Yang You
Cranfield University/INSA-Rouen
+33782658879
y.you@cranfield.ac.uk

To whom it may concern

Thank you very much for your reading, if I may, I would like to introduce my research experience about dual robotic arms' manipulation which uses the deep reinforcement learning in laboratory of Complex System and Intelligence Science, Institute of Automation, Chinese Academy of Sciences.

The robot grasping/manipulation task is one of the most difficult problem when designing an intelligent robot in the world now, the current machine learning methods are not powerful enough to solve this problem perfectly, especially when it comes to two robot arms with collaboration. Our team's goal is realizing the End-to-End Vision-Based Robotic Manipulation with collaboration. The structures of robot control process and the state of the art "End to End robot control with DRL" are as follows:

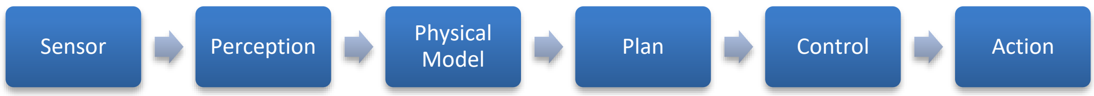


Figure 1 - Robot Control Process

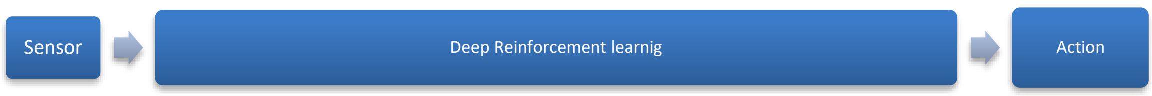


Figure 2 - End to End Robotic Manipulation which uses DRL

My main works can be divided into two parts, the first part is using the exist robot model (Fetch Robot) to test the deep reinforcement learning methods and training dual models learn to collaborate with another. After a success in the first part, the second part is remodeling a robot which is total different one compared with Fetch, and then transferring the deep reinforcement learning methods we have used to the new model.

Part 1:

1.1 One robot arm manipulation:

We used the physical engine Mujoco as our simulation environment to train the agent, and the first task is make sure that the two arms can reach to the object, for letting the robot learn this process, we used DDPG and HER to train one arm first and the other arm remained steady, then I combined the two arms as one agent and modified the inputs and outputs' dimensions to train them together, after few attempts of modifications in parameters in algorithm, we have successfully trained the robot realizing the first task.

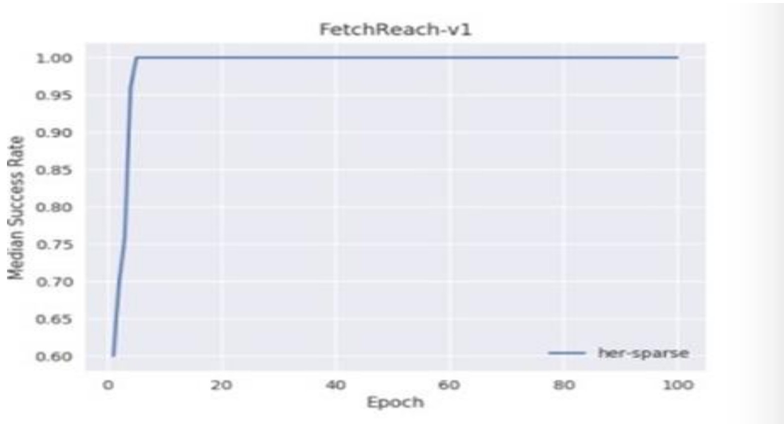


Figure 3 - One Fetch Robot Reach

1.2 Two robot arm manipulation:

The second task is to pick and place the object with two robot arms with collaboration. This part was the most difficult one, because the object was subject to the gravity and if two arms didn't pick them in the same time and in the perfect positions, this object would fall, besides the dimensions of action increased dramatically when adding the two robot hands to pick and place the object, those problems made the algorithm much more difficult to converge. But we had successfully made a breakthrough in this part, we learned from the AC cross inputting from MADDPG and combined it with HER in our algorithm. The brief structure of our algorithms is like this:

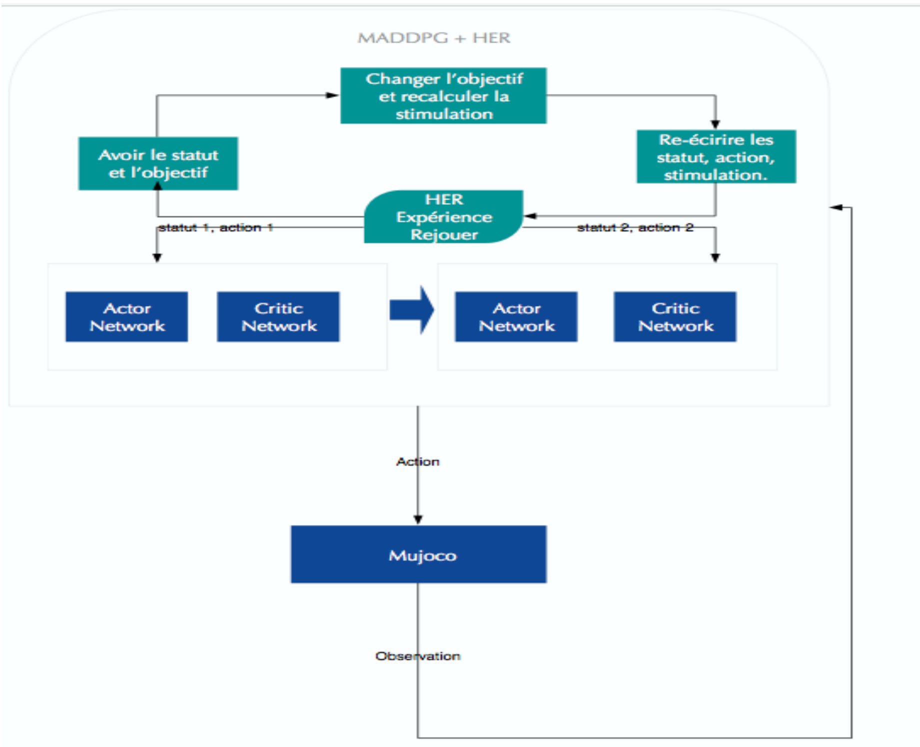


Figure 4 - MADDPG +HER Algorithm's Structure

With few changes in the reward function, we had a very promising result.

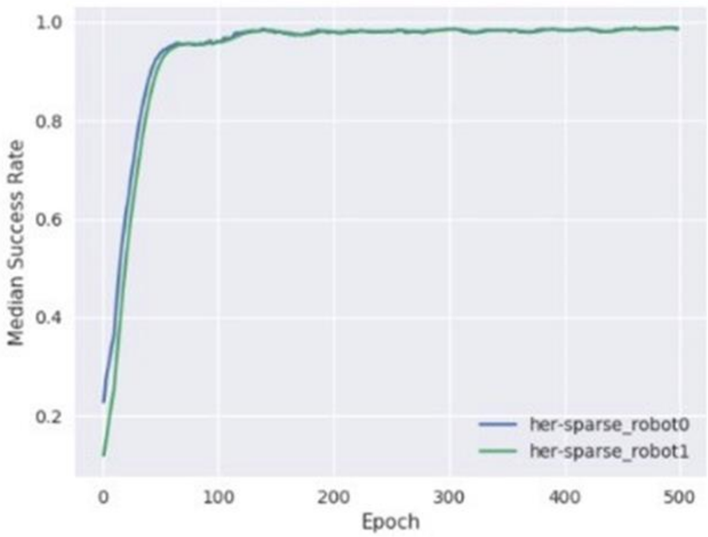


Figure 5 - Two Fetch Robots - Pick and Place

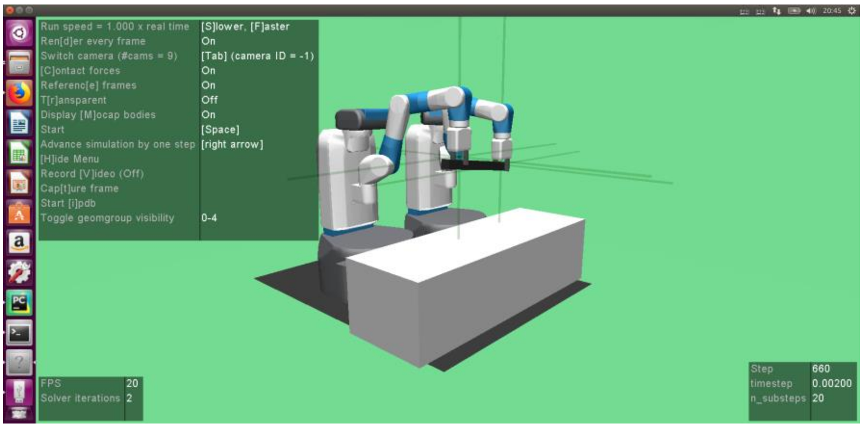


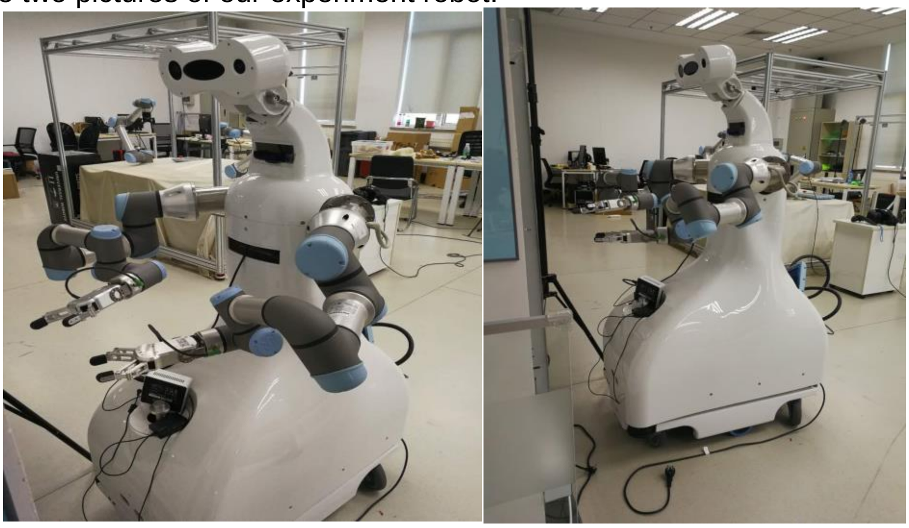
Figure 6 – Simulation Visualization

The next task is changing the Fetch robot model to our experimental robot model, and it comes part 2.

Part 2:

2.1 Robot Modelling:

There are two pictures of our experiment robot.



In this part, we need to modelling the whole structures and mechanisms of our robot. A brief process of the modeling is presented as follows:

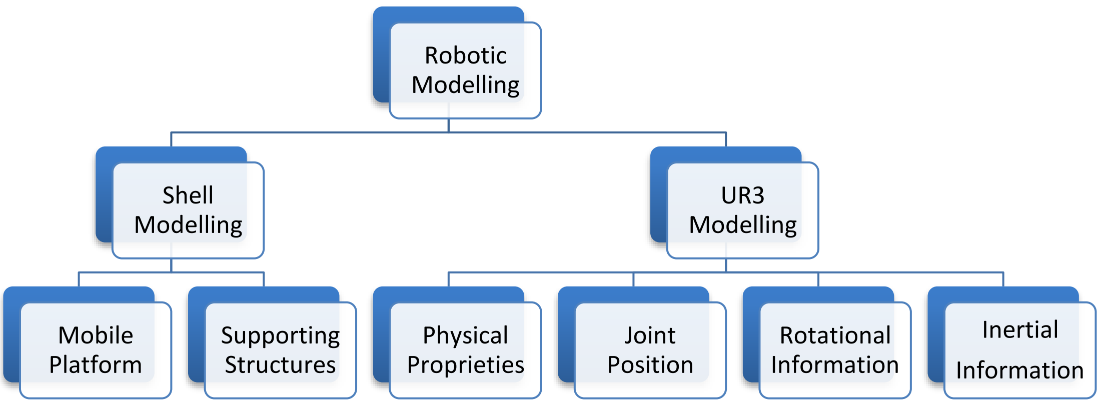
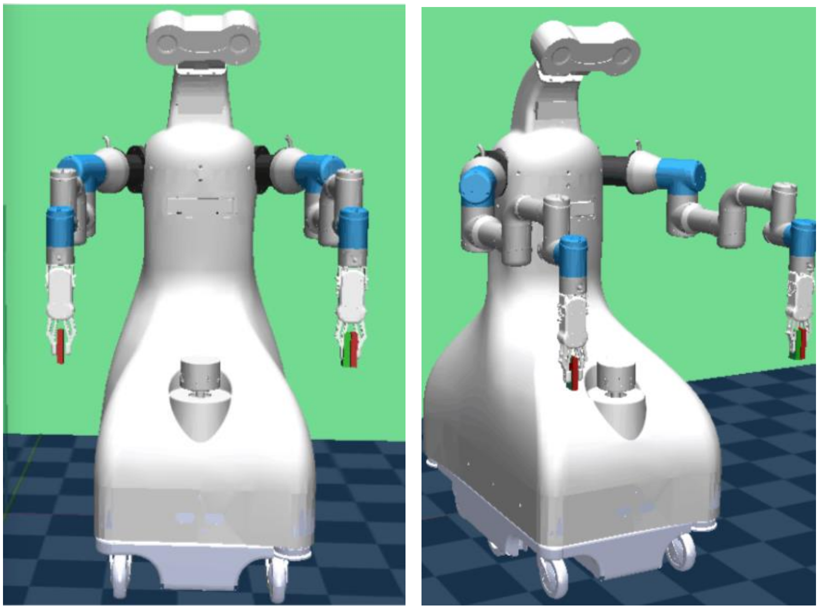


Figure 7 - Modelling Structure

The final model of our experiment robot is presented as follows:



The model files are in type ".STL" and we need to define the physical proprieties such as the mass, inertial and positions in order that the physical engine can simulate all the movements. Thus, in the virtual environment (Mujoco), we redefined the .XML files which are used to describe the proprieties and relations for each part, an example of the definition in .XML file is like this:

```
<body pos="1 0 0">
  <inertial pos="0 0 0" mass="1000" diaginertia="166.667 166.667 166.667"/>
  <geom type="box" pos="0 0 0" size="0.5 0.5 0.5"/>
</body>
```

Figure 8 - Example of properties' definition

2.2 Implement DRL to the new model:

The final task is like the part 1, we need to train our new model which is also consist of two agents (the left and right arms) to learn collaborate to pick and place objects to a random target position. Despite we have succeed training two Fetch robot model collaborated with our algorithm, we still need to modify a little bit of the reward function and the dimensions of inputs & outputs in order to satisfy the real situation of the robot. This is a example of the test which uses our experimental robot model:

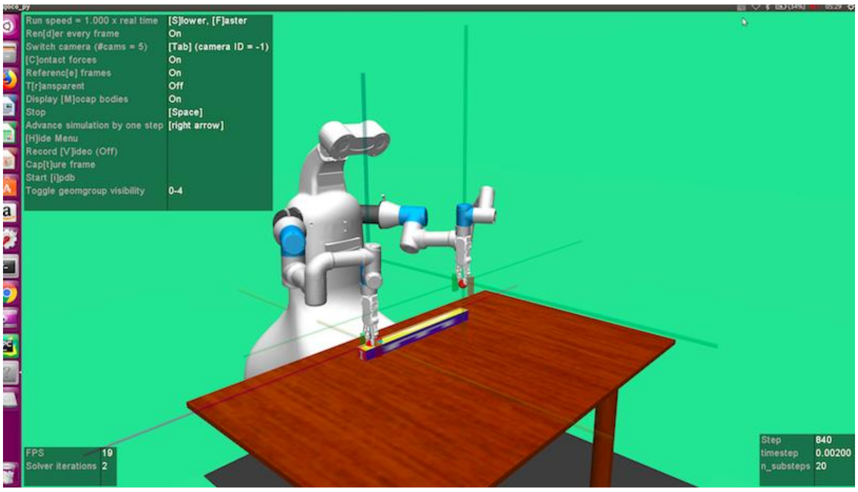


Figure 9 - Simulation Visualization of experimental robot model

The result of the pick and place for this new model situation is presented as follows:

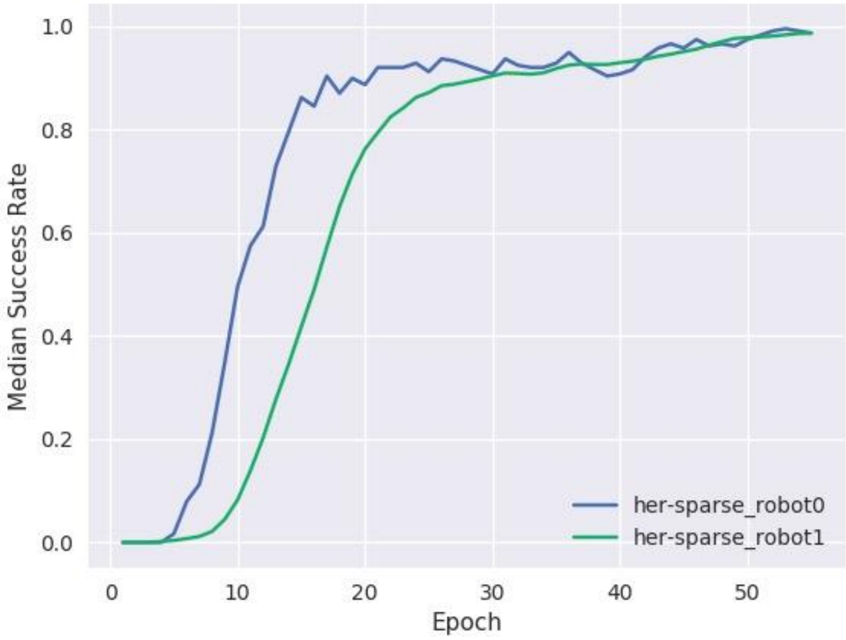


Figure 10 - Pick and Place result for the experimental robot

We can see the result is very promising, but the agent0 is more unstable than agent1 in the training process, I think it's maybe because of the cross-inputting feature in the MADDPG that leads to the first agent explore the environment and share some of the information of policy gradients to the second agent, which can make the second agent more easier to find an optimal policy compared with first agent.

In the end, this research experience is really challenging and exciting, it makes me understand better in machine learning methods especially in deep reinforcement learning field. After this research internship, I am now working with my teachers in Cranfield university to carry out more research in the related fields in this academic year and preparing my first article in this domain.