Cooperative Industrial Multi Robot System using Multi Agent Reinforcement Learning

Xi Lan, Dr.Brian Lee, Dr.Yuansong Qiao

Software Research Institute (SRI), Athlone Institute of Technology (AIT)

Introduction

This poster describes the research definition process for this early stage PhD.
Our goal is to investigate the feasibility of using multi-agent deep learning (MADRL) as an approach for solving cooperative multi robot system (MRS) challenges in smart manufacturing. More specifically we pose the following research questions:

- What are the types of MRS for which an MADRL approach would be suitable?
- What are the RL challenges for the specific MRS application chosen for the PhD research?
- What MADRL methods are most suitable to apply to the chosen MRS application?
- To what extent, if any, can the results of the research be generalized to the use of MADRL for MRS in smart manufacturing?

Problem

In recent years heterogeneous robot systems based on Multi Agent System (MAS) are being used to collaboratively solve problems in a variety of domains including smart manufacturing. Industrial robots and conveyors have been widely applied in manufacturing to create product lines. Compared with robot systems with a single robot, a system with multiple robots may be preferable because it may improve productivity [1].

This study focuses on the use of MADRL to optimize multi-robot pick-and-place system for industrial operation. The multi-robots have the ability to cooperate and collaborate with each other for pick-and-place task in the shared environment (Figure 1). Each robot picks up parts moving through its workspace and places them in the box and has to decide which to pick and which to let pass.

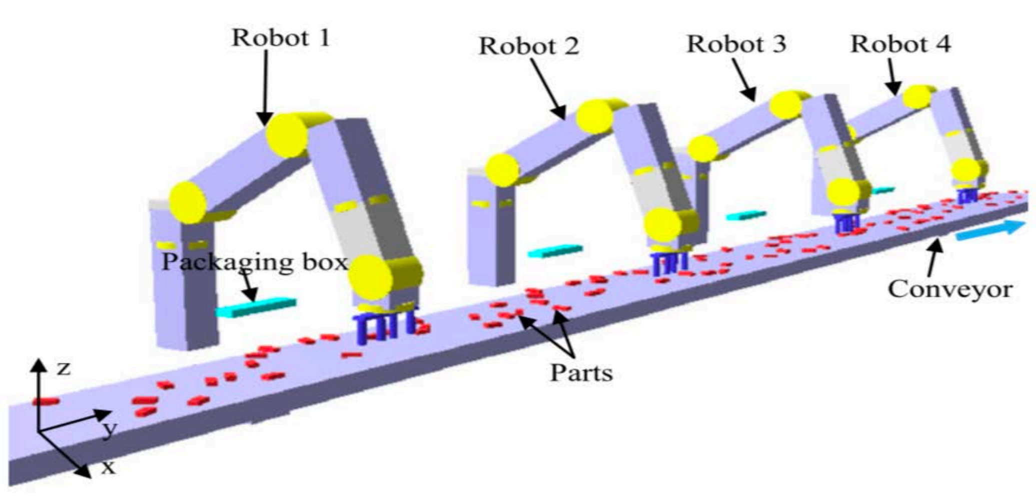

Figure 1: Multi-robot system for pick-and-place tasks.[1]

Methodology

Reinforcement learning (RL) has long been used in robotics for the design of sophisticated and hard to engineer behaviours [2]. **In order to learn within a reasonable time suitable approximations of state, policy and value function need to be introduced to reduce the" curse of dimensionality" MAS using RL (MARL),** have been researched for many years. One of the most significant advancements in RL in recent years has been the use of "Deep RL" to leverage deep learning as a promising approach to solve real world problems including in manufacturing[3]. Deep learning helps to reduce the curse of dimensionality for large state-space RL problems and to reduce the need for manually designing features to represent state information.

DRL algorithms have been designed on both value-based and policy gradient approaches. Some DRL methods that have been used to help with RL challenges include Deep Recurrent Q-Networks (DQRN) to improve partial Observability or Deep Deterministic Policy Gradient (DDPG) based **Asynchronous Advantage Actor Critic (A3C)** algorithm [4] for RL applications involving physical control where the action space is high dimensional and continuous – classic robotic challenges identified by [2] above. A3C trains an agent by asynchronously executing multiple agents in parallel, on multiple instances of the environment.

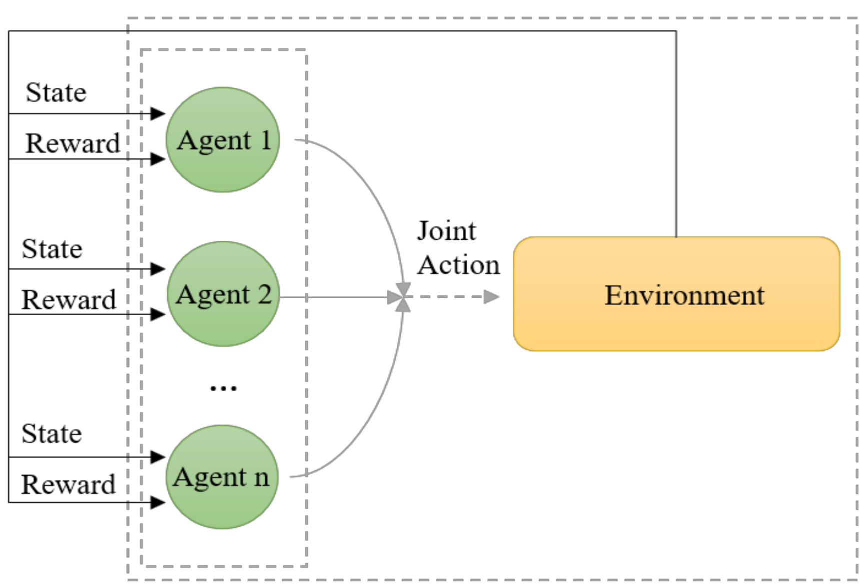

Figure 2 : Multi Agent RL

This extends the many benefits of A3C for the MADRL based robotic MAS teams and is therefore the chosen approach for our research, which we elaborate in the next section.

Approach

To carry out our research we intend to explore the A3C multi agent architecture discussed in [5] – see Figure 4 from [5]- which was designed for a robotic agent block assembly problem. This is a shared learning framework, with 6 agents contained in a meta-agent interacting with a shared environment.

The worker agents in our version of the architecture will represent the pick and place robots and the composite action space will be product of the individual robotic pick and place operations. The composite state space will represent the conveyor belt and each agent space will represent the workspace of each robot.The global Actor-Critic network is stored in the meta-agent. Each internal agent (worker) in Figure 1. draws its current state from the shared environment and selects an action to be performed next. Individual agents' actions are performed synchronously, in any order, and yield individual rewards. Each worker regularly pushes gradients to the global network, after which it pulls the most recent weights from the updated network.

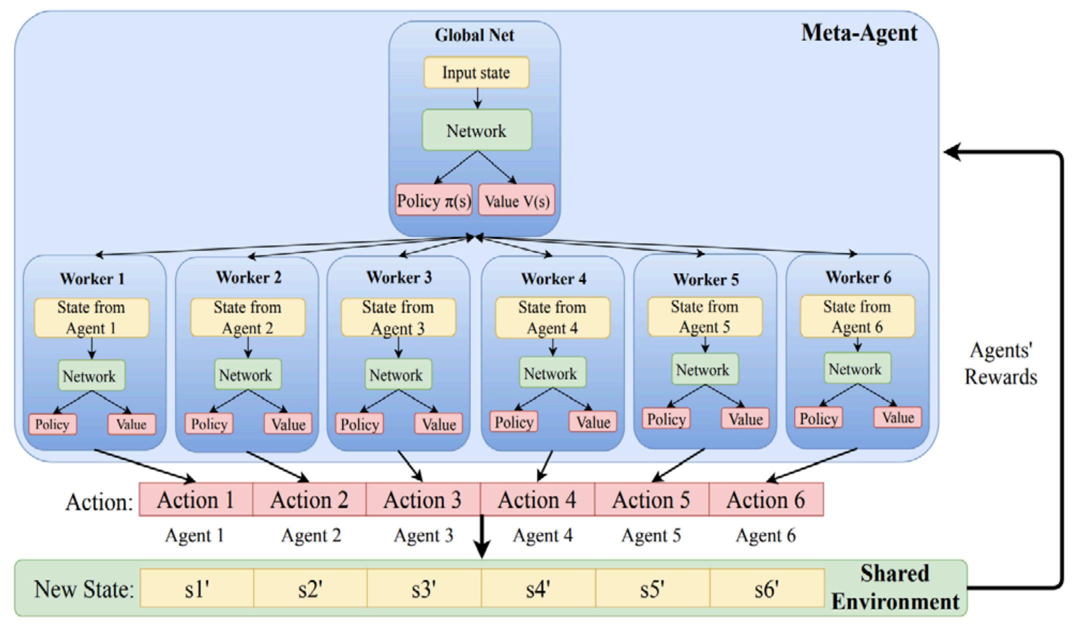

Figure 4: Sharing learning framework Adopted from [5]

we extend the single-agent asynchronous advantage actor-critic (A3C) algorithm to let multiple agents (multi robots) learn a homogeneous, collaborative policy in a shared environment.

We will explore the parameter sharing architecture for actor-critic cf. Figure 5. The framework shares the weights of a certain number of layers between the actor and critic networks of two or more agents. This figure depicts two shared layers in both the actor and critic networks. For our case this depicts two robots sharing parameters for two layers. By sharing parameters from selected layers, the robots can strike a balance between having unique policies and sharing parameters. This will help solve the challenge of how to dispatch parts to the appropriate robot so that each robot picks up as many parts moving through its workspace as possible.

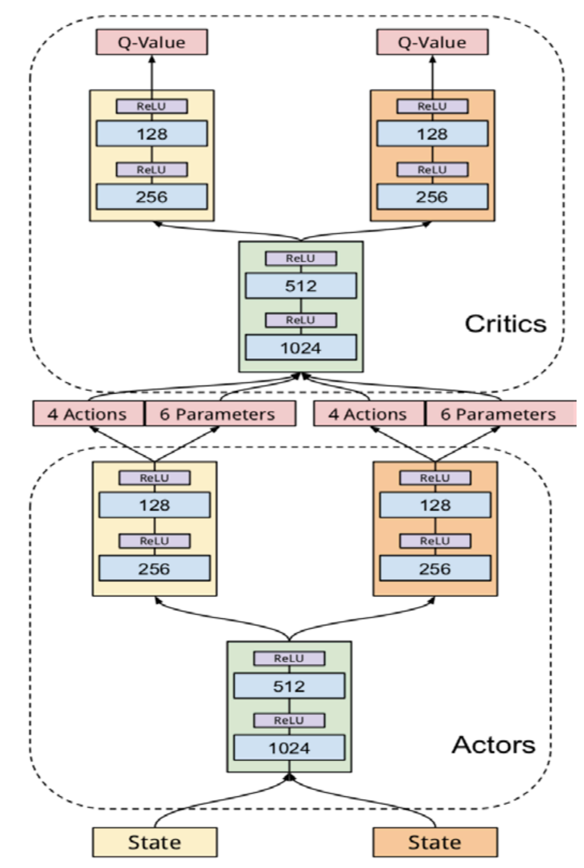

Figure 5: AC Parameter Sharing Architecture

Conclusion

This poster describes our proposed research to investigate multi robot pick and place task. Deep reinforcement learning has shown recent success on many fronts and a natural next step is to test multiagent scenarios. In this study, we seek to solve the problem of multi robot's coordination and cooperation using MADRL. We base our approach on the use of the A3C algorithm which has proven to be very versatile to perform well on a variety of continuous motor control problems.

REFERENCES

[1] Y. Huang et al. / Robotics and Autonomous Systems 64 (2015) 70–83
[2] Kober, J., Bagnell, J. A., & Peters, J. (2013). Reinforcement learning in robotics: A survey. The International Journal of Robotics Research, 32(11), 1238–1274.
[3] Nguyen, T. T., Nguyen, N. D., & Nahavandi, S. (2019). Deep Reinforcement Learning for Multiagent Systems: A Review of Challenges, Solutions, and Applications.
[4] V. Mnih, A. Mirza,A.Graves,T,Harley,T.P. Lillicrap,D,Silver,K,Kavukcuoglu(2016) , Asynchronous Methods for Deep Reinforcement Learning.
[5] Sartoretti, G., Wu, Y., Paivine, W., Kumar, T. K. S., Koenig, S., & Choset, H. (2019). Distributed Reinforcement Learning for Multi-robot Decentralized Collective Construction. Distributed Autonomous Robotic Systems.

THE TIMES THE SUNDAY TIMES GOOD UNIVERSITY GUIDE 2020 INSTITUTE OF TECHNOLOGY OF THE YEAR

AIT Research