Huangyuan Su Fall 2024

Personal Statement

My past research interest lies in the intersection of reinforcement learning (RL) with robotics, computer vision, language models, and game theory, etc. What excites me the most is to develop agents or algorithms that are robust and generalizable.

Applied Reinforcement Learning and Robot Learning

As a MS in Machine Learning student at Carnegie Mellon University, I have had the privilege to work under the supervision of Prof. Jeff Schneider and Prof. Ruslan Salakhutdinov. My first project at CMU found that the anchor embeddings used by recent trajectory prediction approaches can parameterize distinct discrete modes representing high-level driving behaviors. We propose to perform fully reactive closed-loop planning over these discrete latent modes, allowing us to tractably model the causal interactions between agents at each step. We validate our approach on a suite of dynamic merging scenarios, finding that our approach avoids the frozen robot problem which is pervasive in conventional planners. Our approach outperforms the previous art in CARLA on challenging dense traffic scenarios when evaluated at realistic speeds.

In the next project [4], we use diffusion models to perform language-guided closed-loop planning for autonomous driving. Counter to conventional wisdom, we show that excessive conditioning information can hinder downstream planning performance due to poor generalization and lack of diversity. Additionally, we show that gradient-based optimization (e.g. classifier guidance) can actually perform worse than sampling-based optimization (e.g. CEM/ES) for guided diffusion sampling. To this end, we introduce a novel sampling-based guidance method DiffusionES which handles arbitrary non-differentiable black-box objectives without retraining. We use LLMs to map language instructions to programs which adaptively control the behavior of our planner through reward shaping. With language supervision, our approach can synthesize highly complex behaviors (e.g. aggressive lane weaving) not present in the training data. Finally, we show that our approach can be used to solve the hardest nuPlan scenarios through language feedback from a human expert.

After submitting these two projects to ICRA 24 and CVPR 24, we started work on using video games for studying human-AI or multi-agent collaborative decision making in disaster management.

With Prof. Salakhutdinov, we have been working on solving robotic control problems under domain randomization across tasks and robot morphologies (hand, dog, arm) by distilling classical optimization algorithms (MJPC) into visuomotor policies.

Improving Reinforcement Learning Algorithms using Theoretical Tools

In my first research project [1] during my bachelor at Nanyang Technological University, we tackled the failure of standard Policy iteration (PI) relies on Bellman's Principle of Optimality under time-inconsistent (TIC) objectives, such as non-exponentially discounted reward functions. Specifically, we consider an infinite-horizon TIC RL setting and formally present an alternative type of optimality drawn from game theory: subgame perfect equilibrium, that attempts to resolve the aforementioned questions. Drawing on these observations, we propose backward Q-learning, a new algorithm in the approximate PI family that targets SPE policy under non-exponentially discounted reward functions.

Then [2], I noticed that the learned representation of a Q-network and its target should, in theory, satisfy a favorable distinguishable representation property. Specifically, there exists an upper bound on the representation similarity of the value functions of two adjacent time steps in a typical DRL setting. However, in experiments, the DRL agents may violate this property and obtain a suboptimal policy. Therefore, we propose a simple yet effective regularizer called Policy Evaluation with Easy Regularization on Representation, to maintain the distinguishable representation property. These two works are published in TMLR and CVPR 23, respectively.

Future plan

In my previous research endeavors, a significant portion of my time was dedicated to refining model architectures and hyperparameters. This challenge is exacerbated in the realm of RL, where the optimization of distinct parameter sets for different scenarios within the same task is common practice. This intricate process poses a significant obstacle when attempting to deploy our agents in real-world settings. Issues such as data shifts and long-tail scenarios often lead to system failures. Given the current momentum in industry towards integrating AI workers to complement or substitute human roles, it becomes imperative for these AI entities to possess specific qualities, including robustness, safety, and consistency. Additionally, an essential attribute is their capacity for reasoning, enabling them to produce decisions that are not only refined but also explainable to humans. The ability to incorporate feedback for continuous improvement is crucial, as the consequences of inadequate performance can be severe.

While certain studies [5, 6] suggest that agents based on LLMs outperform RL agents or exhibit zero-shot generalization, it is evident that their generalization capabilities remain constrained. Consequently, I strongly advocate for the development of better architectures, frameworks, or training schemes. For the first year of my PhD, I plan to explore various avenues, including leveraging generative (pre-)training for robust 3D recognition, distilling knowledge from pretrained LLMs to mitigate the intensive sample requirements for training autonomous agents, and addressing the paucity of theoretical analysis in deep learning methods.

Looking ahead, my career goal is to work as a professor, contributing to research that yields trustworthy AI agents or systems, while mentoring and guiding motivated students.

Why Columbia University and Computer Science Ph.D.?

Based on my experiences, I've come to recognize that effective problem-solving and the discovery of new, compelling challenges require guidance from the most talented, insightful, and experienced researchers. Among them, I am particularly drawn to the distinguished faculty at Columbia University, including Junfeng Yang, Carl Vondrick, Zhou Yu, Richard Zemel, and Tony Dear. Their groundbreaking contributions in vision, NLP, and robotics, occasionally influenced by RL, strongly resonate with my research interests. A thorough examination of their research, coupled with insights gained from perusing their lab environments, has strengthened my desire to collaborate with them. In enrolling in this CS PhD program, my foremost goal is to undergo systematic and comprehensive research training, honing my skills in problem-solving, enriching my expertise in my specific area while fostering a broad understanding of the field. The program's resources for students planning faculty careers is particularly attractive, as it aligns seamlessly with my long-term aspirations of contributing meaningfully to academia.

Huangyuan Su Fall 2024

References:

[1] Lesmana, Nixie S., **Huangyuan Su**, and Chi Seng Pun. "Reinventing Policy Iteration under Time Inconsistency." *Transactions on Machine Learning Research* (2022).

[2] Qiang He, **Huangyuan Su**, Xinwen Hou, Yu Liu. “Frustratingly Easy Regularization on Representation Can Boost Deep Reinforcement Learning”. 2023 IEEE Computer Society Conference on Computer Vision and Pattern Recognition.

[3] Brian Yang, **Huangyuan Su**, Nikolaos Gkanatsios, Tsung-Wei Ke, Ayush Jain, Jeff Schneider, Katerina Fragkiadaki. “Controllable diffusion-based planning for language-guided driving”. 2024 In submission.

[4] Adam Villaflor, Brian Yang, **Huangyuan Su**, Katerina Fragkiadaki, John M. Dolan, Jeff Schneider. “Tractable Joint Prediction and Planning over Discrete Behavior Modes for Urban Driving”. 2024 In submission.

[5] Wang, Guanzhi, et al. "Voyager: An open-ended embodied agent with large language models." *arXiv preprint arXiv:2305.16291* (2023).

[6] Wu, Yue, et al. "SPRING: GPT-4 Out-performs RL Algorithms by Studying Papers and Reasoning." *arXiv preprint arXiv:2305.15486* (2023).