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

IEEE Robotics And Automation Letters (RA-L)

May 27, 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 this 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.

We apologize the confusion that arose to the reviewer, but we would like to point that computing the disambiguation mode is different from predicting the intent itself. Algorithm 1 is about characterizing intent disambiguation capabilities of each control mode.

We would like to emphasize that the robot assistance, intent inference and intent disambiguation all are working in tandem as the human (simulated) interacts with the robot. In order to bring this point across, we have added a new figure which depicts how the different components interact with each other. We sincerely hope that this will clarify how the output of Algorithm 1 is integrated within the robot assistance system.

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.

We have made every attempt to make the simulation details clear. Tables I and II provides all the simulation parameters and their ranges. We had already specified how u_h (Equation 7) and u_r were generated (Section V.A.1). We have modified Section V.B.3 to back reference Equation 7 and V.A.1. RA-L does not allow Appendix and supplementary material and therefore we won't be able 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 disagree with the reviewer on this point. The goal with our simulation is to be able to control for various confounding factors and be able to randomize them properly so that the analysis can focus on the factor of interest, which in our case is efficacy of our intent disambiguation algorithms. Simulated data serve as a proxy/an approximation of what real data might look like. for example, this is similar to how modern deep learning techniques use data augmentation methods to create 'synthetic data' to train their systems which can then be deployed in real scenarios. Training on simulated data does not invalidate the models learned but only improves their performance in the real world.

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 accordingly for improved readability throughout the paper as well as in the figures.

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 compared 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 hand-crafted heuristic features to perform intent disambiguation. Our goal was to identify at least one information-theoretic approach than could possibly be better than the heuristic approach. We are in the process of extending the current study into a full fledged human subject study in which we will be testing both ENT and KLD against the GRD. We were however, unable to finish the experiment in time for the submission deadline.

We would still like to point out that the KLD-based intent disambiguation system was still more effective than the GRD condition thereby marking an improvement over a heuristic based approach and validates the use of information theoretic approaches for intent disambiguation.

We would also like to emphasize that the major difference between the point robot simulations and the robotic arm simulation setup is the nonlinear kinematics. The robot autonomy was still generated using the potential field approach.

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. 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 and references 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 this point. We highlight this in a general manner during the introduction in which we state that intent inference is crucial for the success for many shared-autonomy human machine systems. Any robotic assistance 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 the case of assistive robotic manipulation) 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 think that when the math is presented it is important to start with the most general form of the equations and introduce simplifications based on different assumptions 'on-the-fly'. This way, the logical progression of ideas and simplification is clear for the reader.

Comment 6

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

This is an interesting question and we would like to point out that the 'intent disambiguation' algorithm is designed to improve the accuracy of the intent inference module. The robot autonomy is generated from a policy that is **independent** of the intent disambiguation algorithm. In our case, we utilize a potential field based approach to generate robot autonomy signals. The potential field generator utilizes the intent inference module's prediction of what the current goal and treats the predicted goal as the sole 'attractor in the world'. All the other objects in the world are then treated as obstacles and therefore will have associated repeller fields. We have modified Section V.A.1 to include this information and we hope this would clarify this point.

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 included one of our other works which does a qualitative comparison of dynamic neural field approach to other standard approaches for intent inference. We would like to argue that the reason why we have used a multitude of approaches for each component of the overall system is that, in simulation we wanted to ensure the robustness of our proposed intent disambiguation approach to the choice of different algorithms for other components of the system. We would like to point out that the intent inference, intent disambiguation and shared control assistance are all working in tandem as the human interacts with the robot. This is exemplified in the Figure 1, that depicts how the different components work together. We hope this clarifies how all the different components 'participate in the evaluation'.

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

We have made every effort to describe the simulation environment in greater detail. We hope that the modifications in the current version will be sufficient to make the simulation setup more comprehensible to the reader. 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.