Statement of Purpose

XXX University Computer Science Ph.D.
Qiushi (Max) Lin

My research interests are broadly in theory and applications of reinforcement learning. Reinforcement learning handles the challenges of sequential decision making through interaction with environments. State-of-the-art deep reinforcement learning methods have a wide range of real-world applications, yet most of these methods lack theoretical foundations. In my Ph.D, I plan to advance our theoretical understanding of reinforcement learning, and, based on that, to develop more theoretical principled algorithms for practical purposes.

In fact, such motivation comes from my research experience in the empirical world of reinforcement learning. I have been working with Prof. Hang Ma for the past two years. We have been focusing on designing practical and efficient algorithms for partially observable multi-agent path finding. The essential drawbacks for previous reinforcement learning-based are that they cannot generalize to large-scale instances and cooperation among agents is difficult to achieve. I designed a form of explicit cooperative guidance based on heuristics and extended the multi-agent actor-critic framework to enhance the generalizability of the learned model. Inspired by a graduate-level course, Graph Representation Learning, from Prof. Oliver Schulte at SFU, I also embed a communication mechanism for better cooperation. This proposed algorithm is built upon the marriage of the planning algorithm (heuristic search) and the reinforcement learning framework (actor-critic), and based on the extensive empirical evaluation, this algorithm achieve outperformance compared to other existing works in various benchmarks, which leads to the publication of my first research paper.

For our second line of research, I turned my eyes to another variation of multi-agent path finding, moving agents in formation. This task requires agents to optimize over two different objectives of path planning and formation control, which has never been achieved by learning-based methods. The issues come from the curse of dimentionality and the complication of juggling two objectives in partially observable environments. To address these problems, I investigate in mean field reinforcement learning and multi-objective reinforcement learning, and purpose a novel approach that can handle multi-objective optimization in large-scale multi-agent reinforcement learning framework. The learned policy can lead to greater performance than planning algorithms and have the flexibility to handle dynamic formations. We summarize our results and write another paper which has been submitted and is currently under review.

Although I loved the idea of connecting the field of reinforcement learning with multi-agent path finding, I was not satisfied with the rationale underneath the proposed learning-based algorithms. Oftentimes, I would find the details in code-level implementation have much stronger impact on the performance than the actual algorithm design, which could be somewhat demotivating. I would much rather work in a field in which every concept is rigorously defined the performance of algorithms is measurable, predicable, and analyzable. Carrying the unsatisfactory from my past researches, I took two graduate-level theory courses, Theoretical Foundations of Reinforcement Learning and Optimization for Machine Learning, taught by Prof. Sharan Vaswani at SFU. In these course, I learned a wide range of knowledge for machine learning theory, from convex optimization to online learning, from multi-armed bandits to policy gradient methods, etc. Exploring all related materials opens another door for me to understanding reinforcement learning from a whole new perspective, and I find myself deeply fascinated by the theory literature. I also realize there are so many unanswered open questions in this field, and without solving these questions, researchers would never get a hold of the essence of reinforcement learning algorithm design.

Motivated by my appreciation for theory and my confusion from empiricism, I made my first attempt for doing researches in reinforcement learning theory. Guided by Prof. Sharan Vaswani, my teammates and I work on analyzing the convergence rates for policy gradient methods with linear function approximation. We developed a general recipe that can reduce optimizing over the log-linear policy class to the tabular softmax settings. With this recipe, we extend the theoretical guarantees of those softmax policy gradient methods, and propose new algorithms for log-linear policies that can enjoy the similar convergence rate.

Even though the pivot of my research interests was recent, I would argue that my motivation has always been a natural match for the subject. My favorite problems to think are those that are concise and abstract, and my favorite solutions to problems are those that are simple and elegant, yet require deep and sophisticated thoughts. Such inclination provides more joy for me to do theory research than empirical reinforcement learning. Besides, through hard studying in Prof. Vaswani's courses, I have been equipped with basic knowledge to understand most cutting-edge theory literature and come up with novel ideas, and now, through this application, I aim to find a place to continue doing so.

I am drawn to XXX university for ... [words specific to every university and POIs]