Perspective Taking for Shared Control

Connor Brooks[1] and Daniel Szafir[2]

I. Overview

In teleoperation, humans produce action commands for robotic control that vary in optimality due to factors such as operator skill or understanding of the robotic system [1]. However, if the user has a poor understanding of the state the robot is in (e.g. due to control interface limitations), even particularly skilled operators may struggle to produce optimal actions. We propose a shared control system for mitigating suboptimality in user actions caused by improper understanding of the environment state. To do this, we use *perspective taking* to take into account the user's knowledge of the environment when interpreting their control signals.

Shared control systems that assist a human user with an arbitrary goal often learn about this goal from the actions the user takes [2]. In doing so, the user is often treated as optimal — or at least noisily-optimal — with respect to choosing actions at each state [3]. However, this assumption may not always hold. The user may have some systematic bias due to a misunderstanding about the robot's current state, the system dynamics, or the goal itself. In cases in which the user has a systematic bias, modeling this bias can enable shared control systems to assist users by correcting for this bias before operational commands are sent to the robot.

We are interested in fixing bias which arises due to an incorrect user estimate of the environment state. Intuitively, the systematic bias that we wish to model is due to the limitations of the user's *perspective* in remote teleoperation. In remote teleoperation, the user's knowledge of the environment is limited according to the data presented to them by the control interface (e.g. a single camera's field of view). Additionally, users often have a limited sense of the robot's internal joint limits and the types of singularities which prohibit it from arbitrary motion at certain states in the configuration space. In either case, the user selects an action according to an incomplete or erroneous estimate of the current environment's true state. This has the potential to introduce suboptimality in the control actions the user issues to the robot. Through *perspective taking*, we model the user's actions as optimal only in the environment state the user had knowledge of when issuing the command.

II. Background

Herman et al. [4] demonstrate the concept of modeling a user's systematic bias by creating a framework for estimating an agent's internal model of system dynamics according to their actions. Reddy et al. [5] apply a similar approach to shared control by performing this estimation, then altering the commands the user sends to a robot according to the mapping from this internal dynamics model to the true system dynamics. Both of these methods require first collecting a round of data in which the user interacts with the system, then performing an offline training step in which the parameterization of the user's internal model is estimated. Finally, the internal model is applied to the user's future interactions with the system while assuming that the user's internal model remains consistent. While these existing works focus on a different source of systematic bias than we do, they demonstrate that correcting a systematic bias can lead to improved performance in shared control.

Outside the context of shared control, the strategy of estimating a human's knowledge about environment or task state has been shown to be useful for human-robot collaboration in contexts such as disambiguating human requests [6] or improving robot learning from a human teacher [7]. Trafton et al. [8] use a cognitive architecture to simulate a human's perspective when responding to requests for help from a robot assistant. Hiatt et al. [9] extend this approach by simulating multiple possible human perspectives and determining which is most likely according to the human's observed actions. Likewise, Devin et al. [10] model a human teammate's knowledge about a task's completion status from their actions to improve team communication. These prior works demonstrate that perspective taking can improve understanding of human actions for various types of human-robot interactions. Our proposal utilizes this idea to mitigate systematic bias introduced through teleoperation interfaces.

III. Formalism

Formally, we model the user as acting within a Markov Decision Process (MDP) defined by the tuple (S, A, T, R). We introduce a function representing the user's perspective that maps the true state to the user's understanding of the state $\phi : S \to S$. This mapping represents the systematic error in the user's knowledge of the state. For example, in a control interface that relies on a mounted camera, a state s may contain obstacles outside the camera's field of view while $\phi(s)$ contains only the obstacles in view of the camera.

A. Myopic Correction

Correcting the limitations in the user's perspective involves inferring what action a_t^* the user would have chosen in the environment state s_t given their observed action a_t^u in their erroneous state $\phi(s_t)$. A straightforward way of applying

[1]Connor Brooks is with the Department of Computer Science, University of Colorado Boulder connor.brooks@colorado.edu

[2]Daniel Szafir is with the ATLAS Institute and the Department of Computer Science, University of Colorado Boulder daniel.szafir@colorado.edu

this correction is to determine what local state the user is attempting to reach. This state, s^u_{t+1}, is the immediate goal of the user's chosen action:

$$s^u_{t+1} = T(\phi(s_t), a^u_t) \tag{1}$$

However, this state may not be reachable in the actual environment; instead, it is a reachable state within the user's perspective-mapped set of states reached by ϕ. Consequently, to model the true environment state they were attempting to reach, we need to reverse the bias mapping:

$$s^*_{t+1} = \phi^{-1}(s^u_{t+1}) \tag{2}$$

Then, we choose the action that minimizes some distance metric d_1 to this desired state:

$$a^*_t = \arg\min_{a \in A} d_1(T(s_t, a), \phi^{-1}(T(\phi(s_t), a^u_t))) \tag{3}$$

B. Predictive Correction

This myopic approach performs a local correction, but it does not take advantage of knowledge that the user is acting suboptimally in a systematic manner. Alternatively, we propose utilizing knowledge of ϕ to prevent situations in which a^*_t is drastically different from a^u_t. That is, by looking ahead to states in which the user's selected action and the closest action for achieving their intended state will differ greatly, we can make minor adjustments to the robot's movement proactively to avoid a potentially jarring mismatch between user input and robot action.

In order to do this, we look ahead over a finite time horizon of length n to possible future trajectories. A trajectory ξ consists of state action pairs from each timestep t to $t+n$:

$$\xi_{t \to n} = \{(s_t, a_t), ..., (s_{t+n}, a_{t+n})\} \tag{4}$$

We first generate a user-perspective expected trajectory, $\xi^u_{t \to n}$. In generating this trajectory we assume knowledge of the action the user would choose at each state. In implementation, this action selection could be estimated by a separate module that performs inference over user goal or simply repeatedly applying the user's last observed action.

$$\xi^u_t = (\phi(s_t), a^u_t) \tag{5}$$

$$\xi^u_{t+i} = (T(s_{t+i-1}, a^u_{t+i-1}), a^u_{t+i}) \quad \text{for } 1 \leq i \leq n \tag{6}$$

We create an objective function for a candidate trajectory that takes into account the difference in expectation between the user-perspective expected trajectory and the candidate trajectory. This objective function utilizes two distance metrics: d_1 is a distance metric over states and d_2 is a distance metric over actions. We use temporal discounting to emphasize the states closest to the current timestep.

$$h(\xi_{t \to n}, \xi^u_{t \to n}) = \sum_{i=t}^{n} \lambda^{i-t}(\alpha d_1(s_i, \phi^{-1}(s^u_i)) + \beta d_2(a_i, a^u_i)) \tag{7}$$

The tuning parameters α and β weight the importance given to matching user expected state and user chosen action, respectively. At each timestep, we select the first action from the trajectory that minimizes the objective function.

$$\xi^*_{t \to n} = \arg\min_{\xi} h(\xi_{t \to n}, \xi^u_{t \to n}) \tag{8}$$

$$a^*_t \sim \xi^*_t \tag{9}$$

This approach spreads out a needed correction over several timesteps, minimizing the occurrence of large corrections that drastically alter user controls.

IV. Future Work

Our continued work is primarily focused on developing methods to efficiently perform the minimization in Eq. 8. Additionally, we initially assume perfect knowledge of ϕ for our proposed approach to mimic situations in which the user's misunderstanding of the state is due entirely to the perspective of the control interface. In future work, we will also investigate methods for relaxing this assumption by estimating ϕ from observed user commands.

Through applying perspective taking to shared control, we propose fixing a systematic bias that causes users to issue suboptimal commands. We believe this approach is particularly appropriate to the context of remote teleoperation due to the limited environment awareness provided to a user through control interfaces. In general, perspective taking for shared control allows correction of errors in user understanding of the environment while retaining user authority in determining the desired path through the environment.

References

[1] M. Fdez-Carmona, C. Urdiales, and F. Sandoval, "Reactive adapted assistance for wheelchair navigation based on a standard skill profile," in *2014 IEEE International Conference on Robotics and Biomimetics (ROBIO 2014)*. IEEE, 2014, pp. 2152–2157.

[2] C. Brooks and D. Szafir, "Balanced information gathering and goal-oriented actions in shared autonomy," in *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 2019, pp. 85–94.

[3] A. D. Dragan and S. S. Srinivasa, *Formalizing assistive teleoperation*. MIT Press, July, 2012.

[4] M. Herman, T. Gindele, J. Wagner, F. Schmitt, and W. Burgard, "Inverse reinforcement learning with simultaneous estimation of rewards and dynamics," in *Artificial Intelligence and Statistics*, 2016, pp. 102–110.

[5] S. Reddy, A. D. Dragan, and S. Levine, "Where do you think you're going?: Inferring beliefs about dynamics from behavior," *arXiv preprint arXiv:1805.08010*, 2018.

[6] R. Ros, E. A. Sisbot, R. Alami, J. Steinwender, K. Hamann, and F. Warneken, "Solving ambiguities with perspective taking," in *2010 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 2010, pp. 181–182.

[7] C. Breazeal, M. Berlin, A. Brooks, J. Gray, and A. L. Thomaz, "Using perspective taking to learn from ambiguous demonstrations," *Robotics and autonomous systems*, vol. 54, no. 5, pp. 385–393, 2006.

[8] J. G. Trafton, N. L. Cassimatis, M. D. Bugajska, D. P. Brock, F. E. Mintz, and A. C. Schultz, "Enabling effective human-robot interaction using perspective-taking in robots," *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, vol. 35, no. 4, pp. 460–470, 2005.

[9] L. M. Hiatt, A. M. Harrison, and J. G. Trafton, "Accommodating human variability in human-robot teams through theory of mind," in *IJCAI Proceedings-International Joint Conference on Artificial Intelligence*, vol. 22, no. 3, 2011, p. 2066.

[10] S. Devin and R. Alami, "An implemented theory of mind to improve human-robot shared plans execution," in *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 2016, pp. 319–326.