Bachelor thesis

An application of Reinforcement Learning method in simulations and control of Line Follower mobile robot

Adam Karwowski
Automatic Control and Robotics, Robotics
Academic year 2019/2020

Supervisor: Marcin Pękal, PhD Eng.

1. Introduction

Machine learning is a field of artificial intelligence. Its main assumption is that a computer has the ability to analyze provided data and — based on them — make appropriate conclusions, using the knowledge acquired during the learning process. The subject discussed in this diploma thesis refers to the reinforcement learning method, which is used here to control a Line Follower mobile robot that has to follow the path defined by the black line as quickly as possible. In the considered approach, the system (further called an agent)—in order to perform the established task—has to make specific decisions using the experience collected during the process of continuous exploration of the environment. Intervention of the user is only limited to define how the program perceives the external world (e.g. how it interprets measurements from the sensors), what actions the agent can perform (e.g. the robot can turn or go straight only) or assess which of them lead to the established goal (e.g. the robot should follow the black line) and which of them are bad for the mobile robot.

2. Reinforcement learning method

In every moment, an *agent* is in a particular *state* of the environment and makes an *action* in order to achieve established goal. As a result, the conditions of the surrondings change. After analyzing this new state, a reward is given to the agent. A positive value of the reward indicates that the taken action leads to the established goal while negative value informs the system that the choice is incorrect. To assess if the action made in the particular state is correct, the *Q-value function* $Q(s,a) : S \times A \rightarrow \mathbf{R}$ (where S is a set of all available states and A is a set of all defined actions) is introduced:

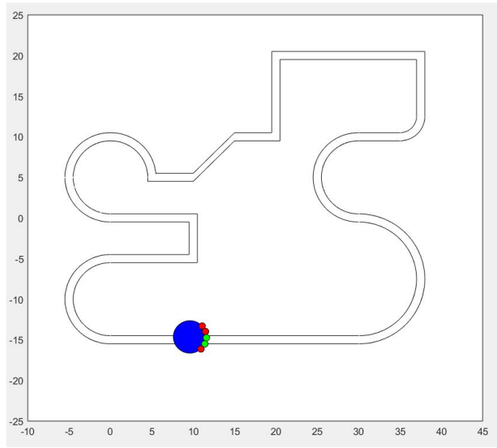
$$Q_{k+1}(s_k, a_k) = Q_k(s_k, a_k) + \alpha(r(s_k, a_k) + \gamma \max_{a \in A} Q_k(s_{k+1}, a) - Q_k(s_k, a_k)), \tag{1}$$

where k is the actual time step, s_k is the state and a_k is the action in the time step k, $r(s_k, a_k)$ is the reward received for making the action a_k in the s_k state, α is the learning rate and γ is the discount factor.

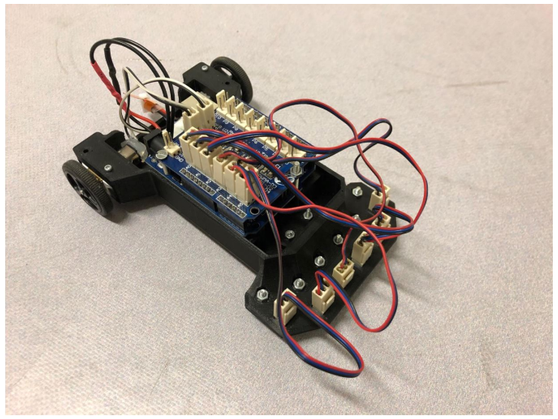
Robot, while moving, receives the reward for actions it makes. Based on this data, it updates the Q-value function, which leads to better knowladge of the environment. After an appropraite number of trials, the system learns how to properly perform the task.

Main goal of the thesis

The main goal of the thesis is to use the reinforcement learning method to control a Line Follower mobile robot and analyze which factors influence its process of learning and moving.



(a) Simulation of the Line Follower mobile robot made in MATLAB computing environment



(b) Line Follower robot used for implementation of the algorithm

3. An application of Reinforcement Learning method in control of Line Follower mobile robot

In this diploma thesis, the simulation of the Line Follower mobile robot is made in MATLAB computing environment (Figure 1a). After applying the reinforcement learning method, an analysis of factors affecting the process of learning the correct way of moving is performed. As a result, an optimal robot operation strategy is determined and implemented on the mobile robot.

Based on the results obtained in simulation, the algorithm is applied to the Line Follower mobile robot (Figure 1b). Observations made during the learning process of the agent meet the conclusions obtained from the simulations. The robot makes a few random actions in the first runs. Consequently, it gets the knowledge about the environment in which it is located. After the appropriate amount of time spent on the learning process, the robot learns how to properly perform the task and it is able to achieve the specified goal.

4. Conclusions

- The reinforcement learning method can be used to control a Line Follower mobile robot. However, this algorithm gives worse results than those obtained from the traditional way of controlling the robot.
- A lot of factors influence the learning process of the robot. The accuracy of the way of moving depends on how the user selects them in the algorithm.