

Reinforcement Learning - Project 4 Outline

Max Brehmer & Florian John

8 February 2024

1 Proposal

1.1 Topic

We are interested in working with a Grid-world style project. Depending on time pressure and how complicated the task gets, we are interested in making a multi-agent cooperation. This could for example be visualized as a taxi attempting to get to its destination as quick as possible, avoiding obstacles like red traffic lights or buildings. This could be expanded upon to include multiple vehicles attempting to get to different locations, perhaps all working for the same company, thus wanting to reduce the total time spent.

1.2 Model as MDP

To model this project proposal as a Markov Decision Process, we need to define a **environment**, an **agent**, all possible **states** and **actions** and a **reward process**.

1.2.1 Environment

We imagine using the **OpenAI Gymnasium API** and more specifically use the **Minigrid** environment to provide a world in which we can modify to meet our demands. To create something with multiple agents, we may use the **Pettingzoo** environment.

1.2.2 Agent

The agent would be one or more moving objects assigned a specific start tile (either at the start of each game or permanently) with the goal of reaching a target tile in as little time as possible.

1.2.3 States

The states would be easily described as the position in the grid that the agent(s) resides and whether the goal(s) is reached.

1.2.4 Actions

The possible actions are to move up/down/right/left if the state is not adjacent to an uncrossable tile, such as an edge, obstacle or perhaps another agent. If there is an uncrossable tile adjacent to the agent, the action space would be smaller.

1.2.5 Reward Process

The largest reward would be for completing the game. An adaptive reward could perhaps take place, in which the agent gets a smaller reward if the time or amount of crossed tiles is large. Smaller (mostly negative) rewards would take place for situations when say the agent wants to pass an uncrossable tile, a tile which has already been crossed is crossed again, or a punishment for moving to tiles far away from the end goal.