

Legible User Input for Intent Prediction

Kenton C. T. Lee
University of Pennsylvania

Anca D. Dragan
Carnegie Mellon University

Siddhartha S. Srinivasa
Carnegie Mellon University

Abstract—In assistive teleoperation, the robot provides assistance by predicting the user’s intent. Prior work has focused on improving prediction by adapting it to the user’s behavior. In this work, we investigate adaptation in the *opposite* direction: training the user’s behavior to the prediction. Results from our user study suggest that users can significantly improve the performance of a simple static predictor after brief exposure to its behavior. In addition, we find this improvement to be more significant when the cognitive load of teleoperation is reduced.¹

Keywords—*expressing intent; prediction; user adaptation*

I. INTRODUCTION

We view intent prediction as a collaborative activity between the actor and the observer. While the observer (here, a robot) tries its best to predict the actor’s intent (here, a user of the robot), the user can also attempt to make his or her intent more legible to the robot. Although much work has focused on the robot’s predictor [1], [7], the quality of the prediction is inherently limited by ambiguity in the user’s behavior. In this work, we investigate the idea of training user behavior to be less ambiguous (more legible or intent-expressive) to the predictor.

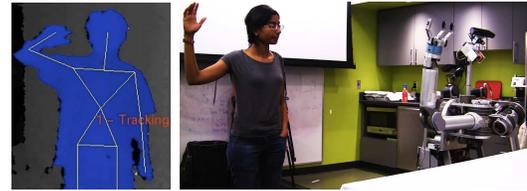
We focus on the assistive teleoperation domain, where control is semi-supervisory. The user provides input to the robot, and the robot blends this input with its own plan to execute the user’s intent [2]. It is this tight coupling between user and predictor that provides incentive to the user to provide legible input. This idea has strong parallels in the IUI literature, in the observation that users can alter their actions when they want their plan to be recognized [6]. In search systems too, users are able to adapt to degraded search systems in order to improve the quality of information retrieval [4].

User adaptation, however, can be harmful if unguided. For example, hyperarticulation in speech, which helps a human listener, can lead to lower accuracy from a speech recognition system [3]. This stresses the importance of guiding adaptation with feedback from the predictor: users must provide input legible *to the robot* rather than to another human.

In addition, since user adaptation is a learning process, cognitive load can hinder its effectiveness. Attention given to teleoperating the robot takes away from the capacity to understand the predictor. For our user study, we test this intuition by modulating the ease of control.

Our results suggest that users can significantly improve the performance of a simple static predictor after brief exposure to its behavior.

¹Parts of this work appear in a paper conditionally accepted to JHRI.



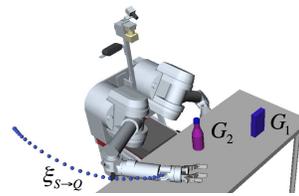
(a) Assistive teleoperation with the Kinect



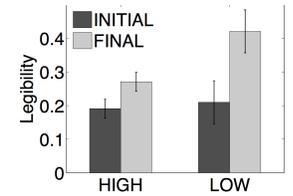
(b) Training interface



(c) Low cognitive load task



(d) Goal prediction



(e) Legibility improvement

Fig. 1: System overview and study results

II. EXPERIMENTAL DESIGN

Hypotheses. We investigate two hypotheses:

H1: *Users can improve the legibility of their input when given real-time feedback on prediction correctness and confidence.*

H2: *Improvement of legibility is greater for teleoperation requiring lower cognitive load.*

Dependent Measure. In order to quantify legibility of user input, we define a metric capturing how well the robot can make predictions based on the input. Given the trajectory of user inputs so far, $\xi_{S \rightarrow Q}$, the robot computes a confidence for each possible goal G , $C(\xi_{S \rightarrow Q}, G)$, and predicts the goal with the highest confidence (Fig.1(d)). A user’s input trajectory $\xi_{S \rightarrow G^*}$ is legible if the robot can predict the intended goal G^* with high confidence at every point Q along the trajectory. Therefore, we define the metric of legibility to be the mean confidence in the goal G^* along the trajectory:

$$M = \frac{1}{Z} \sum_{Q \in \xi_{S \rightarrow G^*}} C(\xi_{S \rightarrow Q}, G^*)$$

with Z being the number of points along the trajectory.

We implement the prediction algorithm proposed by [2] and use its scaled probability estimates as confidence. In our study we use the sum of squared velocity magnitudes as the cost of a trajectory.

Methods. We use a teleoperation interface that allows users to control HERB [5], a robot with 7DOF arms, through a whole-body interface that tracks the user's skeleton (OpenNI, www.openni.org), yielding an arm configuration which serves as the user input (Fig.1(a)).

In order to examine H1, we allowed users to first practice teleoperating HERB by moving towards target locations. After becoming familiar with the interface, users went through an *initial* testing phase, during which we recorded the legibility of their input as they performed a pre-defined set of tasks involving maneuvering HERB's hand towards one of two goal locations. In order to obtain unbiased estimates of the user's trajectories, we did not inject any assistance into the system.

In the training phase, users performed a different set of similar tasks, now with a visual indicator displaying the robot's confidence projected around the predicted goal location (Fig.1(b)). During this phase, the experimenters encouraged the users to find ways to improve this confidence.

Lastly, the prediction indicators were removed, and the users went through a *final* testing phase, in order to test whether training improved input legibility. This phase was simply a replication of the *initial* testing phase.

Participants of the above experiment belonged to the *high* cognitive load group. In order to test H2, we repeated the experiment with an interface requiring reduced cognitive load. The *low* cognitive load group teleoperated a virtual robot with kinematics that matched the user, drastically increasing the ease of control (Fig.1(c)).

Subject Allocation. We recruited 8 participants, 6 female and 2 male. Their ages ranged from 20-33, and none of them had prior experience with robots. The participants were equally and randomly partitioned into the two between-subject groups.

III. RESULTS

Improvement in legibility To test H1, we performed a repeated measures ANOVA to determine whether the training factor was significant. We included two repeated measures, the subject and the task. In order to control for practice effects, the time taken to complete the task was included as a covariate. We found that participants moved significantly more legibly during the *final* condition ($F(1, 22) = 11.71, p = .002$), while the time taken to complete the task was not a significant effect ($F(1, 22) = 1.11, p = .303$). These results suggest that exposure to the intent-predictor allowed users to improve the predictor's confidence by producing more legible motion.

Effect of cognitive load To test H2, we used the *improvement* in legibility as a dependent measure, and performed a repeated measures ANOVA to determine the effect of the cognitive load factor on this measure. We found that the improvement was significantly greater for users in the *low* group ($F(1, 46) = 5.18, p = .027$) (Fig.1(e)).

IV. DISCUSSION, LIMITATIONS, AND FUTURE WORK

Our study shows that users significantly improved their input after a training phase, during which users have access to feedback from the predictor. However, it is still unclear whether users are learning to become legible to the predictor, or they are simply producing motion that would be legible to other humans. If a robot's predictor is intuitive and its predictions are similar to a human's predictions, becoming legible to it may require little or no training. On the other hand, learning to be legible to unintuitive predictors may not be possible at all. Future work must formalize the notion of intuitiveness for prediction algorithms and investigate its effect on input legibility improvement.

We also examined the effect of cognitive load on legibility improvement. Without the cognitive load of teleoperating a real robot with dissimilar kinematics, users achieved a greater improvement in legibility. This effect suggests that it may not be easy for users to naturally adapt to a predictor in a practical setting, where cognitive load is normally high. Instead, a dedicated training session in a simplified setting, where users pay full attention to the predictor, would be more effective.

Though the improvement was significant overall, some subjects were outliers. One subject employed a basic memorization strategy of what type of input was best in what situation, rather than attempting to generalize the mapping found during the training phase. This resulted in poor performance during testing, highlighting the importance of intuitive predictors and appropriate training as opposed to natural adaptation. Another subject explained that she found correct predictions to be "good enough" for the task. Since assistance was removed in the study, the utility of confident predictions may not have been apparent to the participants. For future work it would be worthwhile to understand how the robot's assistance interacts with this user adaptation, and whether explicit feedback is really needed.

Overall, our study took a small step towards a better understanding of this under-exploited resource of improving learners by improving their user's behaviors. We look forward to more in-depth analysis of the factors that affect this improvement.

REFERENCES

- [1] B.D. Argall, S. Chernova, M. Veloso, and B. Browning. A survey of robot learning from demonstration. *Robotics and Autonomous Systems*, 57(5):469 – 483, 2009.
- [2] A.D. Dragan and S.S. Srinivasa. Formalizing assistive teleoperation. *R: SS*, 2012.
- [3] Sharon Oviatt, Margaret MacEachern, and Gina-Anne Levow. Predicting hyperarticulate speech during human-computer error resolution. *Speech Communication*, 24(2):87 – 110, 1998.
- [4] Catherine L. Smith and Paul B. Kantor. User adaptation: Good results from poor systems. In *LAIR Technical Report*, 2008.
- [5] S.S. Srinivasa, D. Berenson, M. Cakmak, A. Collet, M.R. Dogar, A.D. Dragan, R.A. Knepper, T. Niemueller, K. Strabala, M. Vande Weghe, and J. Ziegler. Herb 2.0: Lessons learned from developing a mobile manipulator for the home. *Proceedings of the IEEE*, 2012.
- [6] Annika Waern. Recognising human plans: Issues for plan recognition in human-computer interaction, 1996.
- [7] B.D. Ziebart, N. Ratliff, G. Gallagher, C. Mertz, K. Peterson, J.A. Bagnell, M. Hebert, A.K. Dey, and S. Srinivasa. Planning-based prediction for pedestrians. In *IROS*, 2009.