Augmenting Human Control through Machine Learning

One of the outstanding challenges in robotics is building assistive systems that can help humans achieve their goals in a wide variety of tasks, including driving cars, flying drones, and controlling prosthetic limbs. These systems pose a different set of research problems than autonomous robots, in that they require better methods of modeling human preferences, beliefs, and behavior, in order to fulfill their mission of enabling humans to reach their full potential. Classical approaches to human-machine shared control involve extensive hand-engineering and domain expertise, making them brittle and expensive. Recent advances in machine learning – in particular, in deep reinforcement learning – have the potential to revolutionize shared autonomy through flexible and practical human-in-the-loop learning algorithms. Breakthroughs in computing and sensor hardware have only now made it possible to test these ideas in real-world systems, like driver assist for semi-autonomous cars. In my Ph.D., I hope to develop algorithms that endow machines with the ability to collaborate with humans on challenging sequential decision-making problems in robotic control. My goal is to accelerate the advent of intelligent robotic assistants, by advancing the state of the art in machine learning, robotics, and cognitive science.

Intent inference for human-robot shared autonomy. One of the problems I want to tackle is *shared autonomy*, which augments user control with automated assistance on tasks that are difficult to fully specify, and in which the user has private information about their intent that is difficult to directly communicate to a robot; for example, a human pilot remotely flying an aerial drone, using an onboard camera to track a herd of animals and identify injuries, while dealing with unfamiliar flight dynamics, terrain, and network latency. The key challenge in shared autonomy is inferring the user's intent, and helping them to achieve it.

Existing shared autonomy frameworks tend to require prior knowledge of the dynamics of the environment, a behavioral model of the user, and the set of candidate goals the user may target – assumptions that are often unrealistic in the real world. In a conference paper [3] published at RSS 2018, I proposed a model-free deep reinforcement learning method for shared autonomy that lifts these assumptions. The algorithm learns to implicitly decode the user's intent from their control input and act to achieve it, enabling human-machine centaur teams to outperform individual human pilots and autonomous robot pilots on a quadrotor teleoperation task. This approach could potentially be applied to other challenging domains where models of the physical system and user are difficult to obtain; for instance, prosthetic limb control via brain-computer interfaces – a setting in which our lack of understanding of the neural basis of motor control currently prevents us from designing decoders to translate high-dimensional neural signals into desired motion. In this setting, deep learning and reinforcement learning have the potential to help human patients who have lost limbs regain their independence and flourish.

One of the weaknesses of the model-free deep RL approach to shared autonomy is that it can require many interactions with the environment, which may not be feasible for a human user operating physical robots. I addressed this problem in a conference paper [4] published at NeurIPS 2018, in which I developed a model-based intent inference algorithm that can be used to predict a user's desired trajectory given their control inputs. The method fits a model of the user's internal beliefs about the dynamics of the system, then assists the user by transferring their control policy from the internal dynamics to the real dynamics. In experiments, I showed that this approach enables an assistive agent to help users play the Lunar Lander Atari game, even when the agent does not know the user's final objective. The proposed internal-to-real dynamics transfer algorithm for shared autonomy could potentially be helpful for applications with physical systems that can be accurately modeled and are well understood from an engineering perspective, but human users still struggle to operate – e.g., a robotic arm with many degrees of freedom.

In summary, I have taken preliminary steps [3, 4] toward better shared autonomy algorithms, but many challenges remain. In the following paragraphs, I discuss two of them: teaching robots to act autonomously, and teaching humans to behave more optimally.

Acquiring skills from human demonstrations or feedback. Enabling a robot to infer human intent is only helpful if the robot can actually perform the necessary maneuvers to achieve that intent. One approach to endowing robots with manipulation and locomotion skills is *imitation learning*, where an expert demonstrates near-optimal behavior to an agent that then attempts to replicate that behavior in novel situations. In a conference paper [5] published at ICLR 2020, I proposed an imitation learning algorithm called SQIL that uses off-policy reinforcement learning to train an agent to robustly imitate an expert in playing video games from high-dimensional image observations. This approach could be helpful for training autonomous robots to behave more like expert humans, e.g., automatically detecting hazards and navigating around them, in environments where the robot primarily relies on raw camera images for sensing. SQIL takes a step toward more autonomous robots, but is only applicable to imitation learning problems in which the agent acts in the same environment as the demonstrator. This is often not the case in the real world: the expert may provide demonstrations in a controlled environment with different transition dynamics and initial states than the open-world environment in which the agent is expected to replicate the expert behavior. In a conference paper [6] published at ICML 2020, I proposed a reward learning algorithm called ReQueST that addresses this problem by learning a model of the user's reward function that is robust to state distribution shift.

Teaching humans. Imitating expert behavior can be useful, but human users are not always experts. When the user makes systematic mistakes, the system should be able to identify those mistakes and help the user correct their own behavior. The key challenge here is designing an algorithm to automatically generate a personalized and adaptive curriculum for a human student. In a conference paper [7] published at KDD 2016, I developed a model of spaced repetition systems like Anki and Duolingo, drawing on ideas from queueing theory and psychology to optimize flashcard review scheduling. Insights from building intelligent assistants in educational software may be helpful for designing human-in-the-loop learning systems in robotics; for instance, in a driver assist program that is able to both enhance user control and also teach the user how to optimally execute certain maneuvers under unfamiliar hardware constraints or environmental conditions, like flat tires or icy roads. I plan to explore these ideas in future work.

Summary and future work. Learning to assist, imitate, and teach humans is a challenging and worthwhile pursuit, with applications in robotics, human-computer interaction, and education. In the long run, it will also help us expand our understanding of both artificial intelligence and human intelligence. I have taken preliminary steps toward my research goals [7, 3, 4, 5], but many open questions remain. In particular, I am interested in developing better methods for machine theory of mind [2] and neuroprosthetic decoder training [1].

References

[1] Josh Merel, David Carlson, Liam Paninski, and John P Cunningham. Neuroprosthetic decoder training as imitation learning. *PLoS computational biology*, 12(5):e1004948, 2016.

[2] Neil C Rabinowitz, Frank Perbet, H Francis Song, Chiyuan Zhang, SM Eslami, and Matthew Botvinick. Machine theory of mind. *arXiv preprint arXiv:1802.07740*, 2018.

[3] Siddharth Reddy, Anca D. Dragan, and Sergey Levine. Shared autonomy via deep reinforcement learning. *Robotics: Science and Systems*, 2018.

[4] Siddharth Reddy, Anca D. Dragan, and Sergey Levine. Where do you think you're going?: Inferring beliefs about dynamics from behavior. *Neural Information Processing Systems*, 2018.

[5] Siddharth Reddy, Anca D Dragan, and Sergey Levine. SQIL: Imitation learning via reinforcement learning with sparse rewards. *International Conference on Learning Representations*, 2020.

[6] Siddharth Reddy, Anca D Dragan, Sergey Levine, Shane Legg, and Jan Leike. Learning human objectives by evaluating hypothetical behavior. *International Conference on Machine Learning*, 2020.

[7] Siddharth Reddy, Igor Labutov, Siddhartha Banerjee, and Thorsten Joachims. Unbounded human learning: Optimal scheduling for spaced repetition. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 1815–1824. ACM, 2016.