Research Experience for Excellent Students D - Dkalim Program

Research Experience in Planning and Deep Reinforcement Learning Following the Article:

Solving Hard AI Planning Instances Using Curriculum-Driven Deep Reinforcement Learning

Guide:

Professor Ronen Brafman

Participant

Noam Atia 311394241

1. Background

The Faculty of Natural Sciences of Ben Gurion University of the Negev promotes research in the fields of mathematics, computer science, physics, chemistry, life sciences and earth sciences.

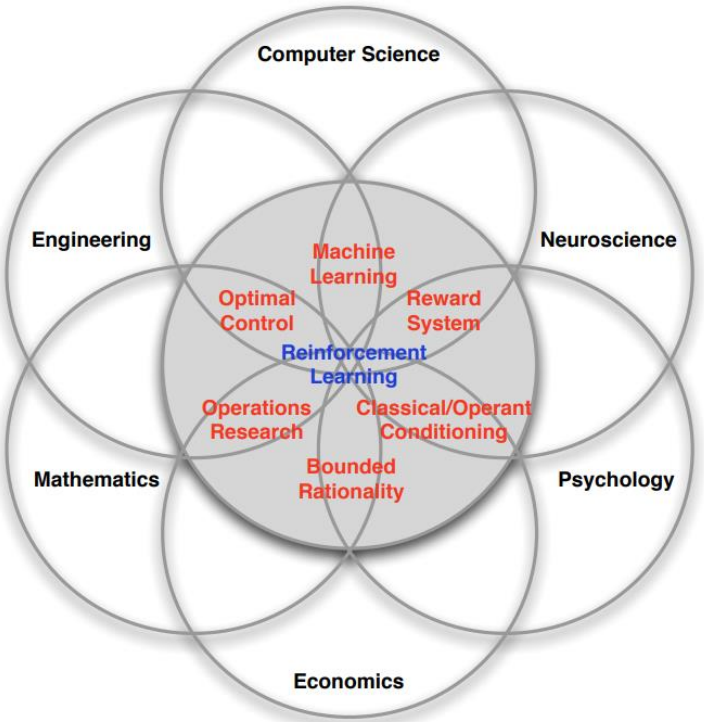
The purpose of Dkalim program is to provide tools for excellent undergraduate students, and to integrate them into research towards postgraduate studies. The research that will be carried out will form the basis for future research in the master's degree.

During the last semester, I had the honor to work alongside Professor Ronen Brafman and to be exposed to research in the fascinating fields of Planning and Deep Reinforcement Learning. Those fields are at the forefront of the world of AI, and it was exciting for me to study them during the bachelor’s degree.

2. Research

2.1 Introduction to Reinforcement Learning with David Silver

The first step was to study the exciting area of Reinforcement Learning by the fantastic course of David Silver. Those ten lectures gave me good basis and understanding of this ML field that I wasn't familiar with before. RL integrates within it so many fields I'm so excited about: CS, ML, Neuroscience and Psychology, and after taking this course I came to the realization that I would be very happy to research RL during my master's degree.



2.2 Artificial Intelligence a Modern Approach

With this wonderful book I studied about Searching and Planning. Those subjects are main core of AI, and it was great to get familiar with them. Although they combine one of my favorite areas, Logic, I can say it was harder for me to study about them, relative to RL.

2.3 Monte Carlo Tree Search

I was reading few articles about Monte Carlo Tree Search to get familiar with that. I was finding this heuristic search algorithm super interesting.

2.4 Planning Seminar

Once every two weeks I attended a planning seminar of the research group of Professor Brafman and Professor Guy Shani.

2.5 Implementing AlphaZero-style Monte Carlo Tree Search

On "Solving Hard AI Planning Instances Using Curriculum-Driven Deep Reinforcement Learning" article, Dieqiao Feng, Carla P. Gomes and Bart Salm, have tried to solve the Sokoban PSPACE-complete planning task, which represents one of the hardest domains for current AI planners. Their AlphaZero-style Monte Carlo Tree Search is the first one to solve hard instances within one day of training while other modern solvers cannot solve these instances within any reasonable time limit. On the last few months, I was trying to reimplement their algorithm by the article.

2.5.1 Sokoban Environment

The first challenge was to write a Sokoban Environment that is adapted to the description of the article and can describe a Sokoban state and it's change by the algorithm decision (State.py). I was based on gym-Sokoban environment, that the main problem with it was that it can't create a new instance by a previous instance and a move, a very important feature for MCTS. In addition, my Sokoban state has an appropriate input for the Neural Network, its actions are pushing boxes and not a single agent move and it's printable.

2.5.2 Monte Carlo Tree Search

The second challenge was to create a MCTS algorithms, which holds a Sokoban State and makes its heuristic for the rollout phases is a Neural Network (MonteCarloTreeSearchNode.py). I was used this article for the implementation, and currently the algorithm can make a decision while rolling out according to the NN, but still can't train the NN after the phase is ending. I was so excited when the algorithm could solve simple instances of Sokoban when it's heuristic was a random choice.

2.5.3 Artificial Neural Network

The last current step was to create a simple ANN that the MCTS can make decisions according to it (Net.py). The net is a simple convolutional Pytorch NN that can take a Sokoban instance and give back a movement.

3. Conclusion

It was honor and great experience for me to work under the direction of Professor Ronen Brafman on the last semester. I learned a lot, I was exposed to fascinating fields that I had not been exposed to before, I tried to be part of a research group and I got great directions for future research. I would like to thank to Professor Brafman and to the Faculty of Natural Sciences for this amazing experience!