Wei Wang Website: *www.waybaba.com* wwang828@uwo.ca

Statement of Purpose

Reinforcement Learning (RL) has experienced remarkable advancements in recent years, particularly in singular tasks such as Go, video games, and robot control. However, its broader application in real-world scenarios hinges on its ability to generalize. For example, robots trained in simulators often underperform in real-world scenarios (environment generalization), autonomous vehicles may struggle with unforeseen destinations (goal generalization) , and agents are typically confined to a single task (task generation). Enhancing the generalization capabilities of agent is critical for the next phase of RL development, a challenge I am eager to tackle through my Ph.D. research. Specifically, my **research interest** target at three key levels of generalization of RL:

- **Macro - Task Generalization:** The rise of large models in fields like Computer Vision (CV) and Natural Language Processing (NLP) marks a pivotal shift. In this context, the development of RL models capable of tackling a variety of tasks represents a significant trend, while also presents unique challenges such as integrating varied embodiments and instructions into a cohesive model, ensure a positive transfer. I am enthused by the fortuitous timing of my Ph.D. commencement, coinciding with a crucial phase in RL's evolution.
- **Medium - Goal Generalization:** In goal-conditioned RL, the agent is expected to accomplish a task g from a set $\mathcal{G}$, rather than a single, fixed task. A potential research avenue is investigating when and how to generalize the RL model for tasks $g \notin \mathcal{G}$. Additionally, I am interested to expand $\mathcal{G}$ broader, such as a set of open-ended language instructions, linking back to my prior research [2].
- **Micro - Environment Generalization:** My interest also lies in theoretically exploring the effects of environmental shift on RL model performance. I am drawn to this research style, which combining theoretical foundations with practical applications, blending mathematical elegance with practical value, as in my prior studies [3, 4].

My background in model generalization and RL is a perfect match for this endeavor. I have both experience of projects focusing on reinforcement learning ([1] at ICLR 2024 as Spotlight and [2] at NeurIPS Workshop) and model adaptability/generalization ([3] at CVPR, [6] at TKDD and [5, 4] at AAAI) .

Discovering My Passion: In truth, my background tells a curve story. During my undergraduate years, I discovered a profound interest in robotics, particularly in the "brain" aspect, which involves complex mathematics and intricate programming. This fascination is further confirmed by my internship in Horizon Robotics, steering me toward Computer Science (CS) for my graduate studies.

Recognizing the need to bridge the educational gap from my undergraduate studies, I enrolled in two additional CS courses during my master's program. The transition, combined with the challenges posed by the pandemic, did not affect my academic pursuit. Instead, it fueled my passion, culminating finishing all courses in two terms and a perfect 4.0 GPA. This accomplishment not only reflects my commitment but also underscores my genuine enthusiasm for my newfound academic path. Particularly, I found that the RL course, enriched by a skilled instructor and the lively examples in Richard Sutton's book, align perfectly with my interest in the "brain" design of robotics.

Research on Generalization: My interest in RL encountered a challenge, as my department lacked a specialized focus in this area. Thus, I chose to collaborate with Prof. Boyu Wang, an expert in machine learning theory. The reason is that machine learning theory encompasses a broad scope. This allows me to study common problems in both RL and ML, and to study the common mathematical tools during research.

Our first project is [6], in which we study how to update model under continually changing environment, which was finally accepted in TKDD. Subsequently, our research concentrated on generalization in evolving environments. For instance, Imagine a scenario where it gets darker and darker, and a robot, which usually relies on visuals, now has to depend more on sounds. Thus exploiting the evolving pattern of environment may help to adjust the model to future trend. This resulted in two papers offering distinct solutions: one employing data augmentation techniques for generating data in unseen domains, published at AAAI 2023 [5], and another is model the trends with prototypes, published at AAAI 2024 [4].

I also actively sought opportunities beyond my university, which led to a collaboration with Zhun Zhong at the University of Nottingham. Generalization can be improved not only by enhancing model robustness during training. In this project, we adapt the models dynamically during the inference stage, resulting a paper [3] accepted at CVPR 2023.

Generalization meets RL: I am always looking for opportunities to the combination of my generalization expertise and intrinsic passion for RL. At the end of my second year of M.Sc., I chose to defer my thesis, although I had enough publications to complete it. This decision allowed me to fully immerse myself in RL research, necessitating temporary departure from campus for collaborative work. Fortuitously, I secured a internship position at Microsoft Research Asia (MSRA) to work alongside Dongqi Han on cutting-edge RL research.

Our first project addressed environment generalization in RL, especially when signal delays. We framed the delay issue within the context of Partial Observation MDPs, conducted theoretical analyses, and proposed methods to preserve Markovian properties. This work [1] has been accepted to ICLR 2024 and selected as spotlight presentation (Top 5.01%).

The second project explores goal generalization in RL. We innovated a gradient-guided mechanism that empowers RL agents to pursue open-ended goals. This approach represents a significant extension of restricted goal sets in conditioned RL. The outcomes [2] was accepted and presented at the NeurIPS 2023 Workshop.

Teaching Experience: Besides the research experience, as a teaching assistant for 7 courses, I developed the ability for explaining complex concepts and aiding student understanding, improving my communication skills and deepening my subject matter expertise.

Career Goals: My long-term aspiration is to lead a research team in academia or industry. I have truly enjoyed collaborating with different people on topics that I'm passionate about. I enjoy the feeling of staying up-to-date with the latest research advancements and making a real-world impact through research.

References

[1] **Wei Wang**, Dongqi Han, Xufang Luo, Dongsheng Li. "Addressing Signal Delay in Deep Reinforcement Learning". In: *International Conference on Learning Representations (**ICLR Spotlight Top 5.01%**)*. 2024. URL: https://openreview.net/pdf?id=Z8UfDs4J46.

[2] **Wei Wang**, Dongqi Han, Xufang Luo, Yifei Shen, Charles Ling, Boyu Wang, Dongsheng Li. "Toward Open-ended Tasks Solving in Reinforcement Learning". In: *Conference on Neural Information Processing Systems Workshop (**NeurIPS Workshop**)*. 2023. URL: https://openreview.net/pdf?id=eZLOaRdlmc.

[3] **Wei Wang**, Zhun Zhong, Weijie Wang, Xi Chen, Charles Ling, Nicu Sebe, Boyu Wang. "Dynamically Instance-Guided Backward-Free Test-Time Domain Adaptation". In: *Conference on Computer Vision and Pattern Recognition (**CVPR**)*. 2023. URL: https://openaccess.thecvf.com/content/CVPR2023/papers/Wang_Dynamically_Instance-Guided_Adaptation_A_Backward-Free_Approach_for_Test-Time_Domain_Adaptive_CVPR_2023_paper.pdf.

[4] Qiuhao Zeng, **Wei Wang**, Gezheng Xu, Ruizhi Pu, Jiaqi Li, Fan Zhou, Changjian Shui, Charles Ling, Christian Gagné, Boyu Wang. "Generalizing Across Temporal Domains with Koopman Operators". In: *AAAI Conference on Artificial Intelligence (**AAAI**)*. 2024. URL: https://arxiv.org/abs/2206.00047.

[5] Qiuhao Zeng, **Wei Wang**, Fan Zhou, Charles Ling, Boyu Wang. "Data Augmentation for Domain Generalization in Non-Stationary Environments". In: *AAAI Conference on Artificial Intelligence (**AAAI**)*. 2023. URL: https://ojs.aaai.org/index.php/AAAI/article/view/26320/26092.

[6] Changjian Shui, **Wei Wang**, Ihsen Hedhli, Chi Man Wong, Feng Wang, Boyu Wang, Christian Gagné. "Lifelong Online Learning from Accumulated Knowledge". In: *ACM Transactions on Knowledge Discovery from Data (**TKDD**)*. 2021. URL: https://dl.acm.org/doi/10.1145/3563947.