ICAISC Zakopane

International Conference on Artificial Intelligence and Soft Computing

June 3-7, 2018

Motivated Reinforcement Learning Using Self-Developed Knowledge in Autonomous Cognitive Agent

Piotr Papież and Adrian Horzyk

AGH University of Science and Technology

AGH

Abstract

This project describes the development of a cognitive agent using motivated reinforcement learning.

The conducted research was based on the example of a virtual robot that placed in an **unknown** maze was learned to reach a given goal optimally. The robot should expand knowledge about the surroundings and learn how to move in it to achieve a given target. The built-in **motivation factors** allow it to focus initially on **gaining experience** and **forming knowledge** instead of trying to reach the goal.

Research and Project Novelty

Project assumptions:

- In the beginning, the robot has no knowledge about its surroundings.
- It must draw conclusions from taken actions.
- The robot should learn how to reach a predetermined location in the maze from any starting point in an optimal way.
- The robot interprets the room as a square grid.
- It can move in four directions: forward, backward, right, or left.
- The robot has sensors that allow him to detect a near obstacle in front of it.
- In the beginning, it only knows the coordinates of the target relative to its starting position.

In the beginning, the robot tries to perform random actions and **draw conclusions** from them as the living beings do. It is possible thanks to defined reward function. For example, the robot knows that falling into a wall is unfavorable. In this way, through the interaction with the surroundings, the robot learns what movements should be performed in specific situations. **Reinforcement learning was expanded by motivated learning** which allows building better knowledge about the surroundings.

Methods

Deep Q-learning:

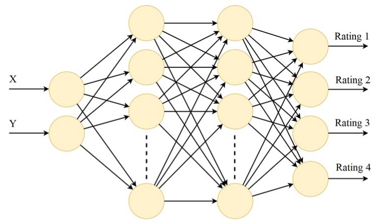
The reward-punishment system motivating the robot to achieve the target location was defined for this research as follows:

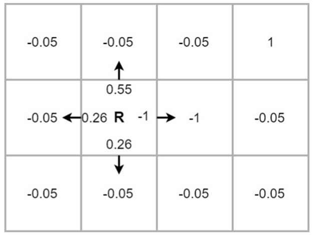
- 1 - for the reaching the goal,
- -1 - for the falling into an obstacle or an object,
- -0.05 - for the movement to the empty field.

In order to satisfy the need to reach the goal, the robot is guided by the Q function defined in the deep Q-learning algorithm.

$$Q(S_t, A_t) \leftarrow R_{t+1} + \gamma \max_a Q(S_{t+1}, a)$$

The role of the Q function is played by the feedforward neural network. At the input, it takes the current coordinates of the robot, and at the output gives four ratings assigned to specific directions of movement.





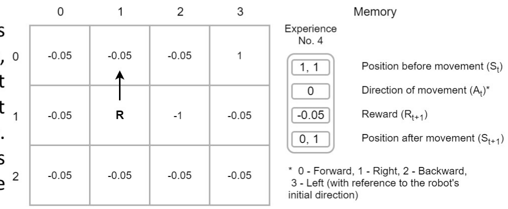
Motivation to explore:

Before the robot begins to correctly choose the subsequent directions of movements on the way to the target location, it is worth using this time for **effective learning**. Therefore, the robot was motivated to explore the surroundings in a situation where there is no sufficient knowledge about it. In the initial phase, this motivation exceeds the motivation to reach the goal. Two exploration methods were defined:

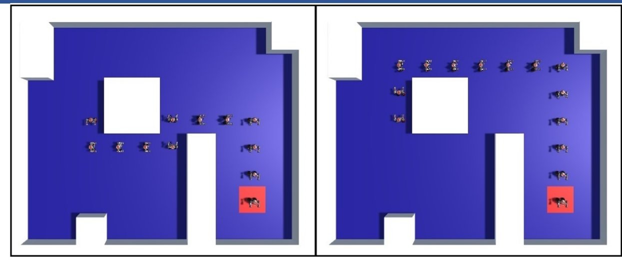
- ε-greedy policy,
- going to unvisited places.

Experience replay:

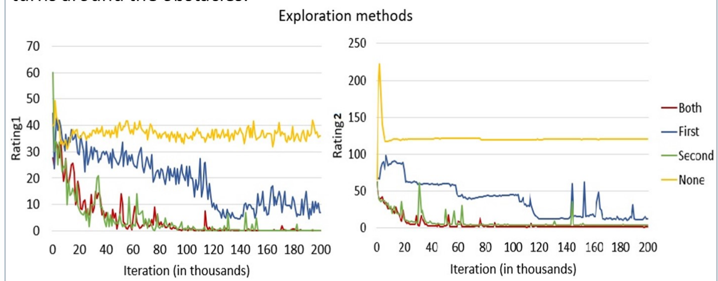
Memory, which stores the robot experiences, is a very important part of the robot cognition system. The robot uses the saved experience to relearn.



Results



The properly learned robot can choose **shorter route** shown on the left. The robot chooses it even though it is more difficult because it must make more turns around the obstacles.



The motivation for exploration is very important because it significantly **improves the learning process**. The learned robot can reach the target from **any place**.

Contributions and Conclusions

1. The developed software module allowed the robot to learn how to reach the designated location from any location in the room in an optimal way.
2. The robot had to collect all information about the surroundings. For this purpose, the robot was given cognitive qualities, and learning was based on the introduced motivated reinforcement learning.
3. By defining the motivation for exploration, the robot could abandon the initial phase of following the neural network and instead gathered experience. Thanks to this, the robot's learning was based on a large database of experiences collected during interactions with the surroundings. As shown in the studies, this had a key impact on the outcome of the project.

Contact Information

Piotr Papież
Email: ppapiez13@gmail

Adrian Horzyk
Email: horzyk@agh.edu.pl

References

P. Papież, A. Horzyk, Motivated Reinforcement Learning Using Self-Developed Knowledge in Autonomous Cognitive Agent, In: Proc. of 17-th Int. Conf. ICAISC 2018, Rutkowski, L., Scherer, R., Korytkowski, M., Pedrycz, W., Tadeusiewicz, R., Zurada, J.M. (Eds.), Artificial Intelligence and Soft Computing, Springer-Verlag, LNAI 10841, pp. 170-182, 2018.