# Research Statement

**Xueguang Lyu**

Ph.D. Candidate, Northeastern University
Advisor: Christopher Amato

## Introduction

I am a Ph.D. candidate in Computer Science at Northeastern University, advised by Prof. Christopher Amato. My primary research interests lie in Multi-Agent Reinforcement Learning (MARL) and robotics. Driven by a motivation to improve agent coordination of multi-agent systems, my doctoral work has focused on first understanding the fundamental trade-offs in MARL, and then developing ways to coordinate agents *optimally* in complex environments and aim to apply reinforcement learning techniques directly to robotic systems.

## Research Contributions

**Independent Learners** In my early research (AAMAS 2020), I used distributional RL in MARL and introduced Likelihood Quantile Networks (LQN) to address non-stationarity challenges in fully decentralized MARL scenarios. This work primarily focused on the pathologies of independent learners in MARL, although the pathologies are known to be impossible to completely solve, the proposed IQN was effective nevertheless by discounting updates influenced by teammates' exploratory actions. This approach was essentially introducing a well-understood and annealing bias to the learning process. To completely solve the pathologies of independent learners, I moved on to centralized training with decentralized execution (CTDE) which in theory provide a way to solve the pathologies.

**Centralized Training, Decentralized Execution: Centralized Critics** A significant part of my research afterward addresses centralized critics in MARL. Initially, I explored trade-offs between centralized and decentralized critics, first time proving that centralized critics *cannot* enhance learning efficiency by mitigating multi-agent credit assignment issues (*AAMAS 2021, Best Paper Nominee*). Expanding this research, our AAAI 2022 paper, *A Deeper Understanding of State-Based Critics in Multi-Agent Reinforcement Learning*, provided a theoretical analysis of how different state-information availability influences policy optimization in MARL. I extended these insights further in a comprehensive journal publication in the *Journal of Artificial Intelligence Research (JAIR 2023)*, titled *On Centralized Critics in Multi-Agent Reinforcement Learning*, formally characterizing the trade-offs inherent in centralized critic approaches. Part of this work was presented as a Spotlight talk at RLDM 2022. This theoretical foundation for understanding centralized critics is crucial for developing robust learning algorithms in complex multi-robot systems where centralized training but decentralized execution is required. The Local Advantage Actor-Critic (LAAC) method, introduced at MRS 2021, enables agents to learn using only local observations and rewards, enhancing resilience to partial observability and non-stationarity. LAAC was shown to stabilize learning effectively in multi-robot coordination benchmarks, achieving robustness against environmental disturbances, and was honored with a Best Paper nomination at MRS 2021. However, the theoretical analysis of centralized critics also yield the conclusion that policy gradient methods cannot find the optimal solution in Dec-POMDP (same goes for value decomposition methods). In short, I find that a complete overhaul of MARL is needed, which leads to my current research.

## On-Going Research

My current research is aimed at solving cooperative tasks *optimally* in MARL. I reformulated current MARL methods as Generalized Multi-Agent Local Policy Iteration (G-MALPI) (encompassing such as VDN, QMIX, QPLEX, MAPPO, IPPO etc.), and show that they do not guarantee policy improvement. Recognizing the fundamental limitations of current methods, I developed a novel approach named Multi-Agent Policy Search (MAPS). MAPS uniquely combines the rigorous optimality guarantees of dynamic programming with the scalability afforded by deep reinforcement learning, effectively bridging the longstanding gap between theoretical completeness and practical applicability.

The MAPS framework allows it to provably converge to globally optimal solutions in Decentralized Partially Observable Markov Decision Processes (Dec-POMDPs). Initially validated in tabular domains, MAPS-based methods provided robust convergence guarantees and demonstrated significant performance improvements over established MARL methods such as VDN, QMIX, and QPLEX, particularly in benchmark environments that emphasize tight coordination. This research also questions the "correct" way to perform policy iteration in MARL. I therefore propose Multi-Agent Generalized Sweeping Policy Iteration (MASGPI), which is general theoretical framework that have convergence guarantees in Dec-POMDPs. First part of this ongoing research, including MAPS and G-MALPI, are finalized and was submitted to NeurIPS 2025.

**Industry Experience: On-Robot Reinforcement Learning** Beyond academia, my role as an Applied Scientist Co-op at Amazon Robotics (2023) involved applying RL **directly train robotic manipulation tasks without relying on simulation**. I developed methods emphasizing sample efficiency and safe exploration, addressing real-world challenges such as noisy sensor data and hardware constraints. This experience solidified my skills in deploying theoretical RL algorithms on practical robotic systems, bridging the gap between simulation and reality. This work aligns with the robotics community's broader goal of making reinforcement learning practical for industrial applications, addressing the sim-to-real gap that has been a persistent challenge in the field.

## Current and Future Directions

My contributions have received recognition in the community, including Best Paper nominations at AAMAS 2021 and MRS 2021, and a Spotlight presentation at RLDM 2022. Looking forward, I am particularly excited about integrating my theoretical MARL advances with real-world robotic applications. The principled approaches I've developed for coordinating agents could be transformative for domains like warehouse robotics, autonomous driving, and human-robot interaction.

In summary, my expertise spans reinforcement learning, multi-agent systems, and robotic control, combining theoretical innovation with practical implementation. I am passionate about advancing efficient and robust learning in multi-agent and robotic systems and look forward to contributing further to cutting-edge research in a postdoctoral capacity, potentially collaborating with established research groups that share this vision of principled, practical multi-agent systems.