Information-Theoretic Characterization of Control Modes for Intent Disambiguation - Response to Reviewers

IEEE Robotics And Automation Letters (RA-L)

May 30, 2019

Deepak E. Gopinath[1, 3], Brenna D. Argall[1, 2, 3],
deepak.gopinath@u.northwestern.edu

[1]Department of Mechanical Engineering, Northwestern University, Evanston, IL
[2] Departments of Computer Science and Physical Medicine and Rehabilitation, Northwestern University, Evanston, IL
[3] Shirley Ryan AbilityLab, Chicago, IL

We sincerely thank the reviewers for their valuable efforts and feedback. We have made every effort to address all the major points raised by the reviewers and have modified the manuscript accordingly. For added ease of reading, we are also submitting a version of the revised manuscript in which the modifications are highlighted in orange color. Detailed responses to the comments follow.

1 Reviewer 3

Comment 1

- I think it is better to define certain terms such as "belief over goals" and "intention disambiguation/inference" in the introduction.

We thank the reviewer for this comment and have modified the text to make the meaning of intent disambiguation and inference more clear to the reader.

Comment 2

As I understand Algorithm 1 explains how intentions are obtained,however it does not explain how this intentions are embedded in the robot assistive controller.

If we are correctly interpreting the reviewer's comment, we believe there is a point of misinterpretation here: Algorithm 1 in fact is about computing the disambiguation mode, which is different from predicting the intent itself (if that is what the reviewer meant by "how intentions are obtained"). This distinction is indeed a potential point of confusion, and we have included a new figure (Figure 1) that explicitly calls out these various components as well as clarifying text throughout the document.

Comment 3

Simulation methods are not described with sufficient details. It is hard to reproduce this. Could you add equations for human and robot control behaviour (u_h and u_r)? Adding detailed equations and parameters to Appendix and supplementary material will be helpful.

Tables I and II now provide all simulation parameters and their ranges. The specification of u_h is given in Equation 7, and u_r in Section V.A.1. (RA-L does not allow Appendix and supplementary material, we therefore were unable to submit those.)

Comment 4

- I am not sure if it is valid to do statistical analysis of simulated data?

We would like to respectfully argue that statistical analyses of simulated data is indeed valid (and happens often within the field of robotics). The goal with our simulation was to be able to control for various confounding factors that can potentially arise in real human subjects studies (such as fatigue, lack of attention etc.) and perform a more extensive evaluation (8000 trials for each metric space) than could ever be possible on real hardware and with real subjects. We do not believe that this replaces hardware evaluation, but rather that it complements it. In this work we chose to do the more exhaustive simulation evaluation first.

2 Reviewer 4

Comment 1

One unclear issue is the presence of Dynamic Neural Field (DNF) approaches. Because the KL and Entropy algorithm utilizes the intent inference algorithms of heuristic, Bayesian, and dynamic neural field approaches. The other two algorithms were applied in this paper, but the usage of the DNS approach is not explicitly written.

We apologize for this confusion. For the simulations we indeed utilized all three approaches (heuristic, Bayesian and DNF). Table I, Row 3 indicates the different intent inference approaches utilized in the simulations.

Comment 2

, most of terms are salient but I suggest changing one of those terms to improve the readability.

We thank the reviewer for this valuable suggestion and have modified the terms used for describing the different metrics accordingly for improved readability throughout the paper as well as in the figures. 'Total Amount of Assistance' has been changed to 'Temporal Ratio of Assistance'.

Comment 3

unclear selection of test condition in the second experiment... KL and ENT should be compared to GRD or ENT should be compared to GRD at least because it showed better performance...

We agree with the reviewer that a full comprehensive experiment would need to compare both ENT and KLD against GRD condition. As the reviewer rightly spotted, this paper is an extension to our earlier paper [21] which just relied on hand-crafted heuristic features to perform intent disambiguation. Our goal was to identify at least one information-theoretic approach that was better than the heuristic approach (if one existed). That KLD-based intent disambiguation system was found to be more effective than the GRD condition validated the use of information theoretic approaches for intent disambiguation. Having validated the promise of the information-theorectic approaches, we are now in the process of preparing a full-fledged human subject study in which one of the conditions indeed will be to test both ENT and KLD against the GRD.

3 Reviewer 9

Comment 1

It seems that the approach is specialised for a specific task, i.e. for a reaching task where goals are discrete. The authors claim that this type of task is the most common which seems to be an overstatement.

We sincerely believe that this is not an overstatement, given the capabilities of today's assistive robotic arms. In the domain of assistive robotic manipulations, robotic arms are typically used by motor-impaired subjects for performance activities of daily living (ADL) tasks. Various types of ADL tasks such feeding, picking up objects from shelves, tabletops and the ground, reaching for doorknobs all involve reaching toward objects in the world. We have added text to the introduction where we clarify this point. Numerous videos of use-cases (available online) in which motor impaired people use robotic arms in their daily lives will further illustrate our point.

Comment 2

There are many tasks in robot assistance or HRI that may benefit from assistance, e.g. obstacle avoidance, trajectory tracking, force tracking, etc. The authors are encouraged to discuss how does the proposed approach generalise to these tasks.

We thank the reviewer for this suggestion and accordingly we have added some text in the discussion section addressing how generalizable our algorithm is. We make further reference to the generalizability in the introduction section where we state that intent inference is crucial for the success for many shared-autonomy human machine systems. Any human-robot system that has an intent inference mechanism under-the-hood can potentially utilize our intent disambiguation algorithm with respect to relevant parameters that affect intent inference (control modes, in our specific example) to enhance the efficacy of the inference mechanism.

Comment 3

Related work rather lists some scarce examples instead of being a structured summation of the most important works related to this paper. There are many approaches for intent inference found in the literature, see the works of e.g. Anca Dragan. Also, the related work should be linked to the present work in a meaningful way.

We thank the reviewer for this suggestion and have accordingly made our related work section more comprehensive. We hope this modification will help the reader to make the connection between past work and the work presented in the paper more clear as well.

Comment 4

It is unnecessary to devote the whole section to mathematical notation. It should only be a paragraph in the introduction. It is highly advisable to rather give a problem setting, with accompanying figure of the task considered where the notation is usefully presented

We have taken this suggestion into account and have accordingly deleted the section devoted to mathematical notation. The content has been integrated with the new Section IV.

Comment 5

Please list all the assumptions you impose on the problem in one place, e.g. that the human doesn't change the goal, that at any given time only one control mode can be activated, etc. Mentioning assumptions "on the fly" is not suitable.

We thank the reviewer for the comment, but we have respectfully chosen not to change the way we present the math. We believe that it is important to start with the most general form of the equations and to introduce simplifications based on different assumptions as the problem specification becomes less general. In our experience such an approach helps the reader to better follow the progression of ideas and makes clear the simplifications.

Comment 6

Can you explain what would happen with the proposed approach if there is an obstacle to a goal?

This is an important observation by the reviewer. An obstacle to a goal can indeed cause 'unintuitive changes' to the direct path to a goal. However, we would like to emphasize that, how well the autonomy is able to reason about these kinds of changes in motion direction depends on the features (directedness, proximity, agreement etc) used by the **intent inference** module. That ability of the intent disambiguation system to disambiguate intent is only as effective as the underlying intent inference mechanism. This is because, during the forward projection of beliefs (Line 7 in Algorithm 1), the BeliefUpdate() utilizes the same inference algorithm that is active during task execution. If the choice of intent inference algorithm ignores the presence of obstacles (probably not the best choice!), intent disambiguation will not be able to reason about the presence of obstacles.

One can think of the intent disambiguation as an add-on module to whichever intent inference scheme is used, such that the disambiguation system is able to rank order the (discrete) parameters on which the inference depends on according to their ability to extract information that will help to enhance inference accuracy.

Comment 7

..Where are the results for the dynamic neural field approach? A multitude of different approaches is proposed which is exceptionally confusing and it is very hard to follow what is relevant in this paper. Or to rephrase, how disambiguation, inference, and assistance concepts participate in the evaluation..?

We have cited one of our other works which does a qualitative comparison of dynamic neural field approach to other standard approaches for intent inference.

The reason why we have used a multitude of approaches for each component of the overall system was to more rigorously evaluate the robustness of our proposed intent disambiguation approach—by demonstrating that its effectiveness was not contingent on a particular choice of intent inference algorithm, shared control algorithm or disambiguation metric. We hope that Figure 1 helps to clarify how all of the different components interact with each other.

Comment 8

..The simulation environment, task, and the complete evaluation setup need to be thoroughly explained. Since the actual user study is missing, it is very important to describe in detail (with equations) how is the human simulated.

Response

Thank you for the suggestion. We have described the simulation environment in greater detail, and hope that the modifications in the current version make the simulation setup more comprehensible to the reader. In particular, Section V.B.3 describes how the human actions are simulated. Further details of how the simulations were conducted are also described in Section V.B.5 and in simulation parameters and their ranges are presented were also presented in Tables I and II. Furthermore, we are also making the point robot simulation code open source and is now available on GitHub. The paper contains the link to the repo as well.