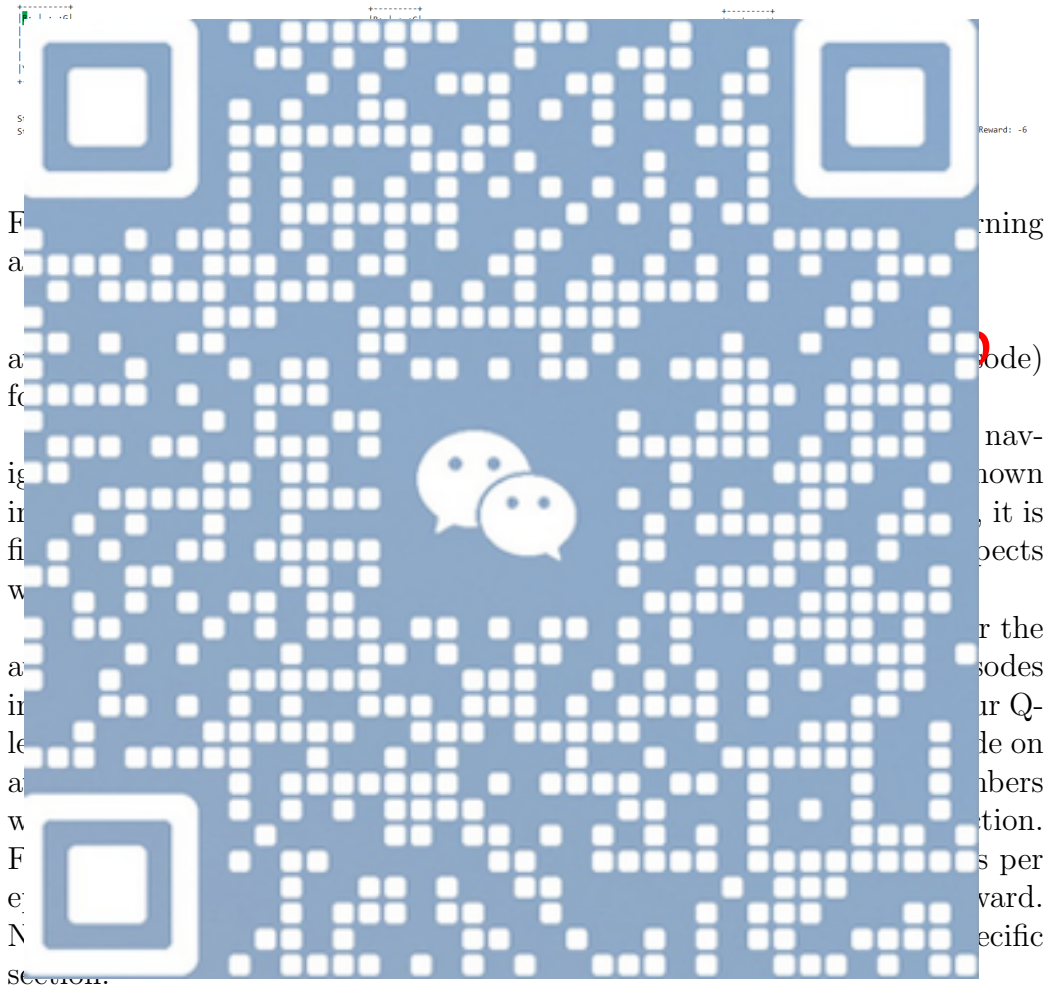
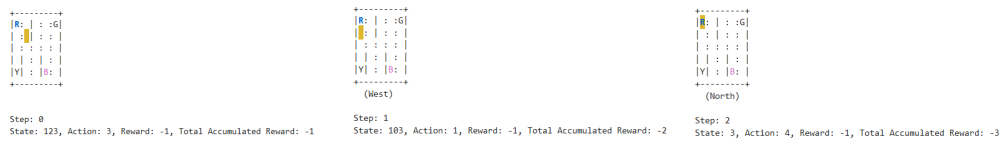


m
 n
 le
 v
 fi
 a
 v
 a

2

As part of the assignment evaluation, your code will be tested by tutors along with you in a discussion carried out in the tutorial session in week 10. The assignment has a total of 25 marks. The discussion is mandatory and, therefore, we will not mark any assignment not discussed with tutors.

Before your discussion session, you should prepare the necessary code for this purpose by loading your Q-table and the “*Taxi-v3*” environment. You should be able to calculate the average number of steps per episode and the



Finally, you will receive 1 mark for code readability for each task, and your tutor will also give you a maximum of 5 marks for each task depending on the level of code understanding as follows: **5. Outstanding, 4. Great, 3. Fair, 2. Low, 1. Deficient, 0. No answer.**