

Sanjiban Choudhury

✉ sanjiban.choudhury@gmail.com | [in linkedin](#) | [github](#) | [google scholar](#) | ☎ 412-251-8061

SUMMARY

I work on interactive AI agents that self-align through few-shot interactions with humans and their environment. Specializing in reinforcement learning (RLHF), imitation learning (IRL), and foundation models for planning, robotics, and code generation. Experienced in leading research groups and deploying production ML models.

WORK EXPERIENCE

- Assistant Professor, Cornell Computer Science** Jun 2022 – Present
- Developing general-purpose, interactive AI assistants aligned with human values. I lead a research team and contribute individually to build self-improving AI and robotics systems that continually learn from interactions.
 - *LLM agents and alignment*: developing agent process reward models^[1], weak-to-strong domain adaptation^[4], learning from privileged AI feedback^[2], analyzing generator-verifier gap^[5] in RLHF with applications to planning^[6], robotics^[9] and code generation^[3].
 - *Interactive learning algorithms*: scalable inverse RL from observations alone^[15], value-aware model-based RL with safety constraints^[14], causal imitation learning^[16], query-efficient active learning with human teachers^[8].
- Senior Staff Research Engineer, Aurora Self-Driving** Sep 2019 – Present
- Developed production foundation models for decision-making and predicting agent interactions in self-driving.
- Postdoctoral Fellow, University of Washington** Mar 2018 – Sep 2019
- Learning from diverse user feedback like demonstrations^[20], interventions^[19], preferences for robot manipulation.
- AI Research Intern, Microsoft Research** May 2016 – Aug 2016
- Developed efficient algorithms for solving POMDPs by imitating privileged MDP oracles.^[21]
- Graduate Researcher, Carnegie Mellon University** Aug 2013 – Feb 2018
- Ph.D. on unified algorithmic frameworks for combining machine learning and planning.

EDUCATION

- Carnegie Mellon University** Pittsburgh, PA
Ph.D. in Robotics, School of Computer Science Aug 2013 – Feb 2018
- Carnegie Mellon University** Pittsburgh, PA
M.S. in Robotics, School of Computer Science Aug 2011 – May 2013
- Indian Institute of Technology Kharagpur** Kharagpur, India
B.S, M.S. in Electrical Engineering Aug. 2006 – May 2011

SELECTED AWARDS / GRANTS

- **NSF Career Award, 2025**
Watch, Interact, Collaborate: Online Learning Paradigms for Robot Apprentices
- **OpenAI Superalignment Award, 2024**
Task Superalignment via Domain Adaptation
- **Google Research Scholar Award, 2024**
Imitation Labs: Interactive Experiment Planning with LLM Apprentices
- **ONR Young Investigator, 2024**
Learning Long-Horizon Dextrous Manipulation from Vision Language Demonstrations
- **NSF Collaborative Grant, 2023**
Inverse Task Planning from Few-Shot Vision-Language Demonstrations
- **NSF Collaborative Grant, 2023**
Superhuman Imitation Learning from Heterogeneous Demonstrations

LLM AGENTS AND ALIGNMENT

- [1] **S. Choudhury**. “Process Reward Models for LLM Agents: Practical Framework and Directions” 2025. [\[paper\]](#) [\[code\]](#)
- [2] **S. Choudhury*** and P. Sodhi*. “Better than Your Teacher: LLM Agents that learn from Privileged AI Feedback” *International Conference on Learning Representations (ICLR)*. 2025. [\[paper\]](#) [\[code\]](#)
- [3] A. K. Jain*, G. G. Pumariega*, W. Chen, A. Rush, W. Zhao[†], and **S. Choudhury[†]**. “Multi-Turn Code Generation Through Single-Step Rewards” *International Conference on Machine Learning (ICML)*, in submission. 2025.
- [4] D. Wu and **S. Choudhury**. “Aligning LLMs with Domain Invariant Reward Models” *International Conference on Machine Learning (ICML)*, in submission. 2025. [\[paper\]](#) [\[code\]](#)
- [5] G. Swamy, W. Sun, **S. Choudhury**, Z. S. Wu, and J. A. Bagnell. “All Roads Lead to Likelihood: The Value of Interaction in Preference Fine-Tuning” *Reinforcement Learning Conference (RLC)*, in submission. 2025.
- [6] G. G. Pumariega, W. Chen, K. Kedia, and **S. Choudhury**. “Query-Efficient Planning with Language Models.” *International Joint Conferences on Artificial Intelligence (IJCAI)*, in submission. 2025. [\[paper\]](#) [\[code\]](#)
- [7] G. G. Pumariega, L. S. Yean, N. Sunkara, and **S. Choudhury**. “Robotouille: An Asynchronous Planning Benchmark for LLM Agents” *International Conference on Learning Representations (ICLR)*. 2025. [\[paper\]](#) [\[code\]](#)
- [8] H. Wang, N. Chin, G. G. Pumariega, J. Bohg, and **S. Choudhury**. “APRICOT: Active Preference Learning and Constraint-Aware Task Planning with LLMs.” *Conference on Robot Learning (CORL)*. 2024. [\[paper\]](#) [\[project\]](#)
- [9] H. Wang, K. Kedia, J. Ren, ···, and **S. Choudhury**. “MOSAIC: A Modular System for Assistive and Interactive Cooking” *Conference on Robot Learning (CORL)*. 2024. [\[paper\]](#) [\[project\]](#)
- [10] W. Zhao, J. T. Chiu, J. D. Hwang, F. Brahma, J. Hessel, **S. Choudhury**, Y. Choi, X. L. Li, A. Suhr. “UNcommonsense Reasoning: Abductive Reasoning about Uncommon Situations.” *NAACL*. 2025. [\[paper\]](#) [\[dataset\]](#)
- [11] H. Wang, G. G. Pumariega, **S. Choudhury**. “Demo2Code: From Summarizing Demonstrations to Synthesizing Code via Extended Chain-of-Thought” *Advances in Neural Information Processing Systems (NeurIPS)*. 2023. [\[paper\]](#) [\[code\]](#)

INTERACTIVE LEARNING ALGORITHMS

- [12] A. K. Jain, H. Wiltzer, J. Farebrother, I. Rish, G. Berseth, and **S. Choudhury**. “Non-Adversarial Inverse Reinforcement Learning via Successor Feature Matching” *International Conference on Learning Representations (ICLR)*. 2025. [\[paper\]](#) [\[code\]](#)
- [13] J. Ren, G. Swamy, Z. S. Wu, J. A. Bagnell, and **S. Choudhury**. “Hybrid Inverse Reinforcement Learning” *International Conference on Machine Learning (ICML)*. 2024. [\[paper\]](#) [\[code\]](#)
- [14] A. Vemula, Y. Song, A. Singh, J. A. Bagnell, and **S. Choudhury**. “The Virtues of Laziness in Model-based RL: A Unified Objective and Algorithms” *International Conference on Machine Learning (ICML)*. 2023. [\[paper\]](#) [\[code\]](#)
- [15] G. Swamy, **S. Choudhury**, J. A. Bagnell, and Z. S. Wu “Inverse Reinforcement Learning without Reinforcement Learning” *International Conference on Machine Learning (ICML)*. 2023. [\[paper\]](#) [\[code\]](#)
- [16] G. Swamy, **S. Choudhury**, J. A. Bagnell, and Z. S. Wu “Sequence Model Imitation Learning with Unobserved Contexts” *Advances in Neural Information Processing Systems (NeurIPS)*. 2022. [\[paper\]](#) [\[code\]](#)
- [17] G. Swamy, **S. Choudhury**, J. A. Bagnell, and Z. S. Wu “Causal imitation learning under temporally correlated noise” *International Conference on Machine Learning (ICML)*. 2022. [\[paper\]](#) [\[code\]](#)
- [18] G. Swamy, **S. Choudhury**, J. A. Bagnell, and Z. S. Wu “Of Moments and Matching: A Game-Theoretic Framework for Closing the Imitation Gap” *International Conference on Machine Learning (ICML)*. 2021. [\[paper\]](#) [\[code\]](#)
- [19] J. Spencer, **S. Choudhury**, M. Barnes, M. Schmittle, M. Chiang, P. Ramadge, and S. Srinivasa “Expert Intervention Learning” *Robotics: Science and Systems (RSS)*. 2020. [\[paper\]](#)
- [20] L. Ke, **S. Choudhury**, M. Barnes, W. Sun, G. Lee, and S. Srinivasa. “Imitation Learning as f-Divergence Minimization” *Workshop on the Algorithmic Foundations of Robotics (WAFR)*. 2020. [\[paper\]](#)
- [21] **S. Choudhury**, M. Bhardwaj, S. Arora, A. Kapoor, G. Ranade, S. Scherer, and D. Dey. “Data-driven Planning via Imitation Learning” *The International Journal of Robotics Research (IJRR)*. 2018. *Best Paper Finalist* [\[paper\]](#)
- [22] **S. Choudhury**, S. Javdani, S. Srinivasa, and S. Scherer “Near-Optimal Edge Evaluation in Explicit Generalized Binomial Graphs” *Advances in Neural Information Processing Systems (NeurIPS)*. 2017. [\[paper\]](#)

PATENTS

- [23] J. A. Bagnell, A. Venkatraman, **S. Choudhury**, V. Narayanan. “systems and methods related to controlling autonomous vehicle (s).” U.S. Patent Application 18/269,209, issued September 12, 2024.
- [24] J. A. Bagnell, **S. Choudhury**, M. L. Phillips, A. Venkatraman, X. Yan. “Systems and methods for autonomous vehicle validation.” U.S. Patent 12,084,085, issued September 10, 2024.
- [25] **S. Choudhury**, S. Kumar, and M. Micol Marchetti-Bowick. “Goal-based Motion Forecasting.” U.S. Patent 11,801,871, issued October 31, 2023.
- [26] J. A. Bagnell, A. Venkatraman, and **S. Choudhury**. “Training machine learning model (s), in simulation, for use in controlling autonomous vehicle (s).” U.S. Patent 11,989,020, issued May 21, 2024.

TECHNICAL SKILLS

Languages: Python, C/C++, JavaScript, HTML/CSS **Libraries:** PyTorch, Transformers, JAX, Gym

TEACHING

Robot Learning, *Cornell University* Spring 23, Spring 24, Fall 24

- Designed a new undergraduate course on robot learning. Looks at fundamentals of MDP, IL, RL, Perception, and case studies of real-world applications from self-driving to manipulation to LLMs.
- Link to course: <https://www.cs.cornell.edu/courses/cs4756/2024sp/>

Learning for Robot Decision Making, *Cornell University* Fall 22, Fall 23

- Designed a new graduate course on learning for robot decision-making. Dives deep into online learning, imitation learning, reinforcement learning and explores frontiers of offline RL, LLMs and diffusion policies.
- Link to course: <https://www.cs.cornell.edu/courses/cs6756/2023fa/>

Imitation Learning: A Series of Deep Dives, *Youtube* 2022

- In this 10-part series, we dive deep into imitation learning, and build up a general framework. A journey through feedback, interventions and more!
- Link to playlist: https://youtube.com/playlist?list=PLQZQ7N26C6ba2BDFVULmmBYC80cX6pNjZ&si=UeEzsNL24kaC_mJj

Core Concepts in Robotics, *Youtube* 2022

- An introductory series that revisits core concepts in robotics in a contemporary light.
- Link to playlist: <https://youtu.be/GDmhrAHxgQE?list=PLQZQ7N26C6bbhXHuongGbbeVLI-PKkJ6K>

Interactive Online Learning: A Unified Algorithmic Framework, *Youtube* 2022

- In this series, we try to understand the fundamental fabric that ties all of robot learning – “How can a robot learn from online interactions?” Our goal is to build up a unified mathematical framework to solve recurring problems in reinforcement learning, imitation learning, model predictive control, and planning.
- Link to playlist: <https://youtu.be/u4HfY Cia50k?list=PLQZQ7N26C6ba-ifTijXwr2NXMkYbQe39O>

Mobile Robots, *University of Washington* Winter 2019

- Designed and instructed an undergraduate course on mobile robots. Students covered a range of topics in perception, planning and control as well as implemented algorithms on a fleet of 1/10th sized rally cars.
- Link to course: <https://courses.cs.washington.edu/courses/cse490r/19sp/>