Implementation of Reinforcement Learning Algorithms for Robotic Pick and Place with Non-Visual Sensing

Babar Imtiaz, Yuansong Qiao & Brian Lee
Software Research Institute (SRI)

Introduction

The term 'Industry 4.0' appeared for the first time in 2011 and saw the introduction of many cutting-edge technologies such as Cyber-Physical Systems (CPS), Internet of Things (IoT), and Digital Twins (DT). World leading scholar Warren G. Bennis was quoted in 2016 as follows

> The factory of the future will have only two employees, a man and a dog. The man will be there to feed the dog. The dog will be there to keep the man from touching the equipment.
>
> Warren Bennis

This quote clearly sums all the expectations and beliefs regarding the implementation of Industry 4.0. Materializing the concept of Industry 4.0 has created a number of global efforts such as Europe's Industry 4.0 [1], America's Advanced Manufacturing [2], China's Made in China 2025 [3], Japan's Super smart society [4], etc. Smart factories realize manufacturing processes with the aid of artificial intelligence (AI), the latest novel sensors, and use of robotics.

When you link the terms 'AI' with 'robotics' in industrial context, what comes to mind is a well-known case of robotic manipulation known as "Pick and Place". Improving robotic manipulation "Pick and Place" with the help of advanced learning techniques has been a focus of the research community for some time now.

Research Focus

The main research focus of this study is making an agent learn robotic manipulation of Pick and Place using reinforcement learning in absence of vision sensors . An industrial setup can lack vision system due to multiple reasons such as :

- High cost
- Less space availability
- High vibrations
- Large amount of dust
- Wash-up from water jets

Reinforcement learning (RL) addresses this task by performing sequential decision-making through a policy learned during the process of maximizing an expected reward.

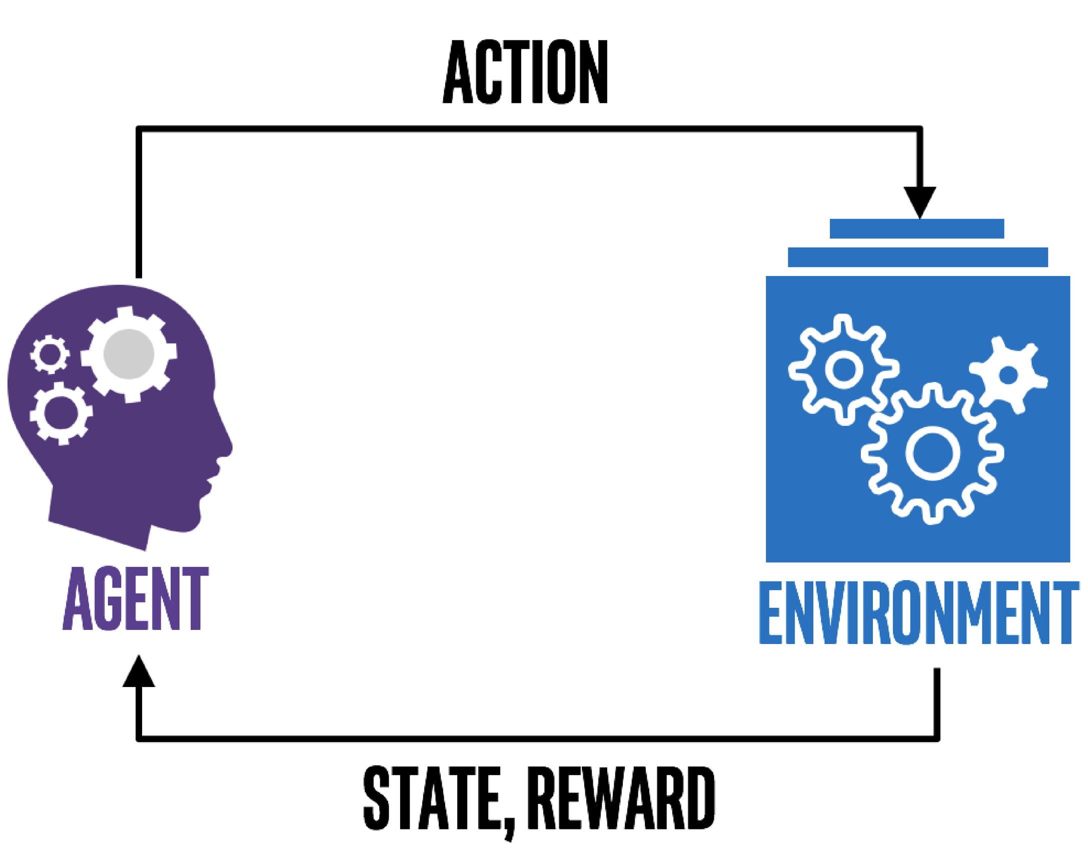

Figure 1

In this study we deploy model-free off-policy TD algorithm : Q-learning and on-policy TD algorithm SARSA. A comparative analysis was conducted in order to find the most suitable learning approach.

Methodology

The approaches presented in this study address the problem of pick and place in a smart production line, where a number of variable-shaped objects are moving on a conveyor belt at different positions and orientation, and where the belt may assume different speeds as shown in Figure 2. The conveyor belt is equipped with ray-type infrared proximity sensors which detect the object and signal the robotic arm to operate.

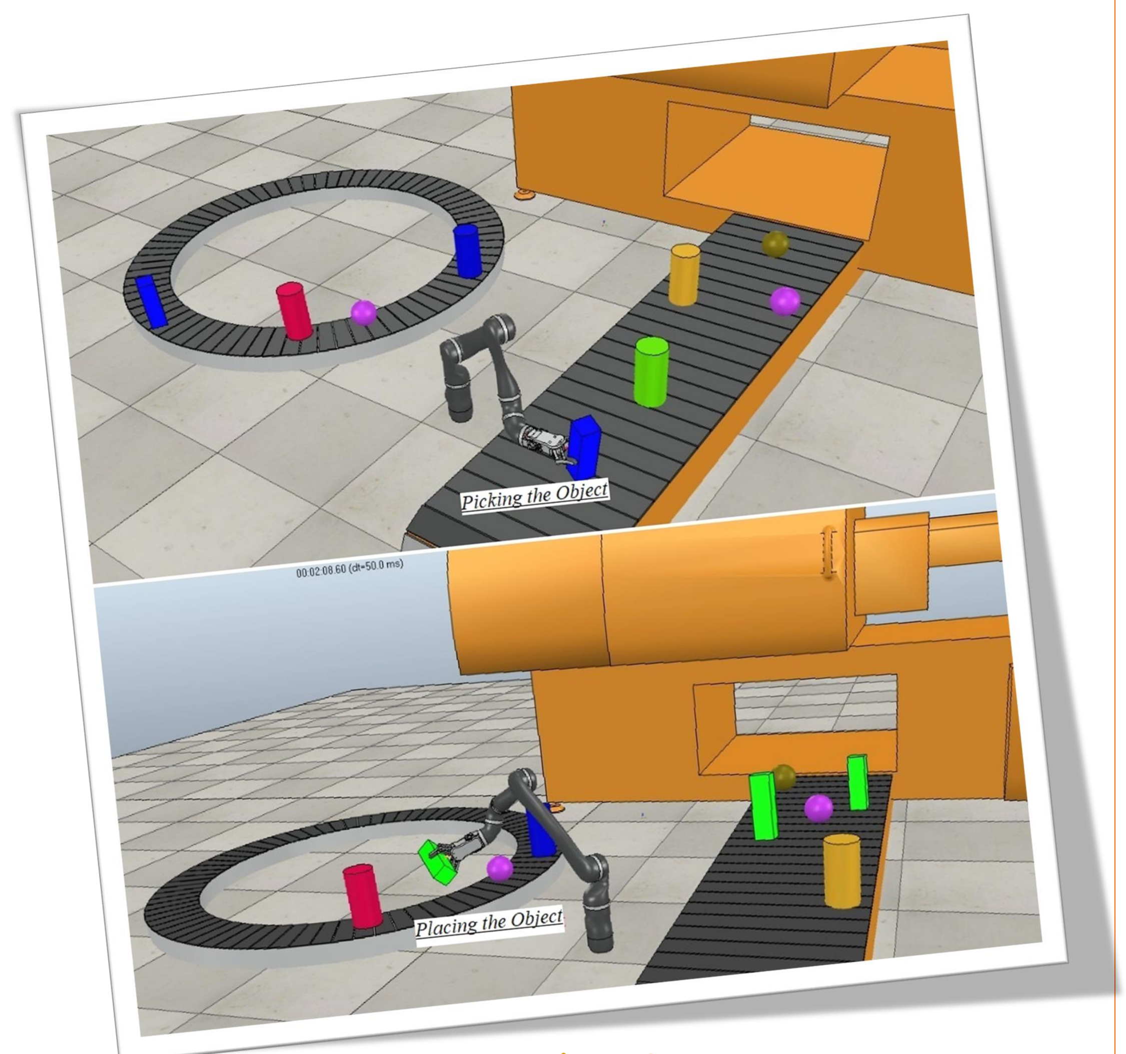

Figure 2

The approach is divided into four different phases, first being 'Initial Phase', second being 'Pre-Pick Phase', third being 'Pick Phase' & last one being 'Place Phase'. The cycle of these four phases can be easily perceived through Figure 3.

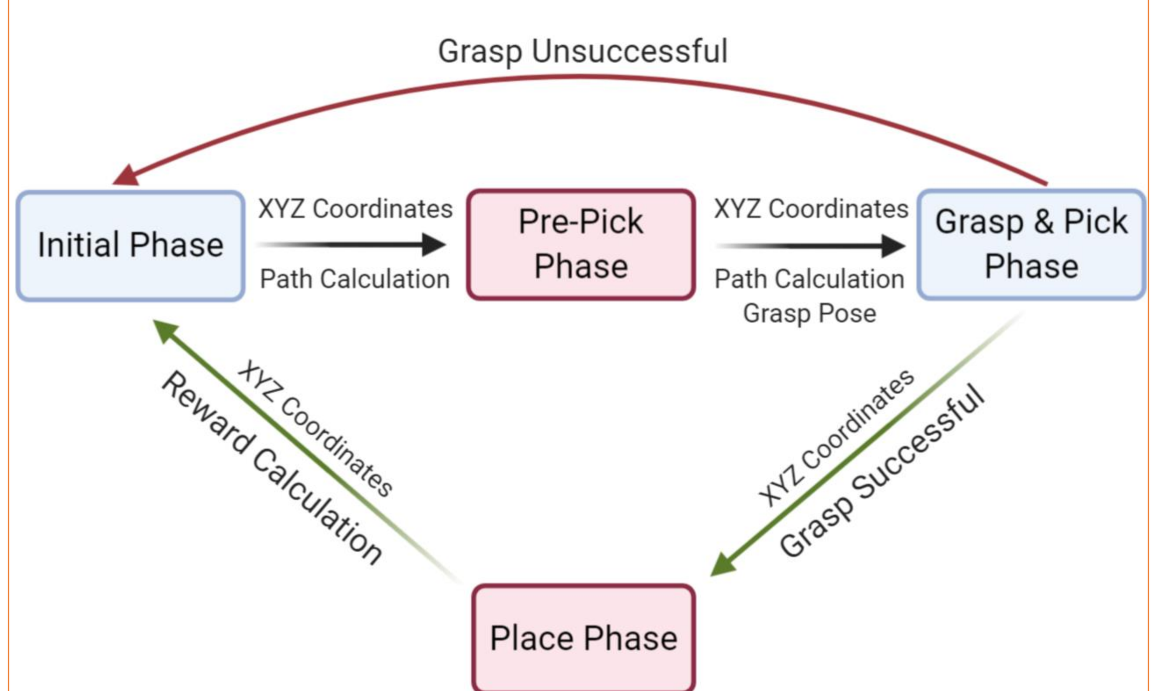

Figure 3

Elements that are being learned in our this non-visual approach are:

- Various XYZ Coordinates for Pick and Place
- Different positions and orientations of Objects on the Belt
- Multiple Speed options of the Belt
- Different Shapes of the Objects

The Virtual Robot Experimentation Platform (V-REP) is a 3D robotic simulator with an integrated development and coding support [5]. Open Motion Planning Library (OMPL) [6] proved to be the best approach as it provided us with a high degree of customization.

Results

The Q-learning and SARSA RL agents described in the previous section were extensively trained and evaluated in our experimentation phase as shown in the Figure 2. Performance comparison of Q-Learning agent and SARSA agent is shown in Table 1 and Figure 4 accordingly.

Table 1

Test Case	Q-Learning Success rate	SARSA Success rate
Test Case 1	93%	82%
Test Case 2	95%	81%
Test Case 3	99%	80%
Test Case 4	83%	77%
Test Case 5	97%	81%

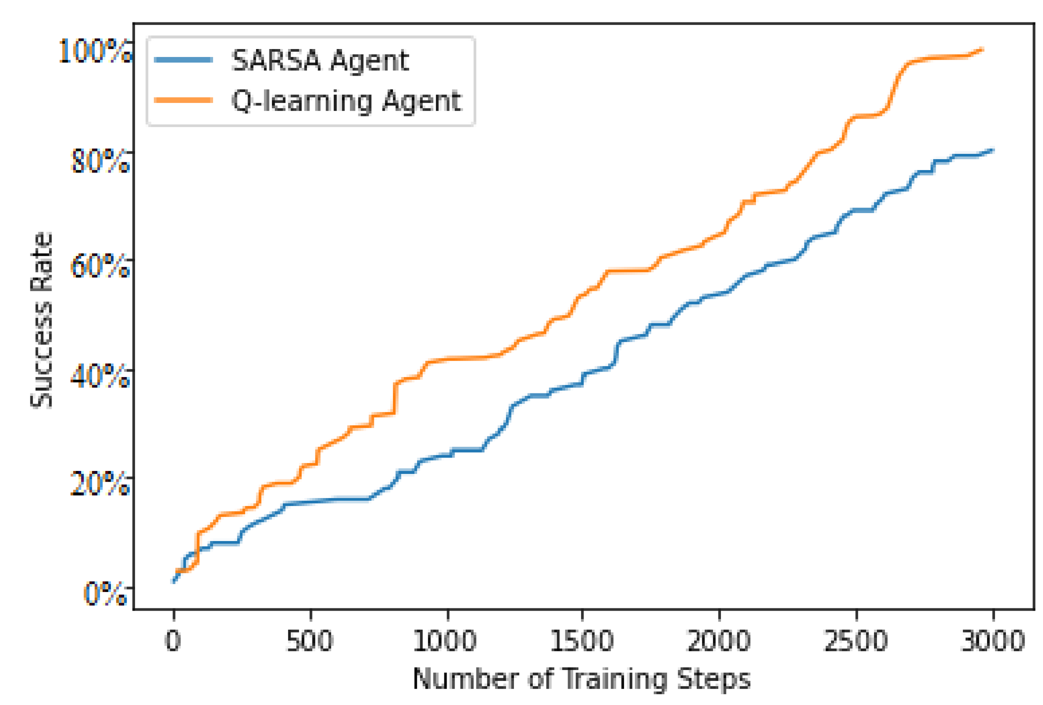

Figure 4

Conclusion & Future Work

In this paper, we have presented an approach to address the problem of industrial robotic pick and place in a non-visual environment. We formulated the problem as an MDP for which Reinforcement Learning provides a very extensive framework for dealing with such tasks. We deployed both model-free off-policy temporal difference RL algorithm (Q-Learning) and on-policy temporal difference RL algorithm (SARSA). We trained and tested Q-learning and SARSA agents on different shaped objects at different position alignments moving at different speeds.

So, for future work, we plan to address these limitations by using a camera, and deploying deep reinforcement learning to handle the state and action complexity. In order to increase the efficiency, a hybrid approach combining on-policy and off-policy algorithms such as backward Q-learning is also being worked on. We also plan to explore the possibility of deploying a multi-query planner for motion planning instead of a single-query planner in order to have multiple options computed on the run, hence increasing the chances of higher efficiency.

Acknowledgement

This publication has emanated from research conducted with the financial support of Science Foundation Ireland (SFI) under Grant Number SFI 16/RC/3918, co-funded by the European Regional Development Fund.

References

[1] B. Melzer, "Reference architectural model industrie 4.0 (RAMI 4.0)," p. 15.
[2] "Pcast-advanced-manufacturing-june2011.pdf." .
[3] C. C. Kuo, J. Z. Shyu, and K. Ding, "Industrial revitalization via industry 4.0–A comparative policy analysis among China, Germany and the USA," Glob. Transit., vol. 1, pp. 3–14, 2019.
[4] Y. Harayama, "Society 5.0: Aiming for a new human-centered society," p. 6.
[5] M. Freese, S. Singh, F. Ozaki, and N. Matsuhira, "Virtual robot experimentation platform V-REP: A versatile 3D robot simulator," in Simulation, Modeling, and Programming for Autonomous Robots, vol. 6472, N. Ando, S. Balakirsky, T. Hemker, M. Reggiani, and O. von Stryk, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 51–62.
[6] I. A. Sucan, M. Moll, and L. E. Kavraki, "The open motion planning library," IEEE Robot. Autom. Mag., vol. 19, no. 4, pp. 72–82, Dec. 2012.
[49] Y. H. Wang, T. H. S. Li, and C. J. Lin, "Backward Q-learning: The combination of Sarsa algorithm and Q-learning," Eng. Appl. Artif. Intell., vol. 26, no. 9, pp. 2184–2193, 2013.
[50] X. Li, Z. Lv, L. Wu, Y. Zhao, and X. Xu, "Hybrid online and offline reinforcement learning for Tibetan Jiu chess," Complexity, vol. 2020, p. 4708075, May 2020.

Confirm Smart Manufacturing

ORCID

Fondúireacht Eolaíochta Éireann
Dá bhfuil romhainn
sfi
Science Foundation Ireland
For what's next