An Information Theoretic Formalism for Intent Disambiguation

Author Names Omitted for Anonymous Review. [Paper ID: **187**]

***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 proposed 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 maintain 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. To thoroughly investigate the impact of these various components, 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.**

I. Introduction

Assistive machines such as robotic arms and smart wheelchairs have the potential to transform the lives of millions of people with motor impairments [12]. These machines can promote independence, enhance the quality of lives and revolutionize the way people interact with society. They can also help extend the mobility and manipulation capabilities of individuals, thereby helping motor-impaired people perform activities of daily living in a more effective manner.

An assistive robotic machine is typically controlled using an interface such as a joystick, a switch-based head array or a sip-and-puff ~~that~~. These interfaces are low-dimensional, low-bandwidth and occasionally discrete and typically motion commands. 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* [22]. 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, ~~and~~ 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 ~~gets~~ is passed through to the machine and 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 environmental cues and task-relevant features such as the robot and goal positions, human control commands and biometric measures [5] as their evidence variables. In contrast, due to reasons of user adoption and cost, our system infers intent exclusively based on the information contained in the constrained control commands issued to the assistive machine.

Our system is inspired by the following insights:

1) 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.
2) The posterior computation of the belief over goals depends on the human policy (treated as an evidence variable) which indirectly depends on the active control mode.
3) Utilizing a model-based approach to forward project the beliefs over goals, the autonomy can *counterfactually* reason about the information gain and content for different control modes.
4) Using this information theoretic characterization of control modes, the autonomy chooses the control mode with the highest intent disambiguation capability *for* the human.
5) ~~By~~ 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 autonomy to step in and *assist the human achieve their desired goal earlier*. This is important, because when the intent inference mechanism infers the incorrect goal (or is simply too slow), the user and autonomy often

be explicit

perhaps just bullet? (not enumerate)

end up competing instead of collaborating and can decrease user satisfaction and result in significantly worse performance. In this paper, we address the problem of accurate 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 Section IV. The study design and experimental methods are discussed in Section V followed, by results in Section VI. Discussion and conclusions are presented in Sections VII and VIII 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.

Information theoretic approaches are widely utilized in the field of machine learning and robotics for optimal experiment design, for efficient data collection processes and for informing search strategies. 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 [24] of human-robot teams and can be an active process in which the robot performs actions (for example, selecting a control mode or executing a robot motion) that will nudge the human to reveal her/his intent more clearly [19, 20]. Information theoretic metrics such as KL divergence can be utilized to identify regions of the sample space that will maximize information gain [25] and subsequently guide the data sampling process. Sensing robots designed for automated exploration and data acquisition tasks can benefit from exploring more information-rich regions in the environment [1]. 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 map [13, 14].

For an assistive machine to provide appropriate kinds of assistance accurately and at the right time, it needs to have a good estimate of the human's underlying intent. Intent inference algorithms, therefore, 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 [2, 3]. 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) [7, 23] 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. To solve this problem, 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 [9]. 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 [18].

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 [4]. The set of all candidate goals is denoted by $\mathcal{G}$ with $n_g = |\mathcal{G}|$ and let g^i refer to the i^{th} goal with $i \in [1, 2, \ldots, n_g]$. $\boldsymbol{p}(t)$ denotes the probability distribution over goals such that $\boldsymbol{p}(t) = [p^t_{g_1}, p^t_{g_2}, \ldots, p^t_{g_{n_g}}]^T$ where $p^t_{g_i}$ denotes the probability associated with goal g^i at time t. The probability $p^t_{g_i}$ also represents the robot's *confidence* that goal g^i is the human's intended goal at time t.

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 $\mathcal{K}$ 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 $n_k = |\mathcal{K}|$. 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.

I know sliding in the i indices is less elegant... but else $g^{t-1} \in G$ is ambiguous as to whether relabeling is happening over time or not. I'm happy to be dissuaded about this however

Let $\mathcal{M}$ denote the set of all control modes with $n_m = |\mathcal{M}|$. 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 $\mathcal{K}$ such that $\bigcup_{i=1}^{n_m} m^i = \mathcal{K}$. The cardinality of each mode $m \in \mathcal{M}$, 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 $m \in \mathcal{M}$, the subspace of $\mathcal{R}^{n_k}$ that is accessible corresponds to $\mathcal{R}^{|m|}$ 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. This constrained set of control commands available in mode m, denoted by $\mathcal{U}^m$, can be written as

do they need to be orthonormal?

$$\mathcal{U}^m = \{\boldsymbol{u} | \boldsymbol{u} \in \mathcal{R}^{|m|} \; and \; \|\boldsymbol{u}\|_\infty \leq v_{lim}\}$$

where v_{lim} denotes the maximum velocity along any dimension k and $\|\cdot\|_\infty$ denotes the L_∞ norm.

IV. Intent Disambiguation

why is this counterfactual?

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

goals. We utilize.

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 counterfactually reason about the disambiguation capabilities.

over goals.

We treat the unknown ~~goal g as a~~ random variable and maintain a probability distribution $p(t)$. 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

$p(g^t)$

undefined

$$b_g^t = p(g^t \mid \boldsymbol{u}_h^{0:t}) \propto p(\boldsymbol{u}_h^t \mid g^t, \boldsymbol{u}_h^{0:t-1})p(g^t|\boldsymbol{u}_h^{0:t-1}) \quad (1)$$

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

$$b_g^t = p(\boldsymbol{u}_h^t \mid g^t) \sum_{g_\lambda^{t-1} \in \mathcal{G}} p(g^t, g_\lambda^{t-1} \mid \boldsymbol{u}_h^{0:t-1}) \quad (2)$$

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

$$\cancel{b_g^t = p(\boldsymbol{u}_h^t \mid g^t) \sum_{g_\lambda^{t-1} \in \mathcal{G}} p(g^t|g_\lambda^{t-1})p(g_\lambda^{t-1}|\boldsymbol{u}_h^{0:t-1})}$$
$$= p(\boldsymbol{u}_h^t \mid g^t) \sum_{g_\lambda^{t-1} \in \mathcal{G}} p(g^t|g_\lambda^{t-1})b_{g^{t-1}}^{t-1} \quad (3)$$

or remove Eqn 2, & explain the marginalization together

Marginalizing over m, the control mode variable, the human policy can be written as

$$p(\boldsymbol{u}_h^t \mid g^t) = \sum_{m \in \mathcal{M}} p(\boldsymbol{u}_h^t \mid g^t, m)p(m) \quad (4)$$

We can simplify Equation 4 using the critical piece of information that due to constraints of the control interfaces 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

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

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

note this is limitation in discussion

$$b_g^t = \underbrace{p(\boldsymbol{u}_h^t \mid g^t, m^t)}_{\substack{\text{human's policy conditioned} \\ \text{on mode}}} \underbrace{b_{g_{t-1}}^{t-1}}_{\substack{\text{belief at} \\ t-1}} \quad (6)$$

From Equation 6 we can see that the belief at time t depends on the current mode. 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.

(E) Our aim is to develop a metric that will capture the "disambiguation capability" of a control dimension/mode. Subsequent user operation of the robot in the control mode with maximum disambiguation capability will likely help the autonomy system to perform better intent inference and ~~may result in the autonomy~~ providing more appropriate assistance.

(A) 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 schemes and the various task relevant features and parameters that contribute to it.

(D) From Equation 6 we can see that the time evolution of the probability distribution is sensitive to the user control command (denoted by $\boldsymbol{u}_h$) and that it will evolve differently as the user controls the robot in different control modes. ~~User-generated motion along certain control dimensions/modes therefore might reveal the human's intent with less ambiguity to the autonomy.~~ By utilizing a model-based approach the autonomy can ~~then~~ counterfactually reason about how the belief over goals will evolve in time for all control modes $m \in \mathcal{M}$ and then select the ~~control~~ 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. ~~Various types of intent inference schemes used in conjunction with the~~

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

B. Algorithm: Information Theoretic Intent Disambiguation

(E) (D) PB PC

proposed disambiguation algorithm in our paper are described in detail in ~~Section IV-F.~~

C B. Model-based Projection of $\boldsymbol{p}(t)$ $p(g_t)$?

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 $\boldsymbol{p}(t)$ 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 Boltzmann rationality assumption is modeled as a von Mises Fisher distribution. That is,

citations

$$\boldsymbol{p}(\boldsymbol{u}_h \mid g) = C_p(\kappa)e^{(\kappa \cdot Q_g(\boldsymbol{u}_h, \boldsymbol{x}_g, \boldsymbol{x}_r))} \tag{7}$$

why chose this?

where $C_p(\kappa)$ is the normalization constant, κ is the concentration parameter, p is the dimensionality of the control space and $Q_g(\boldsymbol{u}_h, \boldsymbol{x}_g, \boldsymbol{x}_r)$ is modeled as the cost of taking action $\boldsymbol{u}_h$ towards goal g at robot configuration $\boldsymbol{x}_r$ when acting optimally. We compute this cost as the *agreement* between the human control command $\boldsymbol{u}_h$ and the vector connecting the current robot configuration $\boldsymbol{x}_r$ and the goal configuration $\boldsymbol{x}_g$. The human policy conditioned on control mode m, $p(\boldsymbol{u}_h \mid g, m)$, is then computed as the projection of $p(\boldsymbol{u}_h \mid g)$ 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 $g \in \mathcal{G}$. 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. The full algorithm is presented in Algorithm 1. We use two different information theoretic measures ~~to capture the disambiguation capability of a control mode~~: (1) information content in the projected probability distributions over goal as encoded by the *entropy* of the distribution 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. (Sec. IV.E)

(B)

C. Entropy Disambiguation Metric (Sec. IV.D)

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, $\boldsymbol{p}(t_b)$, can be used as a measure of how confident the system is in its prediction of human intent. That is, entropy can be used a measure of disambiguation. 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

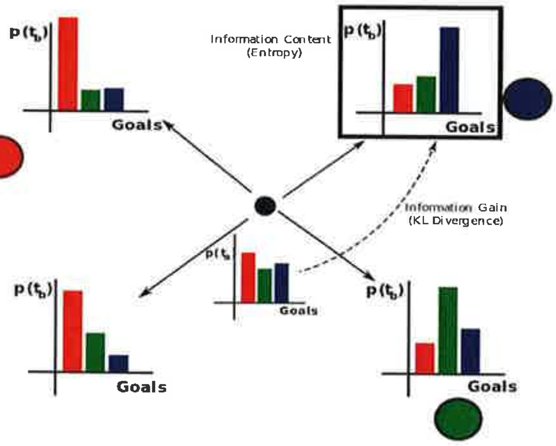

Fig. 1. Illustration of computation of D_m in $\mathbb{R}^2$. The goals are shown in red, green and blue and the robot in black. The probability distributions are forward projected for all $\boldsymbol{u}_h \in Vert(\mathcal{U}^m)$. The change in the overall shape of the probability distribution upon evolving from $\boldsymbol{p}(t_a)$ to $\boldsymbol{p}(t_b)$ amounts to the information *gain* and is captured by the *KL divergence*. The entropy of $\boldsymbol{p}(t_b)$ captures the information *content* in the projected distribution.

by computing a weighted average of the entropies of projected probability distributions. That is,

$$D_m = \frac{1}{|n_g|}\sum_{g \in \mathcal{G}} \mathbb{E}_{\boldsymbol{u}_h \sim p(\boldsymbol{u}_h \mid g, m)}\left[ENT(b_g^{t_b})\right] \tag{8}$$

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. $(p(u_h \mid m))$

D. 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 then be used to characterize the *information gain* regarding the human's intended goal as a result of the application of $\boldsymbol{u}_h$. 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

We first perform a model-based projection of the probability distribution over goals (Lines 5-9, Sec IV.C). We then compute the disambiguation capability D_m of a control mode using information theoretic measures (Lines 10-11).

Algorithm 1 ~~Calculate $p(t_b)$, D_m~~ Info theoretic Intent Disambig

Require: $b_g^{t_a}, \boldsymbol{x}_r^{t_a}, \Delta t, t_a < t_b$

1: **for** $m \in \mathcal{M}$ **do**
2:   Initialize $D_m = 0$
3:   **for** $g \in \mathcal{G}$ **do**
4:     Initialize $t = t_a, \boldsymbol{x}_r^t = \boldsymbol{x}_r^{t_a}, b_g^t = b_g^{t_a}$
5:     **while** $t \le t_b$ **do**
6:       $\boldsymbol{u}_h^t \sim p(\boldsymbol{u}_h \mid g, m)$
7:       $b_g^{t+\Delta t} \leftarrow \text{BeliefUpdate}(b_g^t, \boldsymbol{u}_h^t)$
8:       $\boldsymbol{x}_r^{t+\Delta t} \leftarrow \text{SimulateKinematics}(\boldsymbol{x}_r^t, \boldsymbol{u}_h^t)$
$t \leftarrow t + \Delta t$
9:     **end while**
10:     $D_m \leftarrow D_m + \frac{1}{n_g}\mathbb{E}_{\boldsymbol{u}_h \sim p(\boldsymbol{u}_h \mid g, m)}\Big[ENT(b_g^{t_b})\Big]$
    or
11:     $D_m \leftarrow D_m + \frac{1}{n_g}\mathbb{E}_{\boldsymbol{u}_h \sim p(\boldsymbol{u}_h \mid g, m)}\Big[KL(b_g^{t_b} || b_g^{t_a})\Big]$
12:   **end for**
13: **end for**

gain for all projections of probability distributions $b_g^{t_b}$ ~~That is~~ the disambiguation metric thus can be computed as

$$D_m = \frac{1}{|n_g|}\sum_{g \in \mathcal{G}} \mathbb{E}_{\boldsymbol{u}_h \sim p(\boldsymbol{u}_h \mid g, m)}\Big[KL(b_g^{t_b} || b_g^{t_a})\Big] \tag{9}$$

C

The control mode with highest disambiguation capability m^* then is 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.

I. ~~E.~~ *Intent inference* Section (not subsection)

The exact update rule for the recursive belief updates is determined by the choice of inference schemes (Algorithm 1, Line 6) which impacts ~~as a result of which~~ the projected probability distributions, $b_g^{t_b}$ and the disambiguation metric, D_m. ~~likely are going to be different for different choices. This section describes the different types of intent inference schemes that work in conjunction with the disambiguation algorithms proposed in Section IV.~~

F

1) Heuristic Approaches: Heuristic approaches based on *confidence* functions [7] seek to find direct mappings between instantaneous cues and underlying human intentions. For every goal $g \in \mathcal{G}$, the system maintains an associated set of confidences denoted by $\mathcal{C}$. The system designer has the freedom to choose the set of features that will inform the confidence functions. For example, a simple proximity-based confidence function used extensively in literature is $c \in \mathcal{C}$

$$c(\boldsymbol{x}_r, \boldsymbol{x}_g) = max\Big(0, 1 - \frac{||\boldsymbol{x}_r - \boldsymbol{x}_g||}{r}\Big)$$

where $\boldsymbol{x}_r$ is the robot position, $\boldsymbol{x}_g$ is the goal position, r is the radius beyond which the confidence function is always 0 and $||\cdot||$ is an appropriate distance metric. A slightly more information-rich variant that aims to capture the "directedness" of the human control command to a particular goal position is

$$c(\boldsymbol{x}_r, \boldsymbol{x}_g, \boldsymbol{u}_h) = \boldsymbol{u}_h \cdot (\boldsymbol{x}_g - \boldsymbol{x}_r)$$

where $\boldsymbol{u}_h$ is the human control command. These confidence functions rely on instantaneous features and therefore are amnesic and can exhibit chatter behavior [6]. Using either of the above-mentioned confidence functions, we have

$$p_g^t = c(\boldsymbol{x}_r^t, \boldsymbol{x}_g^t, \boldsymbol{u}_h^t) \tag{10}$$

~~where~~ $\boldsymbol{u}_h^t \sim p(\boldsymbol{u}_h \mid g, m)$.

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 equation for the probability distribution over goals directly uses Equation 6 for belief propagation.

3) Dynamic Neural Field Approaches: Dynamic neural fields are differential equations ~~in time~~ that governs the time evolution of dynamical state variables, with some specific properties such as recurrent interactions between the state variables, robustness to external noise and memory. Dynamic neural fields were originally conceived to explain cortical population neuronal dynamics, based on the hypothesis that the excitatory and inhibitory neural interactions between local neuronal pools form the basis of cortical information processing [21]. 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 $\Sigma_1^{n_g} p^i(t) = 1$ [17].

The full specification of the neural field is given by

$$\frac{\partial \boldsymbol{p}(t)}{\partial t} = \frac{1}{\tau}\Big[-\mathbb{I}_{n_g \times n_g} \cdot \boldsymbol{p}(t) + \underbrace{\frac{1}{n_g} \cdot \mathbb{1}_{n_g}}_{\text{rest state}} \Big] + \underbrace{\boldsymbol{\lambda}_{n_g \times n_g} \cdot \sigma(\boldsymbol{\xi}(\boldsymbol{u}_h; \boldsymbol{\Theta}))}_{\text{excitatory + inhibitory}} \tag{11}$$

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, $\boldsymbol{\lambda}$ is the control matrix that controls the excitatory and inhibitory aspects, $\boldsymbol{\xi}$ is a function that encodes the nonlinearity through which human control commands and task features affect the time evolution, $\boldsymbol{\Theta}$ represents all other task-relevant features and parameters that affect the time-evolution of the probability distribution, and σ is a biased sigmoidal nonlinearity given by $\sigma(\boldsymbol{\xi}) = \frac{1}{1+e^{-\boldsymbol{\xi}}} - 0.5$. Given the initial probability distribution at time t_a, Equation 11 can be solved numerically from $t \in [t_a, t_b]$ using a simple Euler algorithm with a fixed time-step Δt. The design of $\boldsymbol{\xi}$ is informed by what features of the human control input and environment capture the human's underlying intent most effectively. ~~We rely~~ 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.

A More specifically

F Here we present the multiple intent inference schemes evaluated in conjunction with our disambiguation algorithm.

peppers	rice	Fm-hach	chili
onions		tortillas	
		beans	
cilantro	cheese		
mush	butter	bread?	
jam			

Information Theoretic Characterization of Transparency in Human-Robot Interaction using Multivariate Transfer Entropy

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

***Abstract*— Understanding the dynamics of information flow and exchange between humans and robots is critical for the quantification and design of seamless, fluid and intuitive Human-Robot Interaction (HRI) paradigms. Information theoretic characterization can lead to novel approaches to the design of robot behaviors that can likely optimize subjective aspects such as user satisfaction and acceptance as well as objective task performance related metrics such as success rate and effort simultaneously. In this paper, we characterize and quantify the *transparency* of human-robot interaction in a shared-control assistive robotic manipulation setting using the information theoretic concept of *multivariate transfer entropy.* Specifically, we model HRI as a *perception-action* loop that unfolds in time utilizing the mathematical framework of Bayesian Networks and propose that transparency can be quantified as the information flow between the nodes of the relevant autonomy and human variables in the network. We present results from a pilot study in which we modulated robot-to-human transparency and our results indicate that a multivariate transfer entropy based metric to compute information flow can be used to effectively capture the notion of transparency.**

I. INTRODUCTION

Transparency in the context of human-robot interaction is of paramount importance for seamless and fluid human-robot performance. Transparency in the context of shared autonomy can refer to a few different aspects of human-robot interaction, all of which can play a significant role in determining the quality of interaction. For example, in a human-robot team, transparency can be thought of as the *observability* and *predictability* of either agent's behavior; it can also refer to the effective and legible bi-directional communication of internal desires, goals and intentions which in turn facilitates and promotes cooperation and coordinated task execution. Transparency also plays a significant role in user-centered design by having the user aware of the state of the autonomy at all times. Transparency-based metrics the can ~~possibly~~ provide a ~~more~~ high-level approach in which design of robot autonomy is guided by modulating the information flow between the autonomy and the human. As opposed to the typical approach to the design of robot autonomy that relies on optimization of robot behaviors with respect to predefined set of performance related reward functions [1], the ~~hypothesis in the above mentioned is that~~ optimization of information flow might ~~can~~ provide a ~~more~~ high–level, task and platform agnostic metric that ~~will~~ lead to enhanced cooperation and mutual understanding, as a result ~~of which~~ the desired outcomes of improved ~~(better~~ task performance, and ~~improved~~ user satisfaction) will likely naturally emerge. In general, it is widely agreed upon in literature that transparency is a critical factor to improve system performance, reduce human errors due to conflict, and helps to build mutual trust in human-robot teams. Our hypothesis is that A

Information processing is deemed to be a necessary requirement for life; a biological agent needs to first acquire relevant information from the environment before taking any actions [2]. Drawing inspiration from how biological agents interact with their environment and each other via information exchange, we posit that an information theoretic characterization of HRI can offer a novel way of thinking about the design of robot behaviors. To that end, in this work we posit that HRI in the context of shared autonomy can be modeled as a *coupled perception-action loop* that unrolls in time. We utilize the mathematical framework of *Causal Bayesian Networks* to model the perception action loop in which the nodes of the network represent all the relevant variables that are modeled (such as human and autonomy control actions, robot state, environment state, *et cetera*).

In 1948 Claude Shannon originally proposed *information theory* as a mathematical theory of communication to primarily understand the fundamental limitations of a communication channel and what constitutes optimal data encoding [3]. Shannon's work introduced a mathematically sound and meaningful approach to quantify and think about communication (or flow) of information. Shannon's original theory contains all the essential ingredients that are necessary to quantify the information dynamics between various interacting components of systems that exhibit complex dynamics such as a human robot team. The mathematical formulation of information theory is purely based on the *statistics* of random variables and therefore, provides a parameter-free mathematical framework to understand the information flow and exchange between dynamical systems and has been widely used in diverse domains such as economics [4], neuroscience [5], weather forecasting [6] and even animal-robot interactions [7]. Unsurprisingly, information theoretic measures have been widely used to quantify information flow between nodes in a Bayesian network to reveal correlational as well as directed causal influence between components of

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. Deepak Gopinath would also like to thank Dr. Vishnu Sreekumar for initial conversations on transfer entropy.

[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

a complex systemAdd Polani Ay reference.

The contributions of this paper are three-fold:

Science Robotics? ...stronger results

1) We frame HRI in the context of shared autonomy as a coupled perception action loop utilizing the framework of causal Bayesian Networks.
2) We quantify the colloquial notion of transparency in concrete mathematical terms utilizing the concept of information flow as measured by multivariate transfer entropy that purely relies on the statistics of the time-series data recorded during task execution.
3) We perform and present results from a validation experiment in the domain of assistive robotic manipulation in which the above-mentioned metric is used to characterize robot-to-human transparency. Preliminary results indicate a high correlation between the computed metrics and subjective evaluation of transparency.

In Section II we provide a brief overview of different notions of transparency within HRI and, information theoretic concepts as applied to the quantification of information flow in various types of dynamical systems. In Section III-A we describe HRI as a coupled perception action loop and Section III introduces the information theoretic quantities that are essential for understanding information flow/transfer in a Bayesian Network that unrolls in time. Section IV describes the experimental protocol and evaluation followed by results in Section V. Discussion and conclusion are presented in Section VI and VII respectively.

II. RELATED WORK

To add transparency related work.

Schreiber originally proposed transfer entropy as an information theoretic measure that quantifies the directed statistical coherence between dynamical systems that evolve in time [10]. Kaiser et. al also proposed the continuous domain analogue of discrete-domain transfer entropy [11]. In practice, the probability density functions have to be estimated from finite number of data samples. Researchers have developed kernel-based and nearest-neighbor based methods in order to significantly reduce the estimation bias and variance [12].

Ever since its introduction, transfer entropy has found widespread use in diverse domains such as economics [13], neuroscience [14], and to a limited extent in human-robot interaction [15]. Temporal relationships between global financial markets and stock market indices have been studied using transfer entropy and its variants [16]. Information theoretic measures have also been successful in effective network inference in the domains of computational neuroscience as applied to EEG [17], fMRI [18] and spiking neuronal data [19] as well other fields such as supply-chain networks [20]. The potential of transfer entropy to provide key insights regarding the temporal dynamics of biochemical networks have also been recognized [21]. Researchers have also attempted to model joint attention mechanism in a human-robot team using transfer entropy [22]. The importance of information theoretic principles in the context of biological systems, was recognized by Polani when he put forth the notion of information as the fundamental *currency* responsible for the success of a living organism [23]. We take inspiration from this concept and hypothesize that information can also possibly serve as a fundamental currency with which human-robot interaction be quantified and defined.

Fig. 1. Assistive robotic manipulation tasks.

Most information theoretic measures are global metrics averaged over all possible state configurations due to which the local spatiotemporal dynamics of information is ignored. Lizier et. al proposed an information theoretic framework to quantify information storage, transfer and modification at a local spatiotemporal scale [24]. This treatment of information dynamics provides a strong connection to other domains such as a dynamical systems theory and nonlinear time series analysis. We utilize this framework in our work to capture the information dynamics within a human-robot team. Knowledge of information dynamics can provide insight into how to shape the dynamics using control theoretic principles to achieve specific goals.

III. MATHEMATICAL FRAMEWORK

In this section, we present the framework for describing HRI as a coupled perception action loop and a primer on essential information theoretic metrics used for quantifying information flow in a causal Bayesian Network.

A. Causal Bayesian Networks for HRI

In this section, I present causal Bayesian Network (CBN) based mathematical framework for information theoretic analysis of human-robot interaction in the context of shared autonomy. The interaction dynamics and information exchange between the human and the autonomous partner is modeled as a coupled perception action loop. The nodes of the CBN represent the relevant variables pertaining to both human and autonomy (latent and observed) that are relevant for analysis and the edges represent the probabilistic influence they have on each other. The notion of perception-action loops have already been widely used to describe the interaction of an embodied agent (both biological and artificial) and its environment. Perception-action cycle is considered to be the fundamental logic of the central nervous system in which the perception and action processes are closely interlinked. Perception leads to action and action leads to perception (add references). When there are multiple agents involved the notion of a single perception action loop can be generalized giving rise to coupling of independent perception action loops. Within this coupled system, from the perspective of one agent the other agent(s) is part of its own environment.

In order to ground the framing of HRI using CBNs we focus on the scenario of assistive robotic manipulation in which