Information-Theoretic Characterization of Control Modes for Intent Disambiguation

Deepak Gopinath[1] and Brenna D. Argall[2]

***Abstract*— The effectiveness of assistive robots is closely related to their ability to infer the user's needs and intentions *unambiguously* and provide appropriate assistance quickly and accurately. In this paper, we propose a mode selection paradigm that enhances the autonomy's intent inference capabilities by exploiting the fact that how the human operates the robot and reveals his/her intent is intrinsically conditioned on the control mode in operation. We formulate this as a problem of intent disambiguation in information-theoretic terms. We propose two different methods for enhancing intent inference via improved disambiguation using the information-theoretic concepts of *entropy* and *KL divergence*. In our system, the autonomy maintains a probability distribution over goals (beliefs) and we characterize the disambiguation capabilities of the different control modes/dimensions utilizing a model-based approach to compute the average information gain and content. Previous work has shown that the success of a disambiguation algorithm depends on a variety of factors and parameters. We present results from an extensive simulation-based study for both a point robot and physics-based simulation of a six degrees-of-freedom (DoF) robotic arm. Our results indicate that, compared to a baseline, the proposed disambiguation algorithms do enhance intent inference via faster intent disambiguation, which in turn allows autonomy assistance to step in earlier during task execution. We also find that goal inference is more accurate and the total amount of time assistance is engaged is higher.**

48 pt
0.667 in
16.9 mm

I. INTRODUCTION

Assistive machines such as robotic arms and smart wheelchairs have the potential to transform the lives of people with motor impairments [1]. These machines can promote independence, enhance quality of lives, extend the mobility and manipulation capabilities of individuals, and revolutionize the way people interact with society.

An assistive robotic machine is typically controlled using an interface such as a joystick, a switch-based head array or a sip-and-puff. These interfaces are low-dimensional, low-bandwidth and occasionally discrete. For this reason, at any given point in time during task execution they can only operate in a subset of the entire control space. These subsets are referred to as *control modes* [2]. To schedule and execute switches between control modes can be both mentally and physically demanding, and might be alleviated in part by the introduction of robotics autonomy. The efficacy of such autonomy-endowed machines relies on their ability to infer the users' needs and intentions, which is often a bottleneck for providing appropriate assistance accurately, confidently and quickly. Due to the dimensionality mismatch between high-dimensional robots and low-dimensional interfaces, the user is constrained to produce control commands that likely do not express their true underlying intent unambiguously. That is, the control interfaces act like filters that restrict the amount of information regarding the user's intent that is passed through to the autonomy.

In the context of assistive robotic manipulation, the first step of a task is often to *reach* towards discrete objects in the environment. Therefore in this domain, *intent inference* can be cast as a problem of maintaining and updating the probability distributions over all possible discrete goals (objects) in the workspace upon receiving fresh evidence from the human actions and environment. In Bayesian terminology, maintaining and updating the belief is the *recursive Bayesian filtering* problem. Intent inference algorithms typically rely on various task-relevant features such as the robot and goal positions, human control commands and biometric measures [3] as their evidence variables. In contrast, due to reasons of user adoption and cost, our system matches the interfaces available on typical commercial assistive machines (such as powered wheelchairs) and so infers intent exclusively based on the information contained in the constrained control commands issued to the assistive machine.

48 pt
0.667 in
16.9 mm

Our system is inspired by the following insights:

- At any given time, the user is constrained to operate the robot in a specific control mode. That is, the policy followed by the human while executing the reaching task is conditioned on the current active mode.
- The posterior computation of the belief over goals depends on the human policy which indirectly depends on the active control mode.
- Utilizing a model-based approach to project the beliefs over goals forward in time, the autonomy can *counterfactually* reason about the information gain and content for different control modes.
- Using this information-theoretic characterization of control modes, the autonomy chooses the control mode with the highest intent disambiguation capability *for* the human.
- Having the human provide control input within these information-rich control modes can likely improve the accuracy of intent inference. This, in turn, allows the

This material is based upon work supported by the National Science Foundation under Grant CNS 1544741. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the aforementioned institutions.

[1]Deepak Gopinath is with the Department of Mechanical Engineering, Northwestern University, Evanston, IL, USA and Shirley Ryan AbilityLab, Chicago, IL, USA deepak.gopinath@u.northwestern.edu

[2]Brenna D. Argall is with the Department of Mechanical Engineering, Department of Electrical Engineering and Computer Science, Northwestern University, Evanston, IL, Department of Physical Medicine and Rehabilitation, Northwestern University, Chicago, IL and Shirley Ryan AbilityLab, Chicago, IL. brenna.argall@northwestern.edu

autonomy to step in and *assist the human achieve their desired goal earlier*.

In this paper, we address how to enhance the accuracy of intent inference by characterizing control modes according to their information content and proposing a mode switch assistance paradigm that selects the control mode in which a user-initiated motion will *maximally disambiguate* their intent. We formalize the problem of intent disambiguation within the framework of information theory. Specifically, we make use of the information-theoretic notions of entropy and Kullback-Leibler (KL) divergence to characterize and quantify the information content—with respect to intent disambiguation—of each control mode.

In Section II we present an overview of relevant research. Section III introduces our set theoretic treatment of control modes, followed by intent disambiguation and intent inference in Sections IV and V respectively. The study design and experimental methods are discussed in Section VI followed, by results in Section VII. Discussion and conclusions are presented in Sections VIII and IX respectively.

II. Related Work

This section presents an overview of related works in information acquisition in robotics, intent inference in human-robot interaction and robot assistance for modal control.

Robot assistance schemes that seek to elicit more informative cues *from* the human to clarify the underlying intent can be thought of as an information acquisition problem. Intent acquisition can leverage the underlying synergies and shared intentionality [4] of human-robot teams and can be an active process in which the robot performs actions that will nudge the human to reveal her/his intent more clearly [5]. Information-theoretic metrics such as KL divergence can be utilized to identify regions of the sample space that will maximize information gain [6] and subsequently guide the data sampling process. Sensing robots that are designed for automated exploration and data acquisition tasks can benefit from exploring more information-rich regions in the environment [7]. If the spatial distribution of information density is known *a priori*, information maximization can be accomplished by maximizing the ergodicity of the robot's trajectory with respect to the underlying information density [8].

Intent inference algorithms play a vital role in the success of an assistive system. Intent inference and recognition can be classified into two broad categories: heuristic approaches and model-based approaches. Heuristic approaches are often simpler and computationally light-weight and seek to find direct mappings between various task relevant features (such as motion cues) and the human's underlying intention [9]. On the other hand, in model-based approaches the system maintains a model of how a human maps states to control actions. The model can either be learned from data or can be hand-designed based on domain knowledge. For example, the human can be modeled within the Partially Observable Markov Decision Process (POMDP) [10] framework and is assumed to behave according to a control policy that maps states to actions. However, in model-based approaches incorporating the entire history of states requires estimation of joint probability distributions over past states which can become computationally expensive and intractable quickly.

When there is a mismatch between the dimensionality of the problem and the control interface users have to continuously shift their focus from the task at hand to the choice of control mode during task execution, thereby resulting in a higher cognitive load. Various mode switching assistance paradigms have been proposed to alleviate task effort. For example, an automatic time-optimal mode switch assistance has been proposed which has shown to significantly improve user satisfaction [11]. In the area of myoelectric prosthetics, dynamic switching approaches that learn the most effective control option during task execution using temporal difference and reinforcement learning have also been proposed [12].

III. Mathematical Notation

This section describes the mathematical notation used in our intent disambiguation algorithm that computes a control mode that maximally disambiguates between the various goals. We develop this algorithm specifically for robotic manipulation scenarios in which the user is controlling a robotic arm to reach for and interact with various discrete objects in the environment.

A. Probability Distribution Over Goals

In assistive robotic manipulation, intent inference is most commonly the process of estimating the user's intended goal out of a set of discrete objects in the environment [13]. The set of all candidate goals is denoted by $\mathscr{G}$ with $n_g = |\mathscr{G}|$. At time t the random variable $g^t \in \mathscr{G}$ denotes the user's intended goal and b_{g^t} denotes the belief over the goals.

B. Set Theoretic Treatment of Control Modes

The low dimensionality of the control interfaces necessitates the control space to be partitioned into control modes. Let $\mathscr{K}$ be the set of all controllable dimensions of the robot and k^i represent the i^{th} control dimension where $i \in [1,2,\ldots,n_k]$ with $n_k = |\mathscr{K}|$. The number of controllable dimensions (n_k) depends on the robotic platform; for example, a 2D point robot that operates in $\mathbb{R}^2$ has $n_k = 2$ whereas $n_k = 6$ for a 6-DoF robotic manipulator.

Let $\mathscr{M}$ denote the set of all control modes with $n_m = |\mathscr{M}|$. Additionally, let m^i refer to the i^{th} control mode where $i \in [1,2,\ldots,n_m]$. Each control mode m^i is a subset of $\mathscr{K}$ such that $\bigcup_{i=1}^{n_m} m^i = \mathscr{K}$. The cardinality of each mode $m \in \mathscr{M}$, denoted by $|m|$, indicates the number of dimensions that can be controlled when operating in m^1. Furthermore, the user can only operate in one of the n_m control modes at any given time t. That is, for each $m \in \mathscr{M}$, the subspace of $\mathbb{R}^{n_k}$ that is accessible corresponds to $\mathbb{R}^{|m|}$ whose orthonormal basis vectors are given by $e^k \;\forall\; k \in m$. Maximum velocity limits along each dimension impose further constraints on the set of control commands that are available in each mode.

[1]Note that a dimension $k \in \mathscr{K}$ can be an element of multiple control modes and so it is possible that $m^i \cap m^j \neq \emptyset$.

IV. Intent Disambiguation

This section describes how our autonomy reasons about the intent disambiguation capabilities of different control modes by computing the information gain and content of the belief over goals. We utilize a model-based approach in which the human is modeled as a rational agent that seeks to take the shortest distance path to the goal. We also describe the various intent inference schemes that we utilize in our experiments, which work in conjunction with our disambiguation algorithm.

A. Conditional Recursive Bayesian Belief Update

In this subsection we show how the recursive Bayesian update of the probability distribution over goals is explicitly conditioned on the current active control mode and therefore can be leveraged to reason about disambiguation.

We only consider a single evidence variable that corresponds to the human control command denoted as u_h. The goal probability conditioned on the evidence variable can be written as

$$b_{g^t} = p(g^t \mid u_h^{0:t}) \propto p(u_h^t \mid g^t, u_h^{0:t-1}) p(g^t | u_h^{0:t-1}) \quad . \tag{1}$$

Assuming that $(u_h^t \perp u_h^{0:t-1} | g^t)$ (Markovian assumption that the control command at the current timestep is independent of previous timesteps given the current goal)and marginalizing over g^{t-1}, Equation 1 becomes

$$b_{g^t} = p(u_h^t \mid g^t) \sum_{g^{t-1} \in \mathcal{G}} p(g^t, g^{t-1} \mid u_h^{0:t-1}) \tag{2}$$

We also assume that the conditional transition probability of changing to goal g^t given that the goal is g^{t-1} is independent of the control command history. Equation 2, therefore, simplifies to

$$b_{g^t} = p(u_h^t \mid g^t) \sum_{g^{t-1} \in \mathcal{G}} p(g^t | g^{t-1}) b_{g^{t-1}} \tag{3}$$

Marginalizing over m, the control mode variable, the human policy $p(u_h^t \mid g^t)$ can be written as

$$p(u_h^t \mid g^t) = \sum_{m \in \mathcal{M}} p(u_h^t \mid g^t, m) p(m) \tag{4}$$

We can simplify Equation 4 using the critical piece of information that, due to constraint of the control interface, at any given time t only one of the control modes in $\mathcal{M}$ can be active. That is, the probability distribution over modes at any given time t reduces to a delta function. This can be written as

$$p(m) = \begin{cases} 1 & \text{if } m = m^t \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

Using Equations 4 and 5 and under the assumption that the user does not change the goal during task execution, Equation 3 can be further simplified as

$$b_{g^t} = \underbrace{p(u_h^t \mid g^t, m^t)}_{\substack{\text{human policy} \\ \text{conditioned on mode}}} \underbrace{b_{g^{t-1}}}_{\substack{\text{belief at} \\ t-1}} \tag{6}$$

From Equation 6 we can see that the belief at time t depends on the current mode m^t. This is due to the fact that the mode at time t has a direct influence on how the human chooses to operate the robot to accomplish the task.

B. Algorithm: Information Theoretic Intent Disambiguation

Our aim is to develop a metric that will capture the "disambiguation capability" of a control dimension/mode. From Equation 6 we can see that the time evolution of the probability distribution is sensitive to the user control command (denoted by u_h), and that it will evolve differently as the user controls the robot in different control modes. By utilizing a model-based approach, the autonomy can reason about the time evolution of the belief b_{g^t} for all control modes $m \in \mathcal{M}$ and then select the mode that offers the highest information gain and content. That is, motion within maximally disambiguating control dimensions serves as a mechanism to enhance the accuracy of intent inference.

The full algorithm is presented in Algorithm 1. We first perform a model-based projection of the probability distribution over goals (lines 5-9, Section IV-C). We then compute the disambiguation capability D_m of a control mode using information-theoretic measures (lines 10-11) We present two different information-theoretic measures for disambiguation: 1) information content in the projected probability distributions over goal as encoded by the *entropy* of the distribution (Section IV-D) and 2) information gain as a result of the time evolution of the belief as encoded by the *KL divergence* between the posterior and the prior distributions (Section IV-E).

The control mode with highest disambiguation capability m^* is then given by $\text{argmax}_m D_m$. Disambiguation mode m^* is the control mode that system chooses *for* the human to better estimate their intent. Any control command issued by the human when operating in m^* is likely to be more beneficial for the system to determine the human's intended goal.

C. Model-based Projection of b_{g^t}

The first step towards the computation of the disambiguation metric D_m for each $m \in \mathcal{M}$ is the forward projection of the probability distribution b_{g^t} from current time t_a to t_b such that $t_a < t_b$. We rely on a model-based approach in which the human policy under the Boltzmann rationality assumption [14] is modeled as a von Mises distribution [15]. That is,

$$p(u_h \mid g) = C_p(\kappa) \mathrm{e}^{(\kappa \cdot Q_g(u_h, x_g, x_r))} \tag{7}$$

where $C_p(\kappa)$ is the normalization constant, κ is the concentration parameter, p is the dimensionality of the control space and $Q_g(u_h, x_g, x_r)$ is modeled as the cost of taking action u_h towards goal g at robot configuration x_r when acting optimally. We compute this cost as the *agreement* between the human control command u_h and the vector connecting the current robot configuration x_r and the goal configuration x_g. The human thus is modeled as attempting the straight line path to the goal. The human policy conditioned on control mode m, $p(u_h \mid g, m)$, is the projection of $p(u_h \mid g)$ onto the subspace $\mathbb{R}^{|m|}$ spanned by the control mode m.

Given this human model for each control mode $m \in \mathcal{M}$, the autonomy adopts a sampling-based approach to estimate

the time evolution of belief for all goals $g \in \mathscr{G}$. This model-based projection enables the autonomy to counterfactually reason about how much information gain is likely to happen when users operate the robot towards different goals in each of the control modes.

D. Entropy Disambiguation Metric

The entropy of a probability distribution captures the average information content of a stochastic source of data. Lower entropy indicates higher certainty in the value of the random variable, and vice-versa. Therefore, in the context of intent disambiguation, the entropy of the projected probability distribution, $b_{g^{t_b}}$, can be used as a measure of how confident the system is in its prediction of human intent. That is, lower the entropy better the disambiguation, due to higher certainty in the human's intended goal. For a discrete random variable X with possible values $\{x_1, x_2, \ldots x_n\}$, the Shannon entropy is defined as

$$ENT(p(X)) = -\sum_{i=1}^{n} p(x_i) log_2(p(x_i))$$

where $p(X)$ denotes the probability mass function. The disambiguation capability of a control mode m is characterized by computing a weighted average of the entropies of projected probability distributions. That is,

$$D_m = \frac{1}{|n_g|} \sum_{g \in \mathscr{G}} \mathbb{E}_{u_h \sim p(u_h \mid g,m)} \Big[ENT(b_{g^{t_b}}) \Big] \tag{8}$$

where $b_{g^{t_b}}$ denotes the projected probability distribution at time t_b when the control command used for forward projection is sampled from the mode conditioned human policy $p(u_h \mid g, m)$.

E. KL Divergence Disambiguation

Although entropy can capture information *content*, intent disambiguation can likely benefit from the quantification of the information *gain* regarding the human's intended goal as a result of user-initiated motion in a control mode m. KL divergence, also known as relative entropy, measures how a probability distribution differs from another distribution. KL divergence is widely used in the context of Bayesian inference to compute the information gain when the prior is updated to the posterior in the light of new evidence. In the context of disambiguation we can treat the projected probability distribution at time t_b as the posterior and the distribution at time t_a to be the prior. KL divergence can be used to characterize the *information gain* regarding the human's intended goal as a result of the application of u_h.

For a discrete random variable X with possible values $\{x_1, x_2, \ldots x_n\}$ the KL divergence is defined by

$$KL(P||Q) = -\sum_{i=1}^{n} p(x_i) log_2 \frac{q(x_i)}{p(x_i)}$$

where $p(X)$ and $q(X)$ are two different probability mass distributions. More specifically, the disambiguation capability of control mode m is computed by averaging the information gain for all projections of probability distributions $b_{g^{t_b}}$. The disambiguation metric can be computed as

$$D_m = \frac{1}{|n_g|} \sum_{g \in \mathscr{G}} \mathbb{E}_{u_h \sim p(u_h \mid g,m)} \Big[KL(b_{g^{t_b}} || b_{g^{t_a}}) \Big] \quad . \tag{9}$$

Algorithm 1 Information Theoretic Intent Disambiguation

Require: $b_{g^{t_a}}, x_r^{t_a}, \Delta t, t_a < t_b$
1: **for** $m \in \mathscr{M}$ **do**
2: Initialize $D_m = 0$
3: **for** $g \in \mathscr{G}$ **do**
4: Initialize $t = t_a, x_r^t = x_r^{t_a}, b_{g^t} = b_{g^{t_a}}$
5: **while** $t \leq t_b$ **do**
6: $u_h^t \sim p(u_h \mid g, m)$
7: $b_{g^{t+\Delta t}} \leftarrow \text{BeliefUpdate}(b_{g^t}, u_h^t)$
8: $x_r^{t+\Delta t} \leftarrow \text{SimulateKinematics}(x_r^t, u_h^t)$
9: $t \leftarrow t + \Delta t$
10: **end while**
11: $D_m \leftarrow D_m + \frac{1}{n_g} \mathbb{E}_{u_h \sim p(u_h \mid g,m)} \Big[ENT(b_{g^{t_b}}) \Big]$
or
12: $D_m \leftarrow D_m + \frac{1}{n_g} \mathbb{E}_{u_h \sim p(u_h \mid g,m)} \Big[KL(b_{g^{t_b}} || b_{g^{t_a}}) \Big]$
13: **end for**
14: **end for**

V. Intent Inference

At any given time t, the probability distributions over goals encodes the human's underlying intent. The time evolution of the probability distribution, however, depends on the choice of intent inference scheme, and the various task relevant features and parameters that contribute to it.

More specifically, the exact update rule for the recursive belief update is determined by the choice of inference scheme (Algorithm 1, Line 6) which impacts the projected probability distributions, $b_g^{t_b}$, and the disambiguation metric, D_m. Here we present the multiple intent schemes evaluated in conjunction with our disambiguation algorithm.

1) Heuristic Approaches: Heuristic approaches based on *confidence* functions [16] seek to find direct mappings between instantaneous cues and underlying human intentions. For example, a confidence function that captures the "directedness" of the human control command to a particular goal position can be written as

$$b_{g^t} = u_h^t \cdot (x_g^t - x_r{}^t)$$

where $u_h^t \sim p(u_h \mid g, m)$ is the human control command. These confidence functions rely on instantaneous features and therefore are amnesic and can exhibit chatter behavior [17].

2) Bayesian Approaches: Bayesian approaches for intent inference consist of iteratively updating of the belief (probability distribution over goals) as new evidence arrives at every time-step. The Bayesian update procedure directly uses Equation 6 for belief propagation.

3) Dynamic Neural Field Approaches: Dynamic neural fields were originally conceived to explain cortical population neuronal dynamics [18]. When applied to the problem of intent inference, the individual goal probabilities are treated as constrained dynamical state variables whose time evolution is determined by a dynamic neural field such that $p^i(t) \in [0,1]$ and $\Sigma_1^{n_g} p^i(t) = 1$, where $p^i(t)$ denotes the probability associated with the i^{th} goal at time t and $\boldsymbol{p}(t) = [p^1(t), p^2(t), \ldots, p^{n_g}(t)]^T$.

The full specification of the neural field is given by

$$\frac{\partial \boldsymbol{p}(t)}{\partial t} = \frac{1}{\tau}\Bigg[-\mathbb{I}_{n_g \times n_g} \cdot \boldsymbol{p}(t) + \underbrace{\frac{1}{n_g} \cdot \mathbb{1}_{n_g}}_{\text{rest state}} \Bigg] + \underbrace{\lambda_{n_g \times n_g} \cdot \sigma(\xi(u_h^t; \Theta))}_{\text{excitatory + inhibitory}} \quad (10)$$

where time-scale parameter τ determines the memory capacity of the system, $\mathbb{I}_{n_g \times n_g}$ is the identity matrix, $\mathbb{1}_{n_g}$ is a vector of dimension $n_g \times 1$ containing all ones, λ is the control matrix that controls the excitatory and inhibitory aspects, ξ is a nonlinear function through which human control commands and task features affect the time evolution, Θ represents all other task-relevant features and parameters, and σ is a biased sigmoidal nonlinearity given by $\sigma(\xi) = \frac{1}{1+e^{-\xi}} - 0.5$. The design of ξ is informed by what features of the human control input and environment capture the human's underlying intent most effectively. Our implementation relies on the *directedness* of the human control commands towards a goal, the *proximity* to a goal and the *agreement* between the human commands and robot autonomy.

VI. Experimental Evaluation

We evaluate the efficacy of the proposed disambiguation algorithm on simulated point robots and a simulated robotic manipulator that operates in a *shared control* setting. This section describes the shared-control paradigm used in our experiments and the details of the simulation experiment.

A. Shared Control

Shared autonomy in assistive robotics aims to reduce the human's physical and cognitive burden during task execution without having the user to completely cede manual control [19]. In our experimental evaluation, control sharing is achieved with a blending-based paradigm, in which the final control command issued to the robot is a weighted linear combination of the user control command and robot autonomy.

1) Robot Autonomy: The robot autonomy command $u_{r,g}$ is generated using a simple potential field, which is defined in all parts of the state space [20]. Every goal g is associated with a potential field γ_g which treats g as an attractor and all the other goals in the scene as repellers.

2) Blending Paradigm: The final control command u_f issued to the robot then is given by

$$u_f = \alpha \cdot u_{r,g^*} + (1-\alpha) \cdot u_h$$

Variable	Range
n_g	$[2,6]$
$\mathscr{M}$	Randomly selected from Table II conditioned on the space
Intent Inference	['Confidence', 'Bayes', 'Dynamic Neural Field']
Initial Robot Position	Randomized within workspace limits
Goal Positions	Randomized within workspace limits
Intended Goal	Randomly chosen from $\mathscr{G}$

TABLE I

Randomized Factors for Each Simulation Trial

Space	n_k	Control Mode Sets
$\mathbb{R}^2$	2	$\{[1],[2]\}$
$\mathbb{R}^3$	3	$\{[1],[2],[3]\}$, $\{[1,2],[3]\}$
$\mathbb{SE}(2)$	3	$\{[1],[2],[3]\}$, $\{[1,2],[3]\}$
$\mathbb{SE}(3)$	6	$\{[1],[2],[3],[4],[5],[6]\}$, $\{[1,2,3],[4,5,6]\}$, $\{[1,2],[1,3],[4,5],[6]\}$, $\{[1,2],[1,3],[4,5],[6]\}$

TABLE II

Predefined control space partition sets for each of the simulation spaces. The numbers in the rightmost column indicate the control dimensions.

where g^* is the most confident goal and corresponds to the mode of b_{g^t}. The blending parameter, α is a piecewise linear function of the $p(g^*)$ and is specified as

$$\alpha = \begin{cases} 0 & p(g^*) \leq \rho_1 \\ \frac{\rho_3}{\rho_2 - \rho_1} \cdot p(g^*) & \text{if } \rho_1 < p(g^*) \leq \rho_2 \\ \rho_3 & p(g^*) > \rho_2 \end{cases}$$

with $\rho_i \in [0,1] \; \forall \; i \in [1,2,3]$ and $\rho_2 > \rho_1$. In our implementation, we empirically set $\rho_1 = \frac{1.05}{n_g}, \rho_2 = \frac{1.1}{n_g}$ and $\rho_3 = 0.7$.

3) Maximum Potential Mode Switch Assistance: As a baseline for the simulation experiments described in Section VI-B we utilize a simple greedy mode switch assistance scheme. At any given time t, the system selects the control mode which has the *maximum attractive pull* or *maximum potential* towards the predicted goal. The attractive potential of each control mode $m \in \mathscr{M}$ is computed as the total attractive potential of all control dimensions $k \in m$.

B. Point Robot Simulation Setup

1) Robot Workspace: Our simulations were performed on point robots that reside in $\mathbb{R}^2$, $\mathbb{R}^3$, $\mathbb{SE}(2)$ and $\mathbb{SE}(3)$ spaces. The translational workspace limits were set at $[-0.6m, 0.6m]$ along each translational dimension and the orientation limits were set at $[0, 2\pi]$ along each rotational axis. Table I lists all the factors there were manipulated for each trial.

2) Mode Switch Schemes: Four different mode switching schemes were activated during each simulation trial, three of which performed intent disambiguation, and one which served as a baseline for comparison. The disambiguation algorithms were entropy disambiguation metric (Section IV-D), KL divergence disambiguation metric (Section IV-E) and a heuristic approach developed in a related work [21]. We

used the mode switching scheme described in Section VI-A.3 as the baseline for comparison.

3) u_r and u_h^{sim}: The simulated robot autonomy u_r^{sim} is generated using a repeller-free potential field. For an intended goal $g^{intended}$ and a given control mode $m \in \mathscr{M}$, the simulated human control command denoted by u_h^{sim} is sampled from the mode conditioned policy, $p(u_h \mid g^{intended}, m)$.

4) Metrics: The following metrics were used to evaluate the efficacy of the disambiguation algorithm compared to the baseline mode switch assistance:

- *Initial Onset of Assistance*: Earliest time (normalized with respect to individual trial time) at which assistance was triggered towards the intended goal. This measure captures how *early* the robot was able to infer the correct goal *and* provide assistance.
- *Total Amount of Assistance* : Fraction of the total trial time for which assistance towards the intended goal was present.
- *Inference Accuracy* : Fraction of the total trial time for which the intent inference mechanism correctly inferred the human's intended goal. This is a measure of overall accuracy of the system.

5) Simulation Protocol: 8000 simulation runs were performed for each of the scenarios. The maximum trial duration was set at $12s$ or equivalently 120 time-steps with $\Delta t = 0.1$. The forward projection time was set to be 4 seconds and the mode switch assistance algorithm was activated once every 10 time-steps. The trial was ended prematurely if the robot reached within the intended goal as determined by a pre-specified distance threshold. The expectation in Equations 8 and 9 was computed by averaging over five samples. A *dropout* factor was added to the human control command to mimic the signal sparsity exhibited by humans during robot teleoperation. Figure 1 illustrates an example of the simulation for a point robot in $\mathbb{R}^2$.

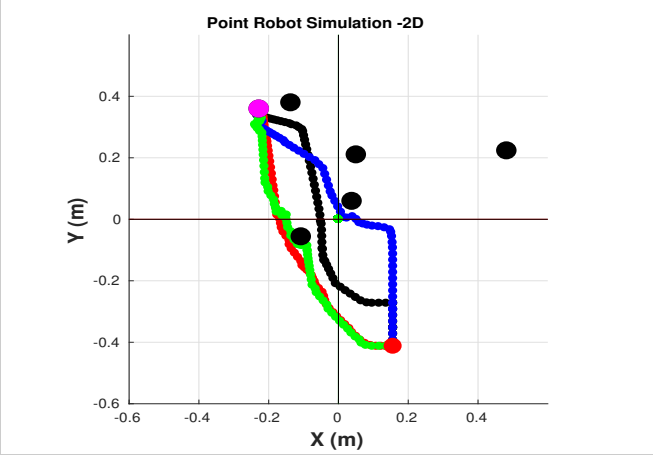

Fig. 1. Example of point robot simulation in $\mathbb{R}^2$. The goals are represented in black, the initial robot position in red and the human's intended goal in magenta. The different colored trajectories correspond to different mode switching schemes. Greedy (Black), Entropy (Red), KL Divergence (Blue) and Heuristic (Green).

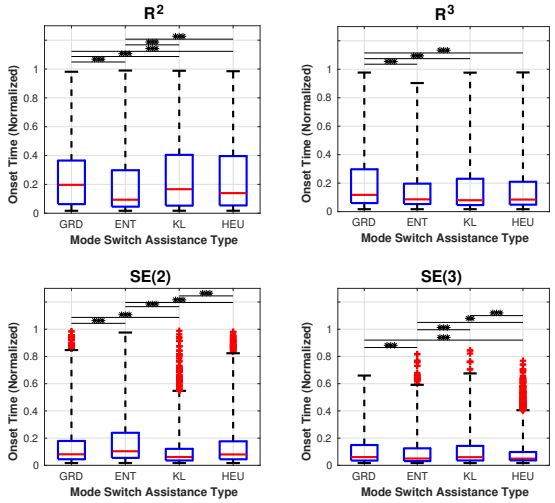

Fig. 2. Initial onset of assistance. Group differences between mode switch assistance schemes (GRD, ENT, KL, HEU), in four operational spaces $\mathbb{R}^2$, $\mathbb{R}^3$, $\mathbb{SE}(2)$ and $\mathbb{SE}(3)$. Box plots show the median and the quartiles.

C. Robotic Arm Simulation Setup

We also evaluated our system in a physics-based simulation of a 6-DoF robotic arm (MICO robotic arm from Kinova Robotics). Although the end effector of the robot lives in $\mathbb{SE}(3)$, the physical constraints of the robot make the kinematics highly nonlinear compared to a point robot in $\mathbb{SE}(3)$.

We evaluated the system on two reaching tasks: 1) Four goals with same orientation and 2) five goals with different orientations. Two mode switching schemes were tested: KL divergence disambiguation and the maximum potential mode switching. During each trial, the mode switching algorithm was activated every $5s$ and the maximum trial duration was set at $40s$. The robot autonomy was generated using a potential field as described in Section VI-A.1 and the control interface mapping was 1D and discrete (similar to that of a headarray). A total of 128 trials were collected for each task.

VII. Results

In this section we present results from our simulation studies conducted on a simulated 6-DoF robotic arm and point robots in $\mathbb{R}^2$, $\mathbb{R}^3$, $\mathbb{SE}(2)$ and $\mathbb{SE}(3)$.

A. Point Robot Simulation

One way ANOVA was performed using Kruskal-Wallis procedure to test for group effects between the four mode switch schemes: greedy (GRD), entropy (ENT), KL-Divergence (KL) and heuristic (HEU). Post-hoc analysis between groups was conducted using a multiple comparison test that used Tukey's HSD criterion. All analysis was performed in MATLAB. In all the data plots, (*) indicates $p < 0.05$, (**) indicates $p < 0.01$ and (***) indicates $p < 0.001$.

1) Initial Onset of Assistance: The Kruskal-Wallis test reveals that the group ranks are different for simulations in all operational spaces. Post-hoc analysis also reveals statistically significant differences between the baseline mode switch

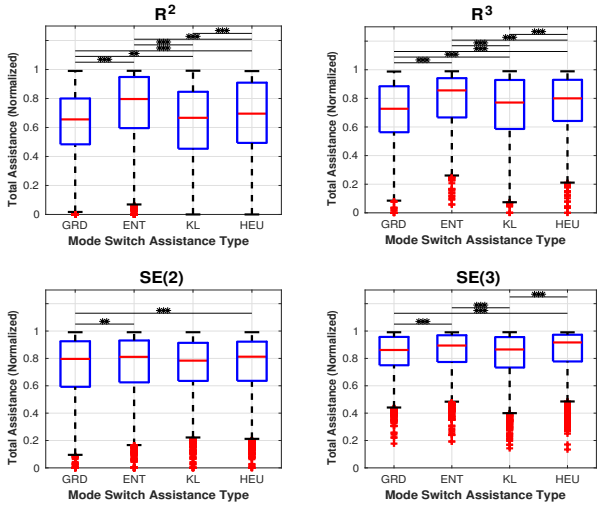

Fig. 3. Total amount of assistance. Group differences between mode switch assistance schemes (GRD, ENT, KL, HEU), in four operational spaces $\mathbb{R}^2$, $\mathbb{R}^3$, $\mathbb{SE}(2)$ and $\mathbb{SE}(3)$. Box plots show the median and the quartiles.

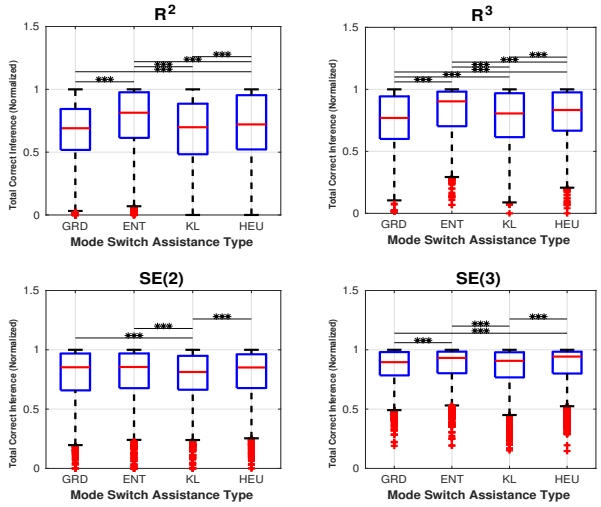

Fig. 4. Inference Accuracy. Group differences between mode switch assistance schemes (GRD, ENT, KL, HEU), in four operational spaces $\mathbb{R}^2$, $\mathbb{R}^3$, $\mathbb{SE}(2)$ and $\mathbb{SE}(3)$. Box plots show the median and the quartiles.

assistance scheme and at least one of the disambiguation schemes for all operational spaces (Figure 2). This shows that when the disambiguation scheme is used, the autonomy is able to provide assistance *earlier* during task execution. This implies that with disambiguation the system is able to elicit more informative control commands from the user and narrow down its prediction to the correct goal quicker during task execution. As a result, the confidence in the prediction becomes higher and the robot assistance is triggered earlier during a trial.

2) Amount of Total Assistance: For all operational spaces, Figure 3 shows a statistically significant increase in the total amount of assistance when disambiguation schemes are active compared to the baseline. The shared control paradigm used in our setup is designed in such a way that the robot

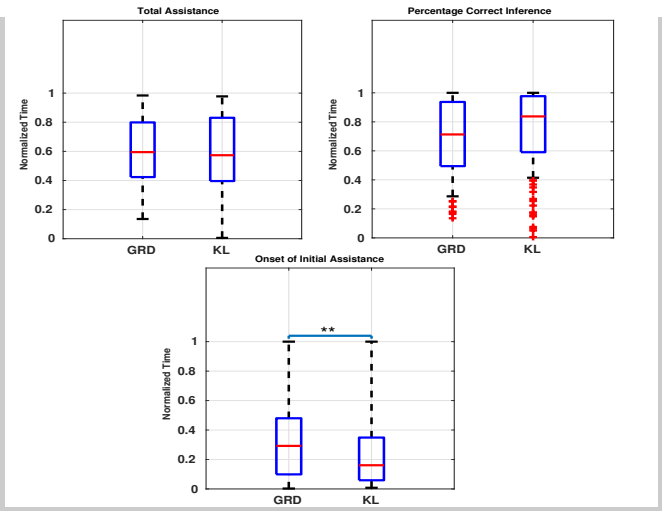

Fig. 5. Results with the simulated 6-DoF robotic arm. Top Left: Fraction of total time for which assistance towards the intended goal was engaged. Top Right: Accuracy of intent inference procedure measured in fraction of total trial time. Bottom: Onset of initial assistance as a fraction of total trial time. Statistical significance was computed using the Wilcoxon Rank-Sum test.

assistance gets triggered when the confidence associated with the goal prediction crosses a predefined threshold. These results indicate that with disambiguation the system is able to maintain the confidence in its prediction of the correct goal above the threshold for a longer duration during the trial, resulting in greater overall assistance during the task.

3) Inference Accuracy: As shown in Figure 4, for all operational spaces, the accuracy of intent inference is higher for at least one disambiguation scheme compared to the baseline. This is likely because of the fact that robot operation in disambiguating modes results in higher information gain regarding the underlying intent thereby helping the system to perform accurate intent inference.

B. Robotic Arm Simulation

Figure 5 reveals that for the simulations conducted on a simulated robotic arm, the onset of initial assistance towards the correct goal is earlier compared to that of the baseline ($p < 0.05$). Similarly, the accuracy of the intent inference procedure is higher when disambiguation algorithm is deployed, although not statistically significant. However, the total amount of assistance provided by the autonomy is comparable between the mode-switching schemes. These results indicate that even with a low-dimensional and discrete control interface that exhibits signal dropouts (due to the sparsity factor), robot operation in control modes selected by the disambiguation algorithm helps the autonomy to elicit more intent-revealing control commands thereby helping the system predict the intended goal more accurately and provide assistance earlier.

VIII. Discussion

The effectiveness of an assistive machine in a shared control setting is closely related to its ability to infer the user's intent

accurately and unambiguously. Accurate goal inference can improve the trust in human-robot interaction. On the other hand, incorrect goal inference can likely lead to suboptimal user experience and result in poor task performance. Our results indicate that disambiguation helps to enhance the accuracy of the intent inference and leads to an increase in the overall assistance provided to the user. Another factor that is crucial for the success of a shared control system, especially for safety-critical applications, is *when* during task execution the assistance gets activated. For example, in the shared control of complex systems that exhibit highly nonlinear and unstable dynamics typically the role of autonomy is to maintain stability and safety. In such situations, if the user operation leads to the destabilization of the system, failure to intervene earlier during task execution can lead to catastrophic results. From our simulation results, we can see that intent disambiguation enhances the inference capabilities of the system and helps the autonomy to provide assistance earlier.

The utility value of a mode-switching scheme can vary during the time course of task execution. For instance, in the beginning stages of task execution when the intent is unclear, disambiguation algorithms are likely to be more useful. That is, a more optimal strategy might be one which is adaptive and chooses different schemes depending on the context. This could possibly be cast within a control theoretic framework as a mode-scheduling problem in which the question of *when* to deploy *which* mode-switching paradigm can be tackled.

Disambiguation algorithms can benefit from high fidelity models of robot kinematics and user teleoperation that can be learned from data using machine learning techniques such as deep neural networks [22]. Furthermore, parallel computation on GPU can also be leveraged for more accurate estimation of the expectation in Equations 8 and 9. Lastly, the disambiguation algorithm can be customized to specific users using real-time learning algorithms that capture individual teleoperation characteristics.

IX. Conclusion

In this paper, we proposed a mode selection paradigm that enhances the autonomys intent inference capabilities by exploiting the fact that how the human operates the robot is intrinsically conditioned on the control mode in operation. We introduced two different methods for intent disambiguation using the information-theoretic concepts of entropy and KL divergence. We also present results from an extensive simulation-based study for point robots and a physics-based simulation of a robotic arm. Results from the study indicate that, compared to the baseline, the proposed disambiguation algorithms facilitated faster intent disambiguation which in turn allowed the autonomy to assist earlier during task execution. Goal inference was more accurate and the total amount of assistance engaged was also higher. In our future work, we plan to perform an in-depth analysis of the data to expose the interactions between various factors. Furthermore, as informed by the simulation results, we plan to evaluate our system with an extensive user study.

References

[1] M. P. LaPlante *et al.*, "Assistive technology devices and home accessibility features: prevalence, payment, need, and trends." *Advance Data from Vital and Health Statistics*, 1992.

[2] T. Simpson, C. Broughton, M. J. Gauthier, and A. Prochazka, "Tooth-click control of a hands-free computer interface," *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 8, pp. 2050–2056, 2008.

[3] D. Croft, "Estimating intent for human-robot interaction," in *IEEE International Conference on Advanced Robotics*. Citeseer, 2003, pp. 810–815.

[4] M. Tomasello and M. Carpenter, "Shared intentionality," *Developmental Science*, vol. 10, no. 1, pp. 121–125, 2007.

[5] D. Sadigh, S. S. Sastry, S. A. Seshia, and A. Dragan, "Information gathering actions over human internal state," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2016, pp. 66–73.

[6] S. Tong and D. Koller, "Active learning for parameter estimation in Bayesian networks," in *Advances in Neural Information Processing Systems*, 2001, pp. 647–653.

[7] N. Atanasov, J. Le Ny, K. Daniilidis, and G. J. Pappas, "Information acquisition with sensing robots: Algorithms and error bounds," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, 2014.

[8] L. M. Miller, Y. Silverman, M. A. MacIver, and T. D. Murphey, "Ergodic exploration of distributed information," *IEEE Transactions on Robotics*, vol. 32, no. 1, pp. 36–52, 2016.

[9] C. L. Baker, J. B. Tenenbaum, and R. R. Saxe, "Goal inference as inverse planning," in *Proceedings of the Cognitive Science Society*, 2007.

[10] T. Taha, J. V. Miró, and G. Dissanayake, "A POMDP framework for modelling human interaction with assistive robots," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2011, pp. 544–549.

[11] L. V. Herlant, R. M. Holladay, and S. S. Srinivasa, "Assistive teleoperation of robot arms via automatic time-optimal mode switching," in *Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 2016.

[12] P. M. Pilarski, M. R. Dawson, T. Degris, J. P. Carey, and R. S. Sutton, "Dynamic switching and real-time machine learning for improved human control of assistive biomedical robots," in *2012 4th IEEE RAS & EMBS International Conference on Biomedical Robotics and Biomechatronics (BioRob)*. IEEE, 2012, pp. 296–302.

[13] B. Calli, A. Singh, A. Walsman, S. Srinivasa, P. Abbeel, and A. M. Dollar, "The YCB object and model set: Towards common benchmarks for manipulation research," in *2015 International Conference on Advanced Robotics (ICAR)*. IEEE, 2015, pp. 510–517.

[14] A. Hula, P. R. Montague, and P. Dayan, "Monte carlo planning method estimates planning horizons during interactive social exchange," *PLoS Computational Biology*, vol. 11, no. 6, p. e1004254, 2015.

[15] K. V. Mardia, "Statistics of directional data," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 37, no. 3, pp. 349–371, 1975.

[16] D. Gopinath, S. Jain, and B. D. Argall, "Human-in-the-loop optimization of shared autonomy in assistive robotics," *IEEE Robotics and Automation Letters*, vol. 2, no. 1, pp. 247–254, 2017.

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

[18] G. Schöner, "Dynamical systems approaches to cognition," *Cambridge Handbook of Computational Cognitive Modeling*, pp. 101–126, 2008.

[19] D.-J. Kim, R. Hazlett-Knudsen, H. Culver-Godfrey, G. Rucks, T. Cunningham, D. Portee, J. Bricout, Z. Wang, and A. Behal, "How autonomy impacts performance and satisfaction: Results from a study with spinal cord injured subjects using an assistive robot," *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, vol. 42, no. 1, pp. 2–14, 2012.

[20] O. Khatib, "Real-time obstacle avoidance for manipulators and mobile robots," *The International Journal of Robotics Research*, vol. 5, no. 1, pp. 90–98, 1986.

[21] D. Gopinath and B. Argall, "Mode switch assistance to maximize human intent disambiguation," in *Robotics: Science and Systems*, 2017.

[22] A. Nagabandi, G. Kahn, R. S. Fearing, and S. Levine, "Neural network dynamics for model-based deep reinforcement learning with model-free fine-tuning," *arXiv preprint arXiv:1708.02596*, 2017.