main

November 7, 2024

0.1 Libs

- numpy
- matplotlib
- gymnasium (openai gym)
- moviepy (for recording video of agent playing)
- tensorflow / tensorflow[and-cuda]

0.2 Documentations / References.

Documentation on gym / gymnasium can be found here: [https://gymnasium.farama.org/api/env/], it has good documentation on all the functions. - Do not use the one "showed" in exam paper: [https://www.gymlibrary.dev/index.html], since it has pretty bad documentation.

1 [TASK-1] Overview of the Environment.

1.1 Grid

The taxi can move on all nodes except X.

For each step it takes, it will be given -1 rewards, if successfully deliver passenger $+20$ reward points is granted. On failure on pickup and dropoff -10 reward points are granted.

$$\begin{bmatrix} \text{RED}_{11} & x_{12} & \text{X} & x_{14} & \text{GREEN}_{15} \\ x_{21} & x_{22} & \text{X} & x_{24} & x_{25} \\ x_{31} & x_{32} & x_{33} & x_{34} & x_{35} \\ x_{41} & \text{X} & x_{43} & x_{44} & x_{45} \\ \text{YELLOW}_{51} & \text{X} & x_5 3 & \text{BLUE}_{54} & x_{55} \end{bmatrix}$$

1.2 Actions

0: move south

1: move north

2: move east

3: move west

4: pickup passenger

5: drop off passenger

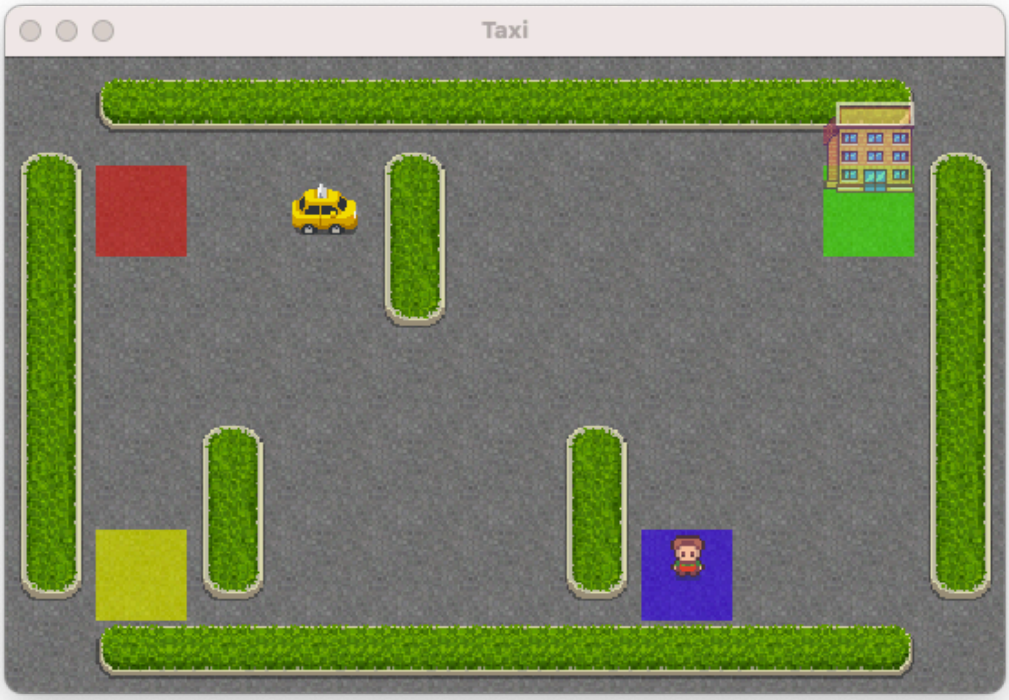
1.3 Observation space.

The taxi can dropoff passanger on (0)RED_{11}, (1)GREEN_{15}, (2)YELLOW_{51}, (3)BLUE_{54}. That means that the passanger can be on all these stats, and additionally (4)in taxi.

The grid is **5x5** resulting in **25** different locations for the taxi.

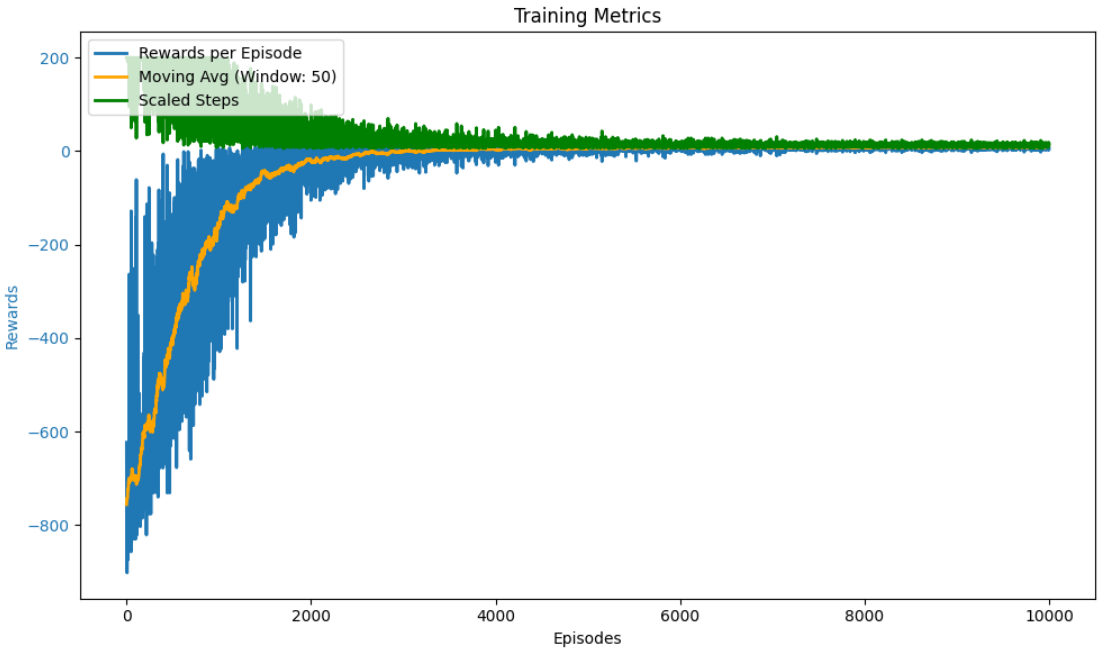
$4_{destinations} * 5_{passanger_locations} * 25_{nodes} = 500_{states}$

[10]:



2 Q-Learning Agent

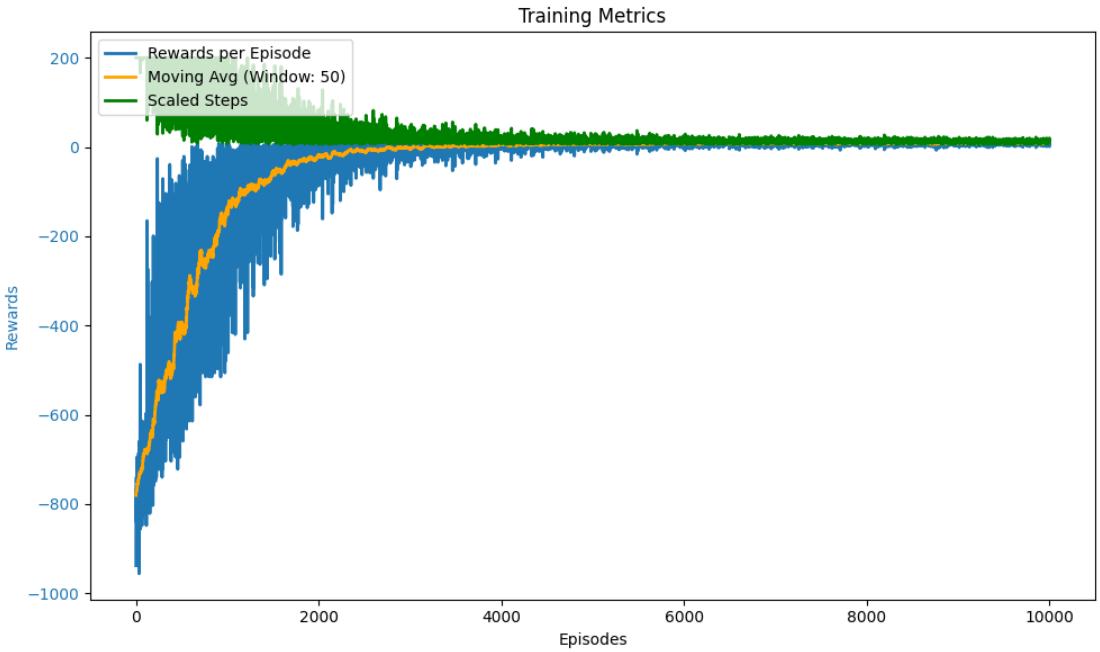
[3]:



[4]: <IPython.core.display.Video object>

3 Sarsa Agent

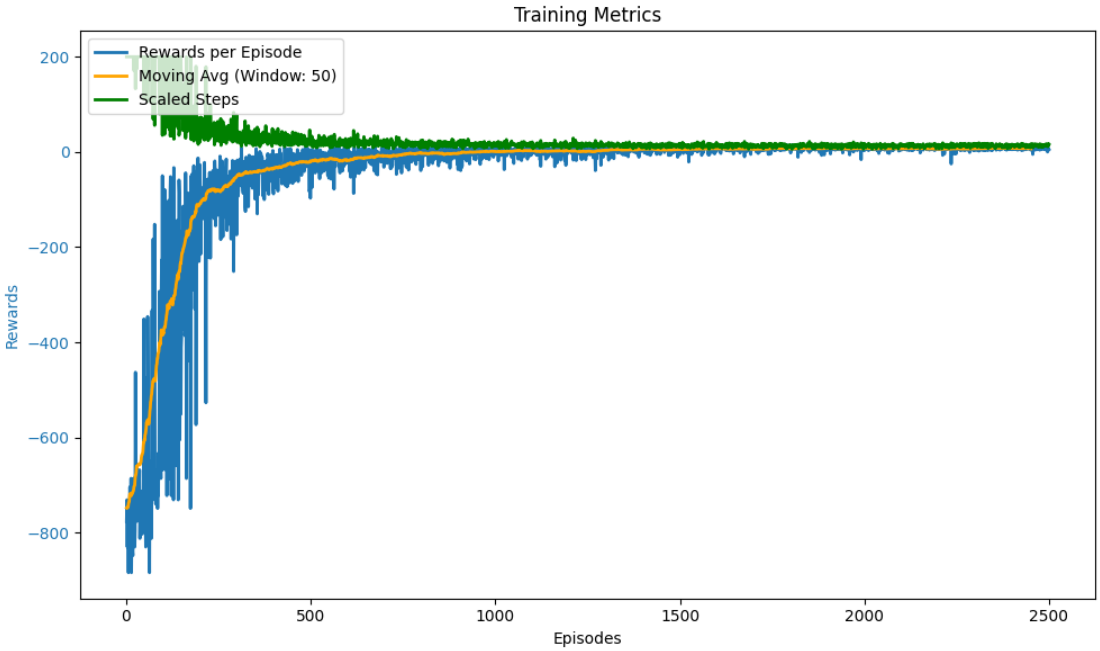
[5]:



[6]: <IPython.core.display.Video object>

4 Deep Q Agent

[7]:



[8]: <IPython.core.display.Video object>