YIQING XU

yiqing.xu@u.nus.edu ⋄ Personal Website ⋄ Google Scholar⋄ LinkedIn

EDUCATION

National University of Singapore   Aug, 2016 - Jun, 2020
B.A. Computer Science   GPA: 4.61
B.A. Applied Mathematics   GPA: 4.79
National University of Singapore   Aug, 2020 - Jun, 2025 (Expected)
Ph.D Computer Science   GPA: 5.00

RESEARCH HIGHLIGHTS

My research on Robot Learning focuses on translating human objectives into signals for machine optimization. I design algorithms and learn representations to align AI agents more closely with human objectives, enabling them to better understand and assist humans.

I've been working on algorithm research and theoretical analysis of inferring reward functions (Inverse RL) from limited data. Recently, I've used large language models to learn reward functions for motor skills through self-alignment. My ongoing work addresses the challenge of defining and solving robotics tasks with ambiguous and multi-factorial objectives, such a "tidying up a table", using more flexible and comprehensive frameworks than conventional reward mechanisms.

RESEARCH EXPERIENCE

Visiting Ph.D Student   Sep, 2023 - present
CSAIL MIT, Prof. Leslie Kaelbling and Prof. Tomás Lozano-Pérez
- Study how to ground the common-sense and multi-factorial objective such as "tidiness" into a set of object relations and optimize them jointly.

Ph.D Student   Aug, 2020 - present
SoC NUS, Prof. David Hsu
- Study how to scale up reward learning (Inverse RL) to long-horizon and high-dimensional tasks.
- Build a simulator and benchmark for the Differentiable Deformable Object Manipulation methods.
- Study how to learn/represent human objectives through various data forms such as demonstration, unstructured play, and foundational models.

SELECTED PUBLICATIONS / PREPRINTS

Grounding Common-sense Objective for Tabletop Object Rearrangement [Link]   **In submission**
Y. Xu, D. Hsu
- Created a method that fuses visual and semantic commonsense reasoning for robotic tasks with vague objectives.

Learning Reward for Physical Skills using Large Language Models [Link]   **CoRL 2023 Workshop**
Y. Zeng, Y. Xu
- Created a self-alignment method to learn reward functions for a diverse set of physical skills from Large Language Models with no human demonstration or feedback.

On the Effective Horizon for Inverse Reinforcement Learning [Link]   **In submission**
Y. Xu, F. Doshi-Velez, D. Hsu
- Presented a theoretical study on how the effective horizon affects reward learning in low data regime. Proposed a new IRL framework to learn both the reward and horizon pair jointly.

Receding Horizon Inverse Reinforcement Learning[Link]   **NeurIPS 2022**
Y. Xu, W. Gao, D. Hsu
- Enhanced inverse reinforcement learning for long-horizon and high-dimensional state/action spaces by sequentially learning and integrating short-horizon rewards.

Benchmarking Deformable Object Manipulation with Differentiable Physics[Link]   **ICLR 2023**
S. Chen, Y. Xu*, C. Yu*, L. Lin, D. Hsu*
- Built a simulator and benchmark for differentiable deformable object manipulation methods.