

---

**Problem Selection:**

---

I selected FetchPickAndPlace-v0 problem (<https://gym.openai.com/envs/FetchPickAndPlace-v0/>) among given problems in homework pdf.

In this problem, a robot arm should find and pick a ball and carry it on top of the target block.

---

**Deep Reinforcement Learning for Robotic Manipulation with Asynchronous Off-Policy Updates:**

---

In this paper, a reinforcement learning algorithm runs on multiple robots simultaneously and experience affects all of the robots. The system handles calculations about Q functions in a deep neural network and tests learning algorithms on door opening events.

**Similarities and Advantages:** Door handle detecting, holding and moving events are similar to ball detecting, picking and moving events. Also parallelism is an advantage in terms of run time performance.

**Diversities and Disadvantages:** Paralel calculations on more than one robot is an important performance speedup but it requires a lot of resources(both in money and development). In addition in FetchPickAndPlace-v0 problem environments, there is only one robot. Calculations should be able to deploy more than one environment in order to make calculations parallel. With low number of robots, there may be lack of data for deep neural network to evolve a proper state.

---

**Robotic Arm Control and Task Training through Deep Reinforcement Learning:**

---

In this paper there is a comparison between Trust Region Policy Opmtimization and Deep Q-network with Normalized Advantage Functions. Note that NAF and Deep Neural Networks are used in paper above as well. It gives better results with TRPO.

**Similarities and Advantages:** One of the crucial parts of this paper is real world experiments. It proves these algorithms can actually work. Also these real world experiments includes object picking and transporting tasks which we have similars in our problem as well.

**Diversities and Disadvantages:** There are not much disadvantages in this paper but it has more specific details than our problem. We can say that experiments in this paper involves our problem. But our problem contains just 3-4 tasks(they are few but not easy), we can just focus on these problems and make a more efficient system just for these tasks.

---

**Modelling:**

---

We know that our goals are in this order:

- Finding the ball.
- Picking the ball.
- Finding the box.
- Placing the ball on the box.

We can create logical functions, preconditions, states etc. in order to model this problem in logical level and implement a search algorithm. But the environment is not known to the robot. There should be explore actions in order to achieve the goals.

**Inputs/Outputs:** Inputs comes from environment. They can be:

- Empty(An empty place).
- Ball.
- Box.
- Grabbed the ball.
- Released the ball.

**Algorithms:** We can run a Forward or Backward chain algorithm on logical statements that we extracted from the problem. But we need exploration actions. Exploration can be done randomly with a proper memory. Also computer vision can be used in order to detect the ball and the box if the environment allows it

**Advantages and Disadvantages:** With a clear paperwork, we can obtain logical sentences and create our knowledge base. After doing this implementation of chain algorithms are very simple and they can achieve the goals(with help of exploration phases). But these algorithms will be specifically implemented for this problem. They will not be generic. And exploration actions should not be stuck in unnecessary loops. Algorithms can be fast but only for this problem.

**Note:** Due to unknown states that came from exploration actions a reinforcement learning technique can also be implemented for this problem.

---

**References:**

---

**Deep Reinforcement Learning for Robotic Manipulation with Asynchronous Off-Policy Updates:** Written by Shixiang Gu, Ethan Holly, Timothy Lillicrap, Sergey Levine - <https://arxiv.org/pdf/1610.00633.pdf>

**Robotic Arm Control and Task Training through Deep Reinforcement Learning:** Written by Andrea Franceschetti, Elisa Tosello, Nicola Castaman and Stefano Ghidoni - [http://www.dei.unipd.it/~toselloe/pdf/IR0S18\\_Franceschetti.pdf](http://www.dei.unipd.it/~toselloe/pdf/IR0S18_Franceschetti.pdf)