

Information Extraction Under Communication Constraints with Assistive Robots

Brenna D. Argall

Department of Electrical Engineering and Computer Science
Northwestern University
Evanston, Illinois 60208
Rehabilitation Institute of Chicago
Chicago, Illinois 60611
Email: brenna.argall@northwestern.edu

I. INTRODUCTION AND MOTIVATION

As robotics technology becomes more ubiquitous and integrated into daily life, we increasingly will see instances of robot operation that shares control with human users and task partners. Robots in manufacturing will perform collaborative tasks with human partners, robotic exoskeletons will provide motor assistance to human workers and humans with motor impairments, robotic cars will take instruction from human riders, and partial autonomy will improve the teleoperation of complex robotic platforms and telepresence robots.

The target domain for this work considers a very particular and salient human-robot team: that of a human teammate with motor impairments working in collaboration with a robot that physically assists them. Such teams are characterized by heterogeneous teammates, extremely high cost of errors, a prioritization on human preference and limited communication channels to receive control signals from the human.

The estimation of user intent is fundamental to robotic systems operating in collaboration with, in close proximity to, sharing control with or assisting humans. Machine learning moreover can be leveraged to enact long-term—rather than simply reactive—changes in autonomy and control sharing.

This extended abstract considers the questions of human intent estimation, human-robot control sharing and adaptation within assistive robot domains with communication constraints. In particular, our work considers only interfaces used to teleoperate the robot, and where moreover the signals do not match the full robot control space. Under such constraints, models of characteristics of the human signals and the interaction between the human and robot are key. Opportunities for natural language to play a role within these models will be highlighted throughout the text.

II. DOMAIN AND FORMULATION

In addition to limited communication channels for human-robot teams within assistive domain, the physical capabilities of the human teammate are extremely non-static: hopefully due to successful rehabilitation, but also possibly due to the degenerative nature of a disease. A given user may even prefer different autonomy allocations on different days or throughout the day, depending on factors like level of pain or fatigue.

Within this domain, the ability to deal with communication bandwidth limitations and shifting autonomy is critical. Many human operators' physical impairments leave them with access only to very limited control interfaces—that perhaps even *preclude* the ability to operate the very assistive devices meant to aid them—and their needed or desired amount of assistance is often in flux. An example interface accessible to those with severe motor impairments controls wheelchair translational and rotational motion separately and in discrete increments—never simultaneously, never with continuous control signals. Figure 1 shows two assistive robot platforms currently under development in our lab.

Our technical formulation assumes the existence of a set \mathcal{F} of automated controllers $f(\cdot)$, and a set \mathcal{B} of control sharing strategies β . Each controller dictates the motion of an autonomous behavior for the assistive robot. Formally, a vector of control signals \mathbf{u}_f^t is generated from state \mathbf{x}^t by function $f(\cdot)$, $\mathbf{u}_f^t \leftarrow f(\mathbf{x}^t)$. Control vector \mathbf{u}_h^t is computed by mapping control signals from the human to the space of robot control.¹ These two signals then are reasoned about within control sharing strategy $\beta(\cdot)$ which generates the signal \mathbf{u}^t executed by the robot platform, $\mathbf{u}^t \leftarrow \beta(\mathbf{u}_f^t, \mathbf{u}_h^t)$.

We furthermore model the frequency of control (communication) signals coming from the human with bandwidth parameter Ω , a running estimate updated online. Changes in the communication bandwidth in turn may trigger a change in the amount of autonomy assumed by the robot.

¹Under direct teleoperation, this mapping is the identity function. With a natural language interface, the human signal would likely be higher-level (e.g. “turn right”) and require a more complex mapping.

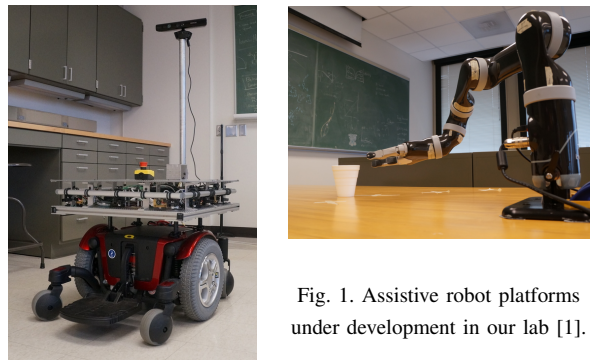


Fig. 1. Assistive robot platforms under development in our lab [1].

III. ESTIMATION OF USER INTENT

To appropriately share control with a human, a robot requires a notion of the human’s intentions or goals—which must be estimated, if they are unknown. The question we aim to address is *how to estimate human intent* from (1) only the control signals used to teleoperate the robot, which moreover (2) are constrained by limitations of the control interface?

The motivation for constraining the problem in this way is two-fold. (1) For common machines like powered wheelchairs, the interfaces used to generate control signals for teleoperation are currently available, broadly employed and their operation is familiar to users. (And *feasible* for users—which is important for assistive domains.) (2) These interfaces do not divert user attention from the task execution. We therefore aim to push the limits of what can be inferred from this constrained set of control signals first, and only if necessary afterwards consider the introduction of additional/augmented interfaces through which an explicit indication by the user is possible.²

Our aim thus is to develop algorithmic approaches to estimate user intent that explicitly encode crucial aspects of the task domain—like control interface limitations—while expecting no additional information from the user other than the control signals used to teleoperate the robot. One such approach is to explicitly model *constraints on the source* of the control signal coming from the user. Specifically, within a path planning environment we model constraints on the formulation of states $s \in S$ and actions $a \in A$, and allowable state transitions $T_{s,a}^{s'} : s \xrightarrow{a} s'$, as imposed by limitations on the control interface. Additionally, we encode within the planner’s cost function $c \leftarrow C(s, a)$ difficulties encountered when operating the interface. To reason about user intent, such an approach then may consider the space of viable plans generated by the constrained planner and to what extent they agree with the forward projection of the user’s control signal.

IV. ADAPTATION

Our motivation in extracting learning cues from the human’s control signals is again in consideration of limited bandwidth—and to extract as much information as is possible from the human teammate signals. We aim to adapt both the formulation of the autonomy behavior and the manner of control sharing. The latter amounts to optimizing the model for how the human and robot interact. The former implicitly models the human and/or their preference, by adapting the behavior to match the human’s control signals or in response to feedback the human provides.

We formulate each behavior $f \in \mathcal{F}$ so as to have an associated set of parameters θ_f which are available for modulation. For example, a path planner [3] used on our mobile robot platform has parameters to modulate how much curvature is in the generated trajectory, and how aggressively the robot attempts to reach the goal position. Exactly what influence the parameters θ_f have on associated behavior f varies greatly

across behaviors. However, the approaches used to *modulate* the parameters can be common across behaviors. Any number of machine learning algorithms may be used to perform this modulation [2, 4]. The key factor to consider is the *feedback signal* received by a machine learning algorithm. Candidate signals include (1) a reward/cost (for use within a Reinforcement Learning (RL) algorithm) or (2) an example/correction (for use within a Supervised Learning algorithm).

For the signal source, one option has the human explicitly provide feedback about their preference or the robot performance. Such a signal might be provided through the control interface, and take the form of a reward/cost or a correction.³

Another option is to infer the feedback signal from the human’s control signals. We consider interpretations that rely on an assessment of *agreement* between signals produced by the human versus the autonomy, and which explicitly considers latency in the communication channel by evaluating a window of prior autonomy decisions. One then might make an optimistic comparison, to the autonomy decision *most similar* to the human signal, or take the *most prevalent* (dominant) autonomy decision over that window, for example. A running measure of agreement $\lambda \in [0, 1]$ is computed, while $\tilde{\lambda}_h$ and $\tilde{\lambda}_\ell$ are respectively upper and lower thresholds on (dis)agreement.

Reinforcement. In the case of *agreement* between the human signal and the autonomy, the learning response is to *reinforce* the autonomy selection. We quantify agreement as $\tilde{\lambda}_h < \lambda$. Reinforcement can be accomplished by generating a positive reward within an RL formulation, for example.

Correction. In the case of *disagreement* between the human signal and the autonomy, the human signal is interpreted as a *correction* for the autonomy. This could be accomplished by treating the execution trace as a demonstration within a Learning from Demonstration paradigm, for example. Disagreement is similarly quantified, according to $\lambda < \tilde{\lambda}_\ell$. While in theory it is possible to treat any individual datapoint discrepancy as a single datapoint correction, it is expected that leveraging instead the aggregate measure λ will prove more stable and robust—especially in the case of bandwidth latency.

Penalization. In addition to correction, we also consider *penalization* within a RL formulation, by generating a negative reward in the case of *disagreement*.

V. CONCLUSION

The domain of robots that provide physical assistance to humans has numerous compelling factors motivating the adaptation of formulations for control sharing and robot autonomy, and to infer human intent. We have proposed mechanisms that rely on models of constraints on the communication of control signals from the human, and extracting as much information as possible from what is communicated—in order to optimize the robot’s autonomy and control-sharing paradigm to match the human’s expectations. Ongoing and future work in our lab explores these ideas.

²A natural language interface would provide a very direct way for the human to indicate high-level goals, alleviating the need to estimate intent.

³Natural language in this case could be quite a straightforward interface for the human to provide both reward-based and corrective feedback.

REFERENCES

- [1] Brenna D. Argall. Turning assistive machines into assistive robots. In *Proceedings of the SPIE Photonics West Congress*, 2015. Keynote presentation.
- [2] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer, New York, NY, USA, 2001.
- [3] Jong Jin Park, Collin Johnson, and Benjamin Kuipers. Robot navigation with model predictive equilibrium point control. In *Proceedings of IROS*, 2012.
- [4] Richard S. Sutton and Andrew G. Barto. *Reinforcement Learning: An Introduction*. The MIT Press, Cambridge, MA, London, England, 1998.