Shared Autonomy and Teaching Tools for Flexible Human-Robot Teaming

Michael Hagenow
Massachusetts Institute of Technology
hagenow@csail.mit.edu

I. Introduction

As of 2025, only approximately *one in ten* manufacturing firms has adopted a robot [1]. This startling lack of adoption is often attributed to a lack of flexibility (e.g., reprogramming robots for new tasks) and technological limitations of current robot platforms [25]. The consequences of poor robot adoption are manyfold. With labor shortages, many manufacturers cannot recruit enough skilled human labor to meet production demands. Furthermore, even when workers are available, subjecting them to the many physically demanding tasks present in manufacturing (e.g., sanding, grinding, torquing fasteners) causes injuries [2]. To address these core adoption challenges, I develop *human-in-the-loop* approaches that can easily adapt to new tasks and reduce human workload, both physically and cognitively. Such approaches will serve a critical role in the future automation landscape for tasks that are underdefined (e.g., lacking data to train autonomous robots), critical, or require human collaboration with robots.

My research explores how to develop technologies and systems for flexible human-robot teaming. I draw on techniques in *controls, optimization, and generative artificial intelligence* and employ *human-centered design and physical prototyping* to design technologies that (1) simplify the process of transferring expert knowledge to robot teammates and (2) allow robots to work effectively alongside skilled humans. My past work has developed physical and algorithmic interfaces for humans to teach and collaborate with robots that have been deployed on realistic tasks (e.g., sanding) and evaluated in real domains (e.g., with manufacturing experts and at aviation-manufacturing facilities). My research builds toward a future of accelerated robot adoption through technologies that enable more general-purpose collaborative robots, which can aid in myriad tasks with little training and minimal input. To realize this vision, I contribute core methods in shared autonomy, end-user programming, and Learning from Demonstration (LfD).

II. Past Contributions: Toward Effective Shared Autonomy and Teaching Tools

I have developed foundational methods toward effective human-robot teaming, such as those illustrated in Figure 1. In particular, I have focused on contact-rich tasks that are difficult for people (e.g., high force) and difficult to fully automate. My past research investigated two key areas: (1) teaming for variable tasks through *informed shared autonomy* and (2) interfaces to *transfer human knowledge to robot behaviors*.

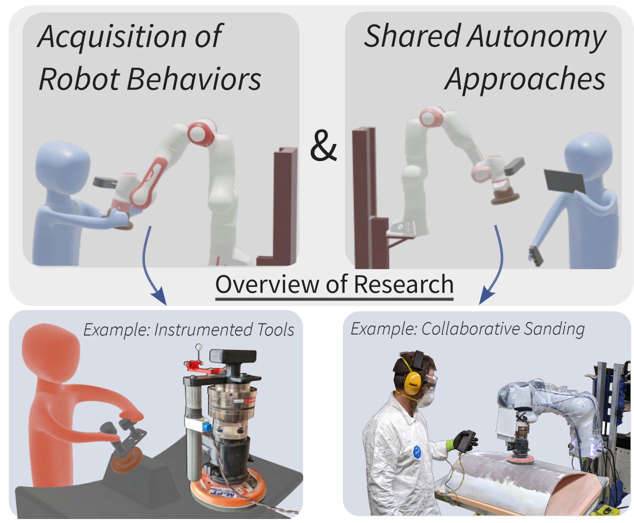

Fig. 1. I create tools and methods for humans to (1) teach robots new skills and (2) collaborate with robots through shared autonomy. My research develops approaches, algorithms, and systems for human-robot teaming. Example from past work include (*Bottom Left:*) instrumented tools for contact-rich LfD and (*Bottom Right:*) collaborative sanding in aviation manufacturing.

More Effective Teaming through Targeted Corrections – From a series of visits to manufacturers, I found that state-of-the-art shared autonomy approaches weren't addressing the needs of many manual tasks. Many popular shared control approaches required either too little (e.g., supervisory interfaces) or too much human input (e.g., constant input) for these semi-structured tasks [29, 22, 19, 20, 7, 10, 28]. To minimize human input while addressing task variability, I developed a shared autonomy approach, *corrective shared autonomy* [13], where skilled workers layer corrections on top of an existing robot behavior. The approach allows for human corrections to any controllable robot state and automatically proposes input mappings for low-dimensional operator corrections (e.g., one dimension to control combinations of force, pitch, and speed to modulate abrasiveness). The method uses expert task demonstrations, principal component analysis, and dynamical systems to automatically extract stable robot behaviors and operator interfaces for users to address likely sources of task error [14]. Through user studies, this approach achieved high usability ratings and enabled users to complete physical tasks (e.g., polishing and fastener insertion) when a simple robot policy lacked the robustness to complete the task [13, 17]. The approach also required less effort and time compared to shared control baselines. While this approach minimized operator

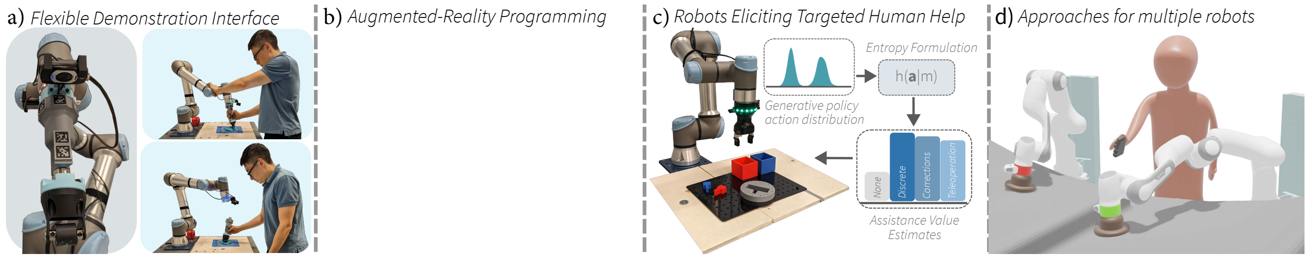

Fig. 2. Examples from past work toward effective human-robot teaming include developing (a) flexible demonstration interfaces (combining teleoperation, kinesthetic teaching, and natural demonstrations) that cater to different users and tasks, (b) augmented-reality user interfaces to program high-level task specifications, (c) approaches to estimate and elicit different types of human assistance for uncertain robots, (d) shared autonomy approaches that scale a skilled worker to sharing control with multiple coordinated robots in tasks with intermittent variability.

input, it wasn't necessarily using the expert human's time effectively (particularly when intervention was rarely needed). I proposed a solution that improves operator utilization by scaling and scheduling multiple executions around robot uncertainty (i.e., likely times of needed assistance). I developed a method to scale shared autonomy to the multi-agent setting where one human corrects multiple robots (see Figure 2d) that are sequenced around probabilistic confidence estimates [15]. Through a study with two sanding robots, our approach decreased task time *by 40 percent* (compared to serial shared autonomy) without significant performance impacts.

Eliciting Targeted Worker Assistance – An effective teammate knows not only *when to ask for help*, but specifically *what help is needed*. There are a range of levels of automation [4] and input mechanisms for a human to assist a robot, including real-time corrections (as introduced earlier), temporary teleoperation, and preference (e.g., discrete) input. However, few previous works have explored combinations of mechanisms [23, 18, 9], and in particular, whether a robot can elicit different levels of human feedback. I developed a method, *Real-Time Estimates of Assistance for Learned Models (REALM)*, that uses uncertainty from a generative AI policy (e.g., diffusion model) to select robot feedback requests that balance simplicity of human input with the informational needs of the robot [12]. I showed how post-intervention entropy estimates can serve as a proxy for robot uncertainty and assess when and how to request human assistance (see Figure 2c). Through a user study on an uncertain manipulation task, I showed how *REALM* can significantly reduce input (*38 percent*) and time on task (*14 percent*) compared to recent approaches that use robot confidence to alternate between teleoperation and autonomy [8, 6, 21].

Physical Interfaces for LfD – A key challenge in robot learning is transferring knowledge from an expert human to a robot, particularly for complex manipulation tasks. LfD has proven to be valuable approach, where a human demonstrates a task and the resulting demonstrations can be used to learn an imitation policy or reward model (through inverse reinforcement learning). However, selecting a demonstration interface that is both informative (for robots) and natural for people remains a challenge [5]. I developed, *Versatile Demonstration Interface (VDI)* [16], a robot tool that fuses demonstration modalities, as shown in Figure 2a. The tool attaches to the robot, interfaces with existing end-of-arm tooling, and allows easy switching between robot teleoperation, kinesthetic teaching, and natural demonstrations (where the tool can be removed and tracked by the robot). By allowing multiple demonstration types; VDI users can easily switch demonstration types based on the task, learning model, or preference of the user. We evaluated VDI in an exploratory user study with manufacturing experts (15 years of average experience) who completed representative manufacturing tasks, identified use cases for VDI, and confirmed the value of flexible demonstration types. For example, one participant stated: *" We have a broad range of right-out-of-school to thirty-years' experience [engineers]. People would pick different methods for sure."*

Research Agenda: Interfaces and Models for Flexible Human-Robot Teaming

My goal is to develop highly flexible robot teammates, that can effectively *learn from people* and are *self aware* to understand when they need help during teaming across a range of complex tasks. In the next several years, I plan to build toward this vision through three main research thrusts. **The first thrust investigates how to lower the barrier to teach robots new tasks**. To address current challenges in LfD [24] requires rethinking both the *physical* and *algorithmic* elements of teaching tools. I plan to develop new physical teaching interfaces, new approach combinations (e.g., natural interfaces and end-user programming [27, 26] – see Figure 2b.), and co-design of algorithms and interfaces to raise the complexity of tasks that end-users can teach to robots (e.g., more dexterous and force-rich tasks). **The second thrust will investigate new paradigms and multi-faceted models for effective assistance requests in high level-of-autonomy robots**. My previous work has explored one element: formalizations relating robot uncertainty and human input. However, by incorporating many other important factors; such as human uncertainty, preferences, and interface capabilities [11]; we can create more adaptive high level-of-autonomy robot assistance. **The final thrust investigates more general guidelines for teaming systems through human-centered design with new application areas**. My past work in aviation manufacturing illustrated how previous shared control methods fell short when faced with new domain challenges. By systematically assessing techniques (e.g., observations, co-design, and user studies) with domain experts [3, 30], I will lead development of design guidelines, taxonomies, and benchmarks that broaden the reach and generalization of teaming approaches.

References

[1] Ben Armstrong. The second robot problem: Obstacles to manufacturing automation at scale. HAMMER NSF Engineering Research Center Seminar, December 2023.

[2] Hamed Asadi, Denny Yu, and John H Mott. Risk factors for musculoskeletal injuries in airline maintenance, repair & overhaul. *International Journal of Industrial Ergonomics*, 70:107–115, 2019.

[3] Jan Auernhammer. Human-centered ai: The role of human-centered design research in the development of ai. 2020.

[4] Jenay M Beer, Arthur D Fisk, and Wendy A Rogers. Toward a framework for levels of robot autonomy in human-robot interaction. *Journal of human-robot interaction*, 3(2):74, 2014.

[5] Aude G Billard, Sylvain Calinon, and Rüdiger Dillmann. Learning from humans. *Springer handbook of robotics*, pages 1995–2014, 2016.

[6] Carlos Celemin and Jens Kober. Knowledge-and ambiguity-aware robot learning from corrective and evaluative feedback. *Neural Computing and Applications*, 35 (23):16821–16839, 2023.

[7] Marco Cognetti, Marco Aggravi, Claudio Pacchierotti, Paolo Salaris, and Paolo Robuffo Giordano. Perception-aware human-assisted navigation of mobile robots on persistent trajectories. *IEEE Robotics and Automation Letters*, 5(3):4711–4718, 2020.

[8] Shivin Dass, Karl Pertsch, Hejia Zhang, Youngwoon Lee, Joseph J Lim, and Stefanos Nikolaidis. Pato: Policy assisted teleoperation for scalable robot data collection. *arXiv preprint arXiv:2212.04708*, 2022.

[9] Tesca Fitzgerald, Pallavi Koppol, Patrick Callaghan, Russell Quinlan Jun Hei Wong, Reid Simmons, Oliver Kroemer, and Henny Admoni. Inquire: Interactive querying for user-aware informative reasoning. In *6th Annual Conference on Robot Learning*, 2022.

[10] Terrence Fong, Charles Thorpe, and Charles Baur. *Collaborative control: A robot-centric model for vehicle teleoperation*, volume 1. Carnegie Mellon University, The Robotics Institute Pittsburgh, 2001.

[11] Gaurav R Ghosal, Matthew Zurek, Daniel S Brown, and Anca D Dragan. The effect of modeling human rationality level on learning rewards from multiple feedback types. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, 2023.

[12] Michael Hagenow and Julie A Shah. Realm: Real-time estimates of assistance for learned models in human-robot interaction. *IEEE Robotics and Automation Letters*, 2025.

[13] Michael Hagenow, Emmanuel Senft, Robert Radwin, Michael Gleicher, Bilge Mutlu, and Michael Zinn. Corrective shared autonomy for addressing task variability. *IEEE Robotics and Automation Letters*, 6(2), 2021. doi: 10.1109/LRA.2021.3064500.

[14] Michael Hagenow, Emmanuel Senft, Robert Radwin, Michael Gleicher, Bilge Mutlu, and Michael Zinn. Informing real-time corrections in corrective shared autonomy through expert demonstrations. *IEEE Robotics and Automation Letters*, 2021.

[15] Michael Hagenow, Emmanuel Senft, Nitzan Orr, Robert Radwin, Michael Gleicher, Bilge Mutlu, Dylan P Losey, and Michael Zinn. Coordinated multi-robot shared autonomy based on scheduling and demonstrations. *IEEE Robotics and Automation Letters*, 2023.

[16] Michael Hagenow, Dimosthenis Kontogiorgos, Yanwei Wang, and Julie Shah. Versatile demonstration interface: Toward more flexible robot demonstration collection. *arXiv preprint arXiv:2410.19141*, 2024.

[17] Michael Hagenow, Emmanuel Senft, Robert Radwin, Michael Gleicher, Michael Zinn, and Bilge Mutlu. A system for human-robot teaming through end-user programming and shared autonomy. In *Proceedings of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*, pages 231–239, 2024.

[18] Borja Ibarz, Jan Leike, Tobias Pohlen, Geoffrey Irving, Shane Legg, and Dario Amodei. Reward learning from human preferences and demonstrations in atari. *Advances in neural information processing systems*, 31, 2018.

[19] Dylan P Losey and Marcia K O'Malley. Trajectory deformations from physical human–robot interaction. *IEEE Transactions on Robotics*, 34(1):126–138, 2017.

[20] Carlo Masone, Paolo Robuffo Giordano, Heinrich H Bülthoff, and Antonio Franchi. Semi-autonomous trajectory generation for mobile robots with integral haptic shared control. In *2014 IEEE International Conference on Robotics and Automation (ICRA)*, pages 6468–6475. IEEE, 2014.

[21] Kunal Menda, Katherine Driggs-Campbell, and Mykel J Kochenderfer. Ensembledagger: A bayesian approach to safe imitation learning. In *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 5041–5048. IEEE, 2019.

[22] Marcia K O'Malley, Abhishek Gupta, Matthew Gen, and Yanfang Li. Shared control in haptic systems for performance enhancement and training. 2006.

[23] Malayandi Palan, Gleb Shevchuk, Nicholas Charles Landolfi, and Dorsa Sadigh. Learning reward functions by integrating human demonstrations and preferences. In *Robotics: Science and Systems*, 2019.

[24] Harish Ravichandar, Athanasios S Polydoros, Sonia Chernova, and Aude Billard. Recent advances in robot learning from demonstration. *Annual review of control, robotics, and autonomous systems*, 3:297–330, 2020.

[25] Lindsay Sanneman, Christopher Fourie, Julie A Shah, et al. The state of industrial robotics: Emerging technologies, challenges, and key research directions. *Foundations and Trends® in Robotics*, 8(3):225–306, 2021.

[26] Emmanuel Senft, Michael Hagenow, Robert Radwin, Michael Zinn, Michael Gleicher, and Bilge Mutlu. Situated live programming for human-robot collaboration. In *The 34th Annual ACM Symposium on User Interface*

Software and Technology, pages 613–625, 2021.

[27] Emmanuel Senft, Michael Hagenow, Kevin Welsh, Robert Radwin, Michael Zinn, Michael Gleicher, and Bilge Mutlu. Task-level authoring for remote robot teleoperation. *Frontiers in Robotics and AI*, page 302, 2021.

[28] Thomas B Sheridan. Human supervisory control. *Handbook of human factors and ergonomics*, 2012.

[29] Keng Peng Tee and Yan Wu. Experimental evaluation of divisible human-robot shared control for teleoperation assistance. In *TENCON 2018-2018 IEEE Region 10 Conference*, pages 0182–0187. IEEE, 2018.

[30] Katie Winkle, Emmanuel Senft, and Séverin Lemaignan. Leador: A method for end-to-end participatory design of autonomous social robots. *Frontiers in Robotics and AI*, 8:704119, 2021.