Research Statement – Sarmad Mehrdad

My research aims to make robot learning data-efficient and interpretable. While much of the field increasingly gravitates toward large end-to-end models trained on massive datasets (as noted in recent panels and perspectives [1,2]), I pursue a complementary direction—one that leverages **Inverse Reinforcement Learning (IRL)** and **Optimal Control (OC)** to infer intent and cost structure from minimal demonstrations. This approach is critical for human–robot interaction (HRI), where demonstrations are scarce, interactions are continuous and individualized, and interpretability is as vital as performance.

The core question that drives my research is not how to move a robot, but *why* and *when* it should move. Modern robotics has mastered fast and adaptive motion generation, yet our notion of "robot learning" remains tied to the size of available datasets and computational power, otherwise risking further limitations in learning. Humans, however, achieve generalization from few examples, learning not only the *task* but the *intention* behind it. Robots, in contrast, often depend on exhaustive datasets to acquire even simple behaviors. Achieving true autonomy requires a deeper understanding of *what* to learn and *how* to represent it efficiently. My work addresses this gap by uniting the quantitative rigor of control theory with the inferential structure of learning from demonstration.

During my Ph.D., I developed a novel **Minimal-Observation Inverse Reinforcement Learning (MO-IRL)** framework that estimates optimal cost functions from minimal expert data. The algorithm iteratively refines the cost estimate by weighting non-optimal trajectories according to their deviation from optimality, yielding a more informative sampling distribution than random local perturbations. This approach achieves **faster, more interpretable learning** without requiring large sets of "bad" trajectories. Our results, accepted at IEEE IROS 2025 [3] and Humanoids 2025 [4], demonstrate significant gains in sample efficiency and convergence stability. Building on this, in my ongoing work I am extending MO-IRL to **physical human–robot interaction (pHRI)** for *online cost learning from human corrections*, translating human feedback into interpretable cost adjustments. In parallel, I am analyzing **multi-stage human intent** across sequential tasks, aiming for a holistic understanding of optimality that can later generalize to humanoids and collaborative robots.

Looking ahead, in my postdoctoral research, I plan to develop **learning frameworks for embodied intelligence** that unify learning and control—enabling robots to learn efficiently, interact naturally with humans, and adapt reliably in real time.

- **Data Efficiency and Learning Optimality**
  Modern AI systems often equate learning success with data abundance. Yet optimality can emerge from *structure*, not scale. I seek to design **sample-efficient algorithms** that infer cost representations and control policies from sparse, high-value demonstrations—bridging IRL and OC to capture intent with minimal data.

- **Human–Robot Coexistence and Intention Adaptation**
  Robots often act conservatively near humans due to limited modeling of shifting human intent. My goal is to develop **multi-agent learning paradigms** for intent-adaptive collaboration that infer and adapt to latent and dynamic human strategies online, enabling robots to respond safely and seamlessly in shared workspaces. [5, 6]

- **Active and Online Learning for Continual Adaptation**
  A persistent limitation of robot learning is the dominance of offline training. I aim to create

proactive and online algorithms that incorporate feedback during execution—updating cost functions, features, or control parameters in real time for safer and more robust adaptation. [7, 8]

- **The Model Dilemma: Beyond Model-Free vs. Model-Based**
  My research seeks to transcend the dichotomy between model-free flexibility and model-based interpretability. Beyond estimating cost *weights*, I aim to infer the **features that define the underlying task structure and characterize what constitutes success in the task**, extending recent efforts in feature discovery (e.g., [9]) and feature selection for control (e.g., [10, 11]). This direction promises more **meaningful learning**—where a robot not only optimizes performance but also understands *what it has learned to care about.*

Ultimately, I envision developing **embodied learning systems** that can interpret, predict, and co-adapt with humans in real time—learning efficiently, acting transparently, and generalizing across tasks. By merging the mathematical rigor of optimal control with the flexibility of modern learning, my long-term goal as a researcher and future faculty member is to advance **explainable, data-efficient, and human-centered robotic intelligence.**

References

[1] Nancy M Amato, Seth Hutchinson, Animesh Garg, Aude Billard, Daniela Rus, Russ Tedrake, Frank Park, and Ken Goldberg. "data will solve robotics and automation: True or false?": A debate. *Science Robotics*, 10(105):eaea7897, 2025.

[2] Ken Goldberg. Good old-fashioned engineering can close the 100,000-year "data gap" in robotics, 2025.

[3] Sarmad Mehrdad, Avadesh Meduri, and Ludovic Righetti. Cost function estimation using inverse reinforcement learning with minimal observations. *arXiv preprint arXiv:2505.08619*, 2025.

[4] Maxime Sabbah, Filip Bečanović, Sarmad Mehrdad, Ludovic Righetti, Bruno Watier, and Vincent Bonnet. Minimal observations inverse reinforcement learning for predicting human box-lifting motions. In *International Conference on Humanoid Robots 2025*, 2025.

[5] Stefanos Nikolaidis, Yu Xiang Zhu, David Hsu, and Siddhartha Srinivasa. Human-robot mutual adaptation in shared autonomy. In *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*, pages 294–302, 2017.

[6] Shen Li, Nadia Figueroa, Ankit Shah, and Julie Shah. Provably safe and efficient motion planning with uncertain human dynamics. 2021.

[7] Dylan P Losey, Andrea Bajcsy, Marcia K O'Malley, and Anca D Dragan. Physical interaction as communication: Learning robot objectives online from human corrections. *The International Journal of Robotics Research*, 41(1):20–44, 2022.

[8] Andreea Bobu, Andi Peng, Pulkit Agrawal, Julie A Shah, and Anca D Dragan. Aligning human and robot representations. In *Proceedings of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*, pages 42–54, 2024.

[9] Andi Peng, Belinda Z Li, Ilia Sucholutsky, Nishanth Kumar, Julie A Shah, Jacob Andreas, and Andreea Bobu. Adaptive language-guided abstraction from contrastive explanations. *arXiv preprint arXiv:2409.08212*, 2024.

[10] Mrinal Kalakrishnan, Peter Pastor, Ludovic Righetti, and Stefan Schaal. Learning objective functions for manipulation. In *2013 IEEE International Conference on Robotics and Automation*, pages 1331–1336. IEEE, 2013.

[11] Sergey Levine, Zoran Popovic, and Vladlen Koltun. Feature construction for inverse reinforcement learning. *Advances in neural information processing systems*, 23, 2010.