Efficient Robot Learning via Interaction with Humans

Erdem Bıyık

Thomas Lord Department of Computer Science
University of Southern California
Los Angeles, CA, USA
biyik@usc.edu

My talk will cover the following research directions I have been working on, with the applications in robot learning:

- Why is reward learning from demonstrations suboptimal?
- How does comparison based human feedback mitigate the problems of demonstration feedback?
- How can we elicit comparison feedback in a more time- and data-efficient way using active learning methods or large pretrained models (e.g., (Wang et al. 2024))?
- What are some complementary human feedback types and how can we enable AI agents to learn from them?

To this end, I will talk about the following papers in detail: (Yang et al. 2024; Liang, Thomason, and Bıyık 2024; Korkmaz and Bıyık 2025), and the following papers very briefly: (Bıyık et al. 2021; Kwon et al. 2020; Ellis et al. 2024).

Having given an overview, let me now talk about the structure of the talk in more detail. I plan to divide my talk into three parts. First, I am going to motivate the reward learning and partner modeling problems: in many human-robot collaboration and multi-agent tasks, it is vital to model the partners and estimate their objectives to efficiently collaborate/interact with them (Zhu, Bıyık, and Sadigh 2020; Bıyık et al. 2022). After giving examples of these, I am going to make a brief introduction to the most common approaches that are used for this purpose, namely learning from demonstrations. I will conclude the first part by showing what is wrong with these existing approaches: (1) they are very data-hungry, which we cannot afford in many settings including robotics (Bıyık et al. 2021), and (2) human demonstrations are unreliable in a surprisingly large number of domains, including those we think humans perform reasonably well, e.g., driving (Kwon et al. 2020).

Having motivated the alternative types of human feedback for learning, I am going to introduce comparison-based feedback (Bıyık et al. 2021; Bıyık 2022) in the second section of the talk, and explain why it does not suffer from most of the problems that demonstrations have, but is still data-hungry. To address this, I will propose comparative language based feedback (Yang et al. 2024) and active learning techniques (Ellis et al. 2024), which will result in (1) a new type of human feedback that has not been considered in the literature before, and (2) an active querying algorithm that optimizes the information the AI agent will elicit from the human.

Finally, I will discuss what other types of human feedback exist. Specifically, this third section will describe how we went about modeling humans' gaze (Liang, Thomason, and Bıyık 2024) and interventions/corrections to the robots (Korkmaz and Bıyık 2025). I will also present our algorithms that enable AI agents to leverage these models to learn from such feedback. This part of the talk will be like a recipe for the researchers who want to leverage multimodal human feedback in learning algorithms, and will conclude my talk.

References

Bıyık, E. 2022. *Learning Preferences For Interactive Autonomy*. Ph.D. thesis, EE Department, Stanford University.

Bıyık, E.; Lalitha, A.; Saha, R.; Goldsmith, A.; and Sadigh, D. 2022. Partner-Aware Algorithms in Decentralized Cooperative Bandit Teams. In *Proceedings of the 36th AAAI Conference on Artificial Intelligence*.

Bıyık, E.; Losey, D. P.; Palan, M.; Landolfi, N. C.; Shevchuk, G.; and Sadigh, D. 2021. Learning Reward Functions from Diverse Sources of Human Feedback: Optimally Integrating Demonstrations and Preferences. *The International Journal of Robotics Research (IJRR)*.

Ellis, E.; Ghosal, G. R.; Russell, S. J.; Dragan, A.; and Bıyık, E. 2024. A Generalized Acquisition Function for Preference-based Reward Learning. In *International Conference on Robotics and Automation (ICRA)*.

Korkmaz, Y.; and Bıyık, E. 2025. MILE: Model-based Intervention Learning. In *(Under review)*.

Kwon, M.; Bıyık, E.; Talati, A.; Bhasin, K.; Losey, D. P.; and Sadigh, D. 2020. When Humans Aren't Optimal: Robots that Collaborate with Risk-Aware Humans. In *ACM/IEEE International Conference on Human-Robot Interaction (HRI)*.

Liang, A.; Thomason, J.; and Bıyık, E. 2024. ViSaRL: Visual Reinforcement Learning Guided by Human Saliency. In *International Conference on Intelligent Robots and Systems (IROS)*.

Wang, Y.; Sun, Z.; Zhang, J.; Xian, Z.; Bıyık, E.; Held, D.; and Erickson, Z. 2024. RL-VLM-F: Reinforcement Learning from Vision Language Foundation Model Feedback. In *International Conference on Machine Learning (ICML)*.

Yang, Z.; Jun, M.; Tien, J.; Russell, S. J.; Dragan, A.; and Bıyık, E. 2024. Trajectory Improvement and Reward Learning from Comparative Language Feedback. In *Conference on Robot Learning (CoRL)*.

Zhu, Z.; Bıyık, E.; and Sadigh, D. 2020. Multi-Agent Safe Planning with Gaussian Processes. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*.