Q-learning path finder
Report

Michał Lubinski
184603

Allan Kirui
183036

Artificial Intelligence
Project

2022

1. *Analysis*

1.1 Introduction

The aim of our project was to implement a simulation of a path finding agent which is trained using a Q-learning Algorithm.

1.2 Q-learning

Q-learning also known as Reinforcement Learning is a model-free problem solving algorithm where decision making is procedural and the solution is a long-term goal. Therefore, it is not a greedy algorithm because it doesn't look for quick and immediate rewards. The q in Q-learning stands for *Quality*

The basic premise is of a growing car manufacturer which is building a factory where delivering of car parts to a specific part of the factory is done by a robot.

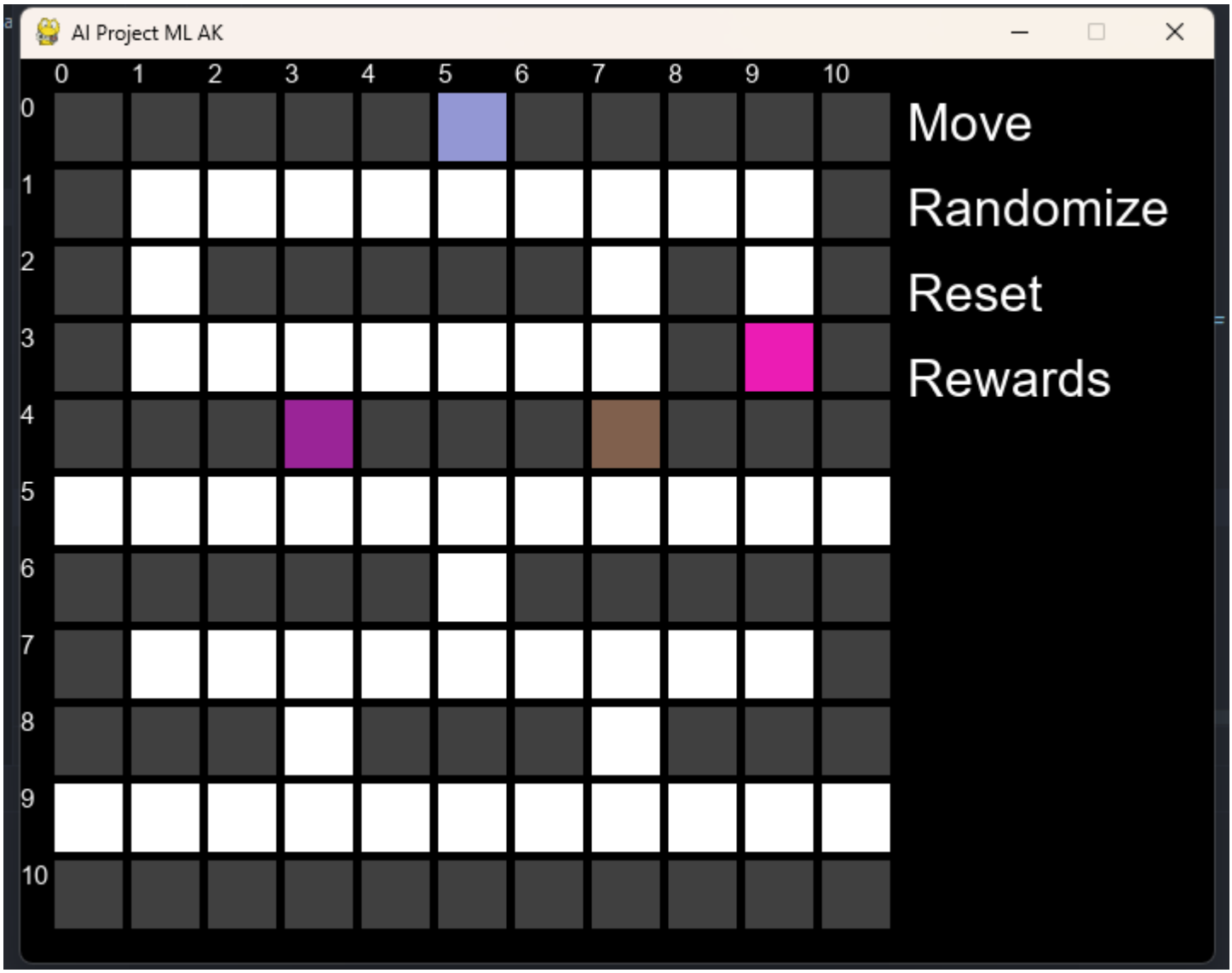

Figure 1.1: The board representing the factory floor

- The pink represents the current location of the agent
- The light purple represents the location of our goal
- The deep purple represents a Fence
- The brown represents Mud
- The Dark Grey represents a wall
- The white represents a path

Each of the fields of the factory floor have values which are used to train the robot on how to find the shortest path to the goal

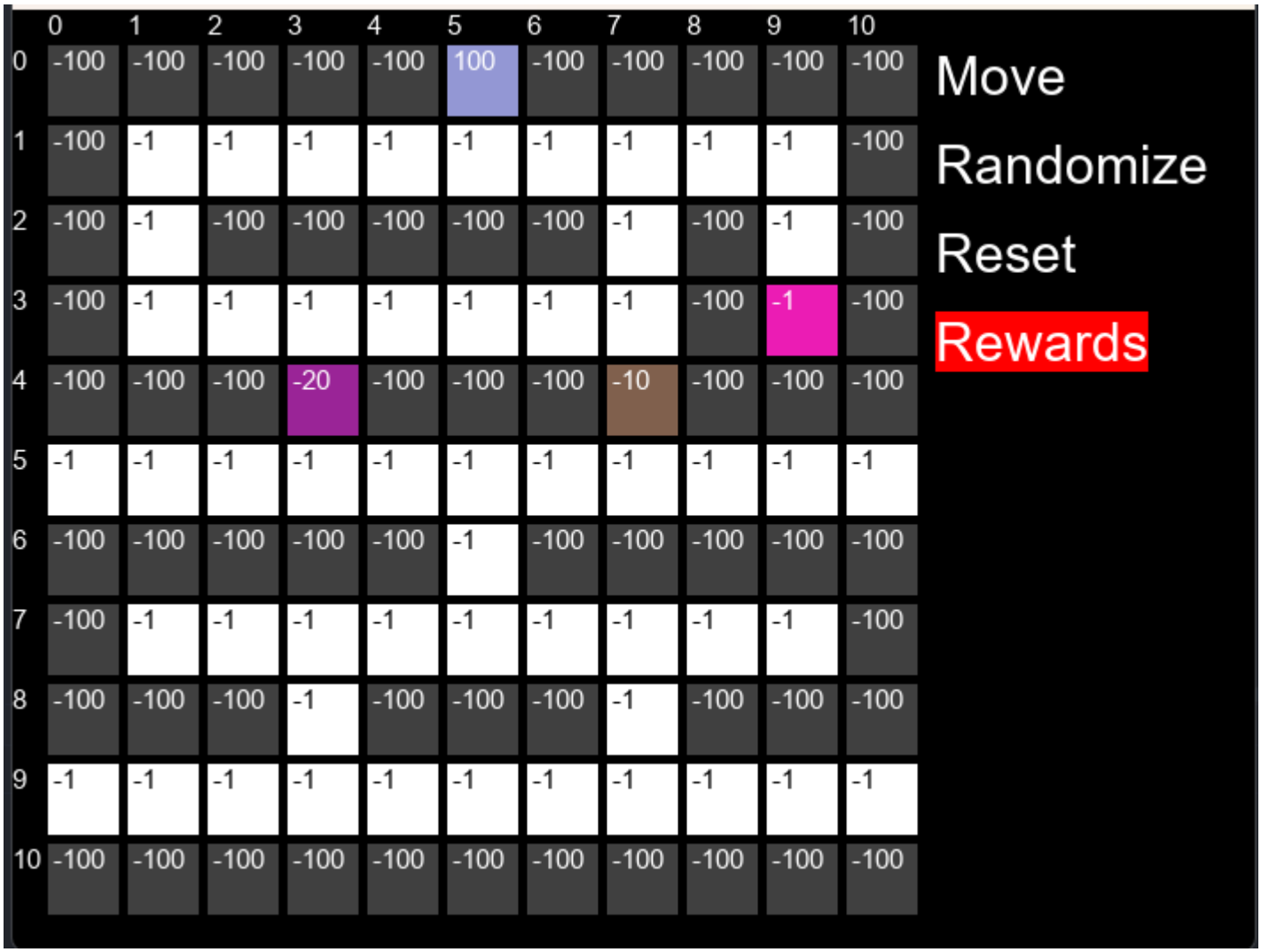

Figure 1.2: The board representing the rewards of each field on a factory floor

The robot will try to maximise its cumulative rewards. Therefore, it will try to avoid the walls and try to move on optimally the paths

1.3 Training the Robot

The training model first selects a random starting position for the robot. This is followed by choosing an action which is movement in either of 4 directions(up,down,left or right). It is selected using an epsilon greedy method which sometimes does not select the best scenario. The training model then moves to the next chosen field, and accumulates the reward of said field, which is then used in calculating the temporal difference The model then checks if we have arrived at our goal, if not it repeats this process from choosing an action.

1.4 Delivering parts

After training our robot, we can now use the knowledge to get the shortest path from a given starting location, as visualized below.

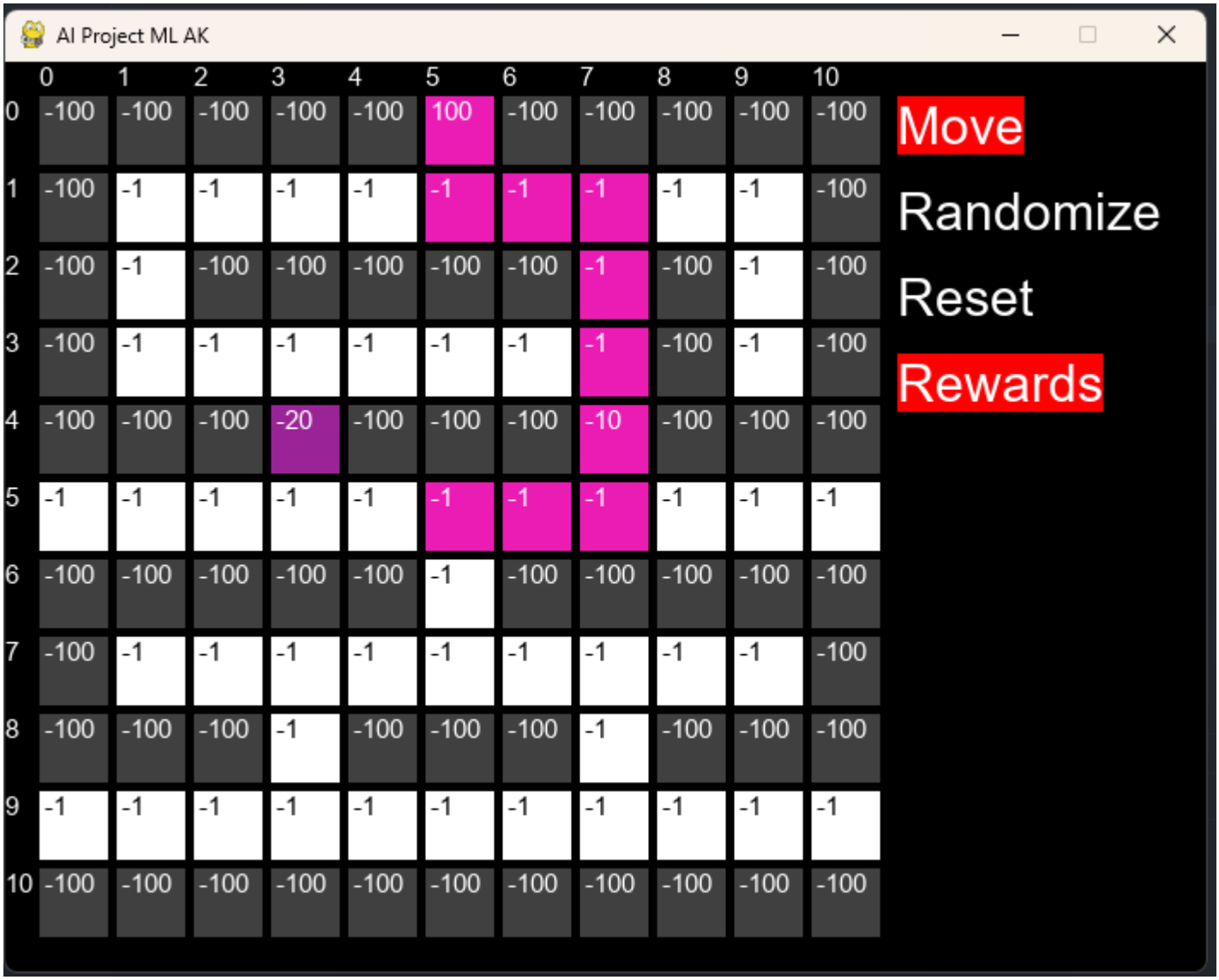

Figure 1.3: The board representing the delivery of a part on the factory floor by the robot