

# Yue (Sophie) Guo

## EDUCATION

---

**Carnegie Mellon University** Pittsburgh, PA, US Aug 2018 - May 2024  
PhD in Computer Science, advised by Prof. Katia Sycara

**Brown University** Providence, RI, US Aug 2014 - May 2018  
BS in Applied Math - Computer Science, Magna Cum Laude, Honors

## INDUSTRY EXPERIENCE

---

**TikTok Inc.** - Machine Learning Scientist San Jose, California, US Jul 2024 - Current  
Recommendation team, PhD Graduate.  
Model adaptation and transfer from the main prediction model.

**PlusAI** - Software Engineer Intern Santa Clara, California, US Jun 2022 - Aug 2022  
Vehicle intention prediction across markets.  
Utilized reinforcement learning to create a parameter selector transferring the prediction network.

**Fields Institute** - Research Intern Toronto, Canada Jul 2017 - Aug 2017  
Simulated a Rapidly Spinning Baseball with fluid dynamics and GPU Computation.

**XiTian Tech** - Research Intern Shanghai, China Jun 2015 - Aug 2015  
Applied Machine Learning methods to detect abnormal patterns in patients' data.

## RESEARCH EXPERIENCE

---

**Carnegie Mellon University** School of Computer Science Pittsburgh, PA, US Oct 2018 - May 2024  
Advanced Agent-Robotics Technology Lab, directed by Prof. Katia Sycara

1. Thesis Work: reinforcement learning policy transfer (collaborated with Honda Research Institute)  
Developed an advisor policy to generate explanations with advice and evaluate its transferability in unseen environments, allowing introspection based on outcome estimation.

Leveraged large language models to advise human students, reducing hallucination, building trust, and improving prediction networks for adaptive advising.

## 2. Intelligence for Successful Teams

Developed a system for urban search and rescue missions, modeling rescuers' navigation as hierarchical graphs, deploying a transferable navigation network, and estimating coordination time.

## 3. Ethical Norms for Autonomous Agents

Created a normative reasoning framework, deriving rewards from norm-compliant trajectories and resolving norm conflicts for ethical reasoning in new scenarios.

**Brown University** Department of Computer Science Providence, RI May 2016 - May 2018

Data Science & Database Lab, directed by Prof. Tim Kraska and Prof. Carsten Binnig

Developed efficient methods for Simpson's Paradox detection.

Intelligent Robot Lab, directed by Prof. George Konidaris

Worked with policy optimization in multi-task reinforcement learning for optimal action priors.

Honors Thesis: Explored unfixed transition functions for reinforcement learning.

## PUBLICATION

---

1. M. Lin, S. Shi, **Y. Guo**, B. Chalaki, V. Tadiparthi, S. Stepputtis, J. Campbell, and K. Sycara. "*Navigating Noisy Feedback: Enhancing Reinforcement Learning with Error-Prone Language Models*" In the Conference on Empirical Methods in Natural Language Processing (EMNLP), 2024
2. **Y. Guo**, X. Zhang, S. Stepputtis, J. Campbell, and K. Sycara. "*Adaptive Action Advising with Different Rewards*" In the Proceedings of Conference on Lifelong Learning Agents (CoLLAs), 2024.
3. M. Lin, S. Shi, **Y. Guo**, B. Chalaki, V. Tadiparthi, S. Stepputtis, J. Campbell, and K. Sycara. "*A Reward Analysis of Reinforcement Learning from Large Language Model Feedback*." In Proceedings of Workshop on Reinforcement Learning Beyond Rewards (RLBrew) Workshop at the Reinforcement Learning Conference (RLC) 2024.
4. **Y. Guo**, J. Campbell, S. Stepputtis, R. Li, D. Hughes, F. Fang, and K. Sycara. "*Explainable Action Advising for Multi-Agent Reinforcement Learning*." In the Proceedings of International Conference on Robotics and Automation (ICRA). IEEE, 2023.
5. X. Zhang, **Y. Guo**, S. Stepputtis, K. Sycara, and J. Campbell. "*Explaining Agent Behavior*

- through Natural Language Interactions.*” In Proceedings of the Human Multi-Robot Interaction (HmRI) Workshop at the 2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, 2023.
6. J. Campbell, **Y. Guo**, F. Xie, S. Stepputtis, and K. Sycara. "*Introspective Action Advising for Interpretable Transfer Learning.*" In the Proceedings of Conference on Lifelong Learning Agents (CoLLAs), 2023.
  7. **Y. Guo**, I. Yang, Y. Wang, K. Sycara. "*Reinforcement Learning Methods for Network-based Transfer Parameter Selection.*" In the Intelligence & Robotics (Journal), 2023.
  8. **Y. Guo**, R. Jena, D. Hughes, M. Lewis, K. Sycara. "*Transfer Learning for Human Navigation and Triage Strategies Prediction in a Simulated Urban Search and Rescue Task.*" In the Proceedings of International Symposium on Robot and Human Interactive Communication (RO-MAN). IEEE, 2021.
  9. **Y. Guo**, B. Wang, D. Hughes, M. Lewis, K. Sycara. "*Designing Context-Sensitive Norm Inverse Reinforcement Learning Framework for Norm-Compliant Autonomous Agents.*" In the Proceedings of International Symposium on Robot and Human Interactive Communication (RO-MAN). IEEE, 2020
  10. D. Hughes, A. Agarwal, **Y. Guo**, K. Sycara. "*Inferring Non-Stationary Human Preferences for Human-Agent Teams.*" In the Proceedings of International Symposium on Robot and Human Interactive Communication (RO-MAN). IEEE, 2020.
  11. D. Abel, Y. Jinnai, **Y. Guo**, G. Konidaris, M. Littman. "*Policy and Value Transfer in Lifelong Reinforcement Learning.*" In the Proceedings of International Conference on Machine Learning (ICML), 2018.
  12. **Y. Guo**, C. Binnig, T. Kraska. "*What you see is not what you get!: Detecting Simpson's Paradoxes during Data Exploration.*" In the Proceedings of Workshop on Human-In-the-Loop Data Analytics (SIGMOD), ACM 2017.

## PATENT & BOOK

---

1. **Book Chapter Author:** "Enhancing Safety and Efficiency Through Explainable Transfer Learning," in *Transfer Learning - Leveraging the Capability of Pre-trained Models Across Different Domains* (working title), ISBN 978-0-85014-247-1, accepted for publication in 2024.
2. **Y. Guo**, I. Yang, Y. Wang. "Reinforcement Learning Techniques for Network-Based Transfer Learning." US Patent accepted, Notice of Allowance received in 2023.