# Reinforcement Learning Application of Multi-Agent Path Planning

**Mohammad Hasan Olyaei Torqabeh[1], Amin Noori[2]**

[1] Sadjad University of Technology/Faculty of Electrical Engineering; mh.olyaei123@sadjad.ac.ir
[2] Sadjad University of Technology/Faculty of Electrical Engineering; amin.noori@sadjad.ac.ir

## Abstract

In this paper, we talk about the path planning problem of two robots. This robot routing has been trained through Reinforcement learning. In this paper, then, we are referring to reinforcement learning and the Q-learning method, and after defining the environment and problem constraints, defining the states vector and the actions vector, and defining the reward function, problem solved and simulation results are shown by using MATLAB software. The environment is a 7x8 cell that robots must learn to move towards the target point and pick up the target without colliding with obstacles, and then reach the goal point.

**Keywords:** Reinforcement learning, Multi agent, Path Planning, Q-learning

## Introduction

Reinforcement learning is one of the methods that used to solve many issues. One of the major applications of Reinforcement learning is the ability to control the robot with 7 degrees of freedom to learning flip pancakes in the air [1], learning Playing Atari games [2], learning control of an UAV and Designing a UAV autopilot [3] and etc.
In this paper discusses the navigation of two robots by reinforcement learning. The chosen policy is e-greedy.

## Reinforcement Learning

Reinforcement learning is part of the science of machine learning. The main body of reinforcement learning [4] is shown in Figure 1.
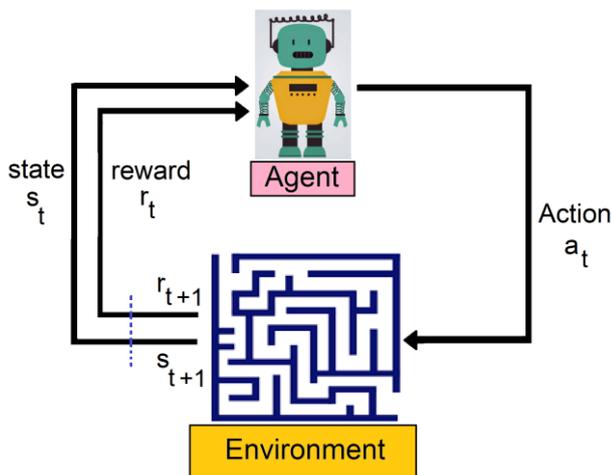
In Figure 1, the main body of the Reinforcement learning and how the agent communicates with the environment is shown. Each of the agents' state in step t, is $s_t$ and after take the $a_t$ action they are in $s_{t+1}$ and receive rewards according to the $s_{t+1}$ state. Again, this process is done in a new state. Reinforcement learning seeks to increase the total rewards received. The total received rewards [5] are calculated according to equation (1).

$$R = \sum_{k=0}^{\infty} \gamma^k r_{t+k+1} \qquad (1)$$

In equation (1), ɣ is the discount rate. The Reinforcement learning method used in this paper is Q-learning. Because learning has to be done for two robots, the reinforcement learning problem is presented in the form of Multi-Agent reinforcement learning. The agents in this problem are considered robots. Our goal is to let the robots learn how to reach and pick up the target and then reach to goal point. The agents must reach one of three goal points. The Environment for agents is unknown and agents must be able to identify the environment and make the goal of the problem.
The environment in this paper is a cell with dimensions of 7x8 squares, which includes a number of obstacles [6]. Agents must learn not to encounter with obstacles. The environment and location of the obstacles are shown in Figure 2. In Figure 2, white squares represent the locations where robots can move. Red squares represent obstacles and the green squares correspond to the goal points, and the yellow square represents the target's location.
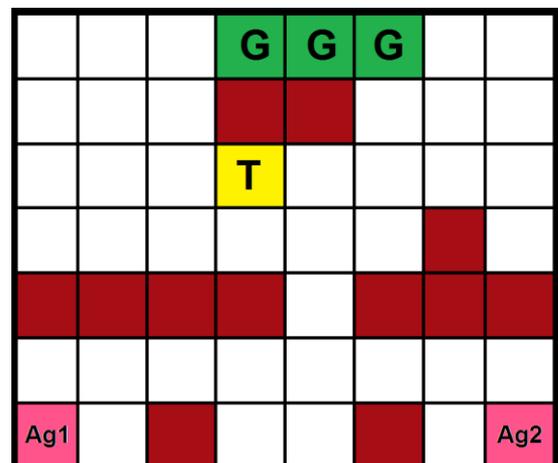


Figure 1. The Main Body of RL



Figure 2. The Environment