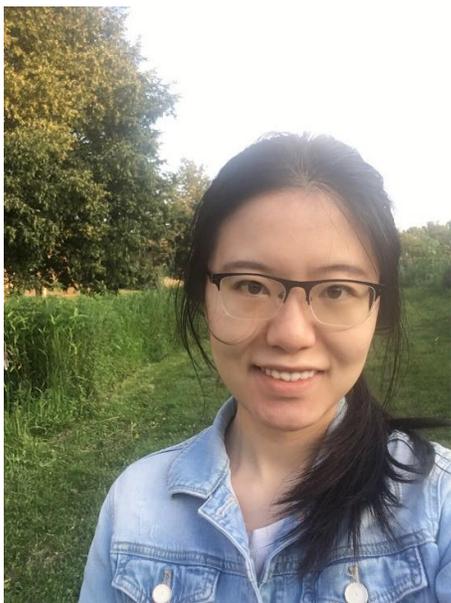




DISSERTATION DEFENSE



Yijie Guo

Sample-Efficient Algorithms for
Hard-Exploration Problems in
Reinforcement Learning

Wednesday, July 6, 2022

4:00 – 6:00pm

3725 Beyster

Hybrid – [Zoom](#)

ABSTRACT: Hard-exploration problems in reinforcement learning are characterized by sparse environment reward, large state and action space, and long time horizon. The issue of sample efficiency plagues deep RL algorithms in hard-exploration problems because it is expensive or even infeasible to explore the entire state and action space without dense and informative reward signals. This challenge calls for algorithms that cost fewer trials to collect rewarding experiences and require fewer samples to learn good policies. In this thesis, I tackle the exploration problem in two scenarios by carefully collecting and exploiting the agent’s experiences. First, I consider policy learning to solve one complex task suffering from extremely sparse rewards. I propose to imitate the agent’s past good experiences of high-rewarding transitions to indirectly drive deep exploration. In addition, I will present a new algorithm to reproduce and augment the agent’s diverse past trajectories to encourage exploration in diverse directions of the environment and increase the chance of finding the (near-)optimal solution to the hard-exploration problem. Second, I address the issue of sample efficiency in learning a shared policy to solve multiple related hard-exploration problems. I propose an action translator to transfer the good policy from one training task to any other tasks with varying dynamics. Thus, the good experiences and policy learned from one training task can benefit the data collection and further policy optimization for other sparse-reward training tasks. Also, I will discuss learning the exploration policy and then generalization on tasks with varying environment structures. I formulate a novel view-based intrinsic reward to maximize the agent’s knowledge coverage in hard-exploration problems. The agent exploits the exploration knowledge extracted from training environments and generalizes to explore unseen test environments well by enlarging its view coverage. Overall, in this thesis, I develop several advanced approaches to handle the hard-exploration problems from the exploitation perspective. This allows us to understand the challenges of hard-exploration problems more deeply and serves as an inspiration to rethink the balance between exploration and exploitation in reinforcement learning.

CHAIR: Prof. Honglak Lee