Information Theoretic Intent Disambiguation via Contextual Nudges for Assistive Shared Control

Deepak Gopinath[1,3], Andrew Thompson[1,3], and Brenna D. Argall[1,2,3]

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

Abstract. The usefulness of shared-control assistive robots frequently relies on the underlying autonomous agent's ability to infer human intentions unambiguously, often from low-dimensional and noisy signals generated by the human through a control interface. In this paper, we propose a strategy in which the autonomous agent nudges the context in which the human generates their control actions. In doing so, the autonomous agent attempts to improve its own ability to infer intent accurately, which in turn allows it to provide more accurate assistance. The contributions of this paper are three-fold. First, we introduce an interface-aware information-theoretic metric for active disambiguation that aims to characterize world states according to their potential to extract maximally intent-expressive control actions from the user. Second, we propose a turn-taking based human-autonomy interaction protocol in which the autonomous agent utilizes the disambiguation metric to help itself reduce the uncertainty of its prediction of human intent. Third, we evaluate our metric and interaction protocol both in simulation and with a 9-person human subject study. Our results suggest that disambiguation (a) helps to significantly reduce task effort, as measured by number of mode switches, task completion times, and number of turns executed by the human, and (b) enables the autonomous agent to provide accurate assistance with greater contribution to the overall control signal.

Keywords: Shared Autonomy, Assistive Robotics, Intent Disambiguation, Human-Robot Interaction.

1 Introduction

A fundamental challenge in robotics is that of state estimation from noisy sensor data [2]. The primary goal of any state estimation algorithm is to reduce the uncertainty that arises from noisy measurements and inaccurate models, which can become difficult due to limited information channels and associated hardware constraints. In human-autonomy interaction scenarios, usually state estimation performed by the autonomous agent not only involves estimation of the environment state but also of the unobserved latent human state that encodes the human goals, beliefs, and intentions [12]. Particularly within the domain of shared autonomy assistive robots, the effectiveness of the autonomy agent depends on how well it is able to infer the user's intentions *unambiguously* from the control interface signals that are generated by the human. Shared

autonomy for assistive robots holds considerable potential to improve the lives of millions of people with motor impairments [15] for whom manual teleoperation of assistive robots is hard because standard control of these robots is enacted via low-dimensional control interfaces such as sip-and-puffs, switch-based headarrays, and joysticks. Due to the low-dimensionality, these interfaces can only operate in subsets of the entire control space (referred to as *control modes*) [10]. Additionally, inherent mechanical limitations of the interface alongside varying motor skill deficits due to injury or disease can result in sparse, noisy, and low-information signals about the user's intended goal; hence, intent inference based on user control signals is difficult for the autonomous agent.

Intent disambiguation algorithms aim to improve the intent inference capabilities of an autonomous agent by rigorously eliciting more information from constrained sensor channels [7, 8]. In this paper, we frame *intent disambiguation* as a problem of optimally *nudging* the user's environment (decision making context) such that their subsequent control interface actions are guaranteed to result in *maximal information gain* regarding the user's latent intentions. In this work, we also explicitly incorporate the impact of the control interface's inherent noisy characteristics on information gain.

The key contributions of this paper are three-fold:

1. We propose an interface-aware information-theoretic framing of the problem of *intent disambiguation*.
2. We propose a turn-taking based Human-Autonomy Interaction (HAI) protocol in which the autonomous agent utilizes the proposed disambiguation metric to *help itself* when uncertain about its prediction of human intent.
3. We present results from a nine person human subject study that evaluated the effectiveness of the proposed disambiguation metric and turn-taking based protocol.

In Section 2 we present an overview of related research. Section 3 presents our mathematical formalism for the proposed interface-aware information-theoretic disambiguation metric. The turn-taking based HAI protocol is described in Section 4 followed by study design and experimental methods in Section 5. Section 6 presents the results. Discussion and conclusions are presented respectively in Sections 7 and 8.

2 Related Work

Information Gathering in HAI: The aim of eliciting more legible and information-rich control commands from the user to improve intent estimation is closely related to active learning. Designing optimal control laws that maximize information gain can be accomplished by having the associated reward structure reflect some measure of information gain [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 [16]. Probing algorithms likewise have been designed to elicit information-rich signals in a collaborative workspace setting; for example, an autonomous car interacting with a human driver at a traffic intersection [18]. Brooks et al. [3] seek to balance information gathering actions with goal-oriented actions within a shared autonomy context. Their objective is to

identify *autonomous actions* to quickly ascertain the user's goal; whereas, in this paper, we optimize over *future states* from which *human actions* will facilitate information gathering about the latent intentions.

Intent Disambiguation as Nudging: In our algorithm, when the autonomous agent uses the intent disambiguation algorithm to nudge the robot into maximally disambiguating states, which is an example of how a *decision-making* context is altered to indirectly affect the decisions taken by the human. This phenomenon, known as *nudging*, is extensively studied in the sphere of behavioral economics [21], public policy [5], and business marketing [13]. Originally proposed by Thaler and Sunstein [22], nudging is the mechanism by which any aspect of the choice architecture is modified in an attempt to influence a person's behavior in a predictable manner. *Choice architecture* refers to the organization and presentation of the choices that a decision maker has access to. Nudging is not the same as introducing an arbitrary number of constraints, but rather it is an attempt to influence the decision maker's choice without limiting the choice set or making other alternatives more costly. In the domain of robotics, applications of nudge theory are explored in the context of social robotics, particularly with respect to the ethics and morality of nudging humans when they interact with robots [17]. More recently, a computational account of optimal nudging is proposed [4].

In the case of assistive robotic teleoperation with low-dimensional interfaces, the *context* is the state of the environment in which the human is required to generate actions. Altering context amounts to changing some aspect of the environment state. Specifically, the autonomous agent *intervenes* to alter the context in specific ways in order to influence the human to act in a certain manner. In previous work [7] on intent disambiguation, changes to the context were restricted to changes to the active mode; in this paper, the changes affect both the robot location and control mode.

Assistance via Turn Taking: In a shared autonomy system in which both the human and the autonomous agent control the *same physical device* at the *same time*, it can become difficult for the human to isolate the autonomous agent's contribution to the overall control signal as feedback is limited. To estimate the autonomous agent's control contribution requires either a good mental model of the autonomous agent's policy and understanding of the control arbitration scheme, or mentally *subtracting* away the effect of their own issued controls. In turn-taking, the bulk of the autonomous agent's actions are executed during its own turn *without any contribution from the human* and as a result it becomes easier for an observer to infer the latent assistance strategies from *uncorrupted* trajectory rollouts initiated by the autonomous agent.

Turn taking is one of more common HAI protocols used in a variety of robotics subfields. Fluency in HAI based on information flow is studied [19, 23]. Assistive robots utilize turn-taking as an interaction paradigm for therapy of children with autism spectrum disorder [20]. Conversational turn-taking robots that rely on gestures and natural language modalities are extensively used to understand social dynamics of human-robot interaction [14]. An additional motivation for the target domain is that turn-taking could provide a natural framework which affords periods of rest for the human. Taking sufficient rest becomes particularly important (especially with limited interfaces such as a sip-and-puff) as continuous manual teleoperation is physically taxing and can quickly result in fatigue.

3 Mathematical Formalism

In this section, we present the mathematical formulation of our interface-aware information-theoretic intent disambiguation algorithm. A probabilistic graphical model of limited control-interface mediated robot teleoperation is presented in Section 3.1. In Section 3.2 we describe the recursive Bayesian intent estimation algorithm used by the autonomous agent to determine the user's intended goal and in Section 3.3 we present the design of our novel disambiguation metric.

3.1 Modeling Limited Control-Interface Mediated Robot Teleoperation

A model for control-interface mediated robot teleoperation by a human is necessary for both goal inference (Section 3.2) and intent disambiguation (Section 3.3).

Let $\mathcal{K}$ denote the controllable dimensions of the robot and d be the dimensionality of the control interface such that $d < |\mathcal{K}|$. Due to the dimensionality mismatch, $\mathcal{K}$ is partitioned into a set of control modes denoted as $\mathcal{M}$, such that $\bigcup m = \mathcal{K}$, where $m \in \mathcal{M}$. Let $s \in \mathcal{S}$ represent the state of the combined robot-interface system. The state s consists of both the robot state $q \in \mathcal{Q}$ and the current active control mode $m \in \mathcal{M}$; i.e., $s = (q, m)$ with $\mathcal{S} = \mathcal{Q} \times \mathcal{M}$. We also define $\Psi_q(s) = q$ and $\Psi_m(s) = m$ as functions that compute the projections of s on to $\mathcal{Q}$ and $\mathcal{M}$ respectively.

Let $a \in \mathcal{A}$ be the set of all actions available for the robot-interface system. $\mathcal{A}$ can be decomposed as $\mathcal{A}_q \times \mathcal{A}_m$, where $\mathcal{A}_q$ is the set of **control actions** and $\mathcal{A}_m$ is the set of **mode switch actions**. Control actions bring about state transitions in the robot state space $\mathcal{Q}$ and result in *robot motion*. Mode switch actions result in transitions in $\mathcal{M}$ and determines the *active control mode* and consequently the dimension(s) in which robot motion is possible at any given time.

Following the interface-aware model presented in [6], let Φ be the set of all available **interface-level** actions. Different control interfaces have distinct sets of interface-level actions that depend on the unique physical activations necessary to operate the interface. The user's *intended* interface-level action, determined by the task-level action a, is denoted as ϕ_i. The *measured* interface-level action, ϕ_m, can be different from the *intended* interface-level action, ϕ_i, because of various factors such as motor noise, electromechanical wear and tear in the interface, and also the user's ability to execute ϕ_i. Due to the lower dimensionality of the interface, the subset of actions available in any given state $s \in \mathcal{S}$ is determined by the mode-switching paradigm Δ used by the interface. Finally, ϕ_m is mapped onto low-level control commands (denoted as u_h) issued to the robot-interface system via a deterministic transformation function given by Υ.

Fig. 1 depicts the probabilistic graphical model of control interface mediated robot teleoperation by a goal-directed human.

We model the human teleoperating the robot towards a goal $g \in \mathcal{G} \subset \mathcal{Q}$ using a limited control interface as an **interface-dependent goal-directed Markov Decision Process** (MDP) denoted by the tuple $(\mathcal{S}, \mathcal{A}, \mathcal{T}, \mathcal{R}_g, \gamma, \rho_0, \Delta)$, where $\mathcal{R}_g : \mathcal{S} \times \mathcal{A} \times \mathcal{S} \to \mathbb{R}$ is the goal-dependent reward, $\gamma \in [0, 1)$ is the discount factor, Δ is an interface-dependent parameter that determines the mode-switching paradigm, and ρ_0 is the initial state distribution. We model the robot-interface system as a deterministic hybrid dynamical system with a transition function $\mathcal{T} : \mathcal{S} \times \mathcal{A} \to \mathcal{S}$ consisting of two parts:

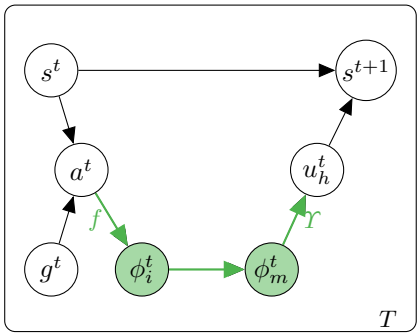


Fig. 1: Probabilistic graphical model depicting a specific user's goal directed robot teleoperation using a control interface at single time step t. The nodes and edges that model the physical aspect of controlling the interface are highlighted in green. In this graphical depiction s^t subsumes both q^t (the robot state) as well as m^t (the control mode). ϕ_i^t and ϕ_m^t are the *intended* and *measured* interface-level physical actions respectively. These variables encode the specific physical activation mechanisms needed to generate a signal using an interface.

$\mathcal{T}_q : \mathcal{Q} \times \mathcal{A}_q \rightarrow \mathcal{Q}$, which determines how the task-level control actions result in *motion* or equivalently in changes to the robot location, and $\mathcal{T}_m : \mathcal{M} \times \mathcal{A}_m \rightarrow \mathcal{M}$, which determines how control-level mode switch actions determine the current active mode.

We solve for the goal-dependent optimal policy (which is a mapping from state s to a distribution over task-level actions a), denoted as π^g, using standard value iteration and treat the goal-dependent stochastic human policy, $p(a|s, g)$, as an ϵ-greedy policy that can be written as

$$p(a|s, g) = (1 - \epsilon) \cdot \pi^g + \epsilon \cdot \pi_{unif} \tag{1}$$

where π_{unif} is a uniform policy and $\epsilon \in (0, 1)$.

In Section 4, we elaborate on how the low-level control commands u_h^t generated by the human are combined with autonomous control commands u_a^t within a linear-blending based shared-control assistance system.

3.2 Recursive Bayesian Goal Inference

At any time t, the human's true intended goal is latent and unobservable to the autonomous agent and as such it maintains a belief over goals denoted as b_g^t. The process of goal inference amounts to computing the posterior over goals given the history of measured interface-level actions, $\phi_m^{0:t}$ and states $s^{0:t}$. More precisely, we can use Bayesian inference to compute the posterior $b_g^t = p(g^t|\phi_m^{0:t}, s^{0:t})$ as

$$b_g^t \propto p(g^t|\phi_m^{0:t-1}, s^{0:t}) p(\phi_m^t|g^t, \phi_m^{0:t-1}, s^{0:t}) \tag{2}$$

and using the conditional independence assumptions encoded in the model shown in Fig. 1 we can remove the dependence of $p(\phi_m^t|g^t, \phi_m^{0:t-1}, s^{0:t})$ in Eq. 2 on $\phi_m^{0:t-1}$ and

$s^{0:t-1}$. Then Eq. 2 becomes,

$$b_g^t \propto p(g^t|\phi_m^{0:t-1}, s^{0:t})p(\phi_m^t|g^t, s^t). \tag{3}$$

Marginalizing over a^t and ϕ_i^t we can express Eq. 3 as

$$b_g^t = \eta \cdot p(g^t|\phi_m^{0:t-1}, s^{0:t}) \sum_{a^t \in \mathcal{A}} \sum_{\phi_i^t \in \Phi} p(\phi_m^t|\phi_i^t)p(\phi_i^t|a^t)p(a^t|s^t, g^t) \tag{4}$$

with η as the normalization factor. $p(\phi_i^t|a^t)$ captures the *user's internal model of the true mapping* (denoted as f) between task-level actions and interface-level actions. $p(\phi_m^t|\phi_i^t)$ is the *user input distortion model*, which captures the stochastic deviations of the *measured* interface-level actions from the *intended* interface-level actions.

$p(g^t|\phi_m^{0:t-1}, s^{0:t})$ can be recast as,

$$\begin{aligned} p(g^t|\phi_m^{0:t-1}, s^{0:t}) &= \sum_{g^{t-1} \in \mathcal{G}} p(g^t, g^{t-1}|\phi_m^{0:t-1}, s^{0:t}) \\ &= \sum_{g^{t-1} \in \mathcal{G}} p(g^t|g^{t-1}, \phi_m^{0:t-1}, s^{0:t})p(g^{t-1}|\phi_m^{0:t-1}, s^{0:t}) \\ &= \sum_{g^{t-1} \in \mathcal{G}} p(g^t|g^{t-1}) \underbrace{p(g^{t-1}|\phi_m^{0:t-1}, s^{0:t-1})}_{b_g^{t-1}} \end{aligned} \tag{5}$$

under the assumption that g^t only depends on g^{t-1} and that the state at time t does not influence the goal at time $t-1$. Combining Eq. 5 with Eq. 4 the recursive Bayesian update for goal inference is given by

$$b_g^t = \eta \left[\sum_{g^{t-1} \in \mathcal{G}} b_g^{t-1} \cdot p(g^t|g^{t-1}) \right] \sum_{a^t \in \mathcal{A}} \sum_{\phi_i^t \in \Phi} p(\phi_m^t|\phi_i^t)p(\phi_i^t|a^t)p(a^t|s^t, g^t). \tag{6}$$

Under the assumption that the goal transition probability is a delta distribution, the above equation can be further simplified as

$$b_g^t = \eta \cdot b_g^{t-1} \sum_{a^t \in \mathcal{A}} \sum_{\phi_i^t \in \Phi} p(\phi_m^t|\phi_i^t)p(\phi_i^t|a^t)p(a^t|s^t, g^t). \tag{7}$$

3.3 Disambiguation Metric

We formalize intent disambiguation as a characterization of the states in the state space $\mathcal{S}$ according to their potential to *extract* interface signals ϕ_m from the user that provide the most information about the latent goal $g \in \mathcal{G}$. Intent disambiguation is particularly useful when the measured control signals ϕ_m are very noisy and sparse. An autonomous agent that uses the intent disambiguation algorithm can nudge the robot into *maximally disambiguating* state(s), such that *subsequent* actions executed by the user will help the autonomous agent to infer human intent more accurately.

In order to perform intent disambiguation, the autonomous agent needs to have a notion of the *amount of information contained* in ϕ_m^t *about* g^t, conditioned on the current state and the history of states and actions. To be precise, at any given time t, if $\phi_m^{0:t-1}$ and $s^{0:t-1}$ represent the history of interface-level actions and states that the autonomous agent have observed, then the **conditional mutual information** $I(\phi_m^t; g^t | \phi_m^{0:t-1}, s^{0:t-1}, s^t)$ between ϕ_m^t and g^t, conditioned on $\phi_m^{0:t-1}$ and $s^{0:t-1}$ and the current state s^t, measures the amount of information obtained about g^t by observing ϕ_m^t. Using the standard definition of conditional mutual information we then have

$$I(\phi_m^t; g^t | \phi_m^{0:t-1}, s^{0:t-1}, s^t) = \sum_{\phi_m^t \in \Phi} \sum_{g^t \in \mathcal{G}} p(\phi_m^t, g^t | \phi_m^{0:t-1}, s^{0:t-1}, s^t) \log \frac{p(\phi_m^t | g^t, \phi_m^{0:t-1}, s^{0:t-1}, s^t)}{p(\phi_m^t | \phi_m^{0:t-1}, s^{0:t-1}, s^t)}. \quad (8)$$

The term $p(\phi_m^t, g^t | \phi_m^{0:t-1}, s^{0:t-1}, s^t)$ can be rewritten as

$$p(\phi_m^t, g^t | \phi_m^{0:t-1}, s^{0:t-1}, s^t) = p(g^t | \phi_m^{0:t-1}, s^{0:t-1}, s^t) p(\phi_m^t | g^t, \phi_m^{0:t-1}, s^{0:t-1}, s^t). \quad (9)$$

By using the conditional independence assumptions encoded in the model, we can further simplify $p(\phi_m^t | g^t, \phi_m^{0:t-1}, s^{0:t-1}, s^t)$ to $p(\phi_m^t | g^t, s^t)$ and $p(\phi_m^t | \phi_m^{0:t-1}, s^{0:t-1}, s^t)$ as $p(\phi_m^t | s^t)$ by removing the dependence on $\phi_m^{0:t-1}, s^{0:t-1}$. Therefore, Eq. 8 becomes

$$I(\phi_m^t; g^t | \phi_m^{0:t-1}, s^{0:t-1}, s^t) = \sum_{g^t \in \mathcal{G}} p(g^t | \phi_m^{0:t-1}, s^{0:t-1}, s^t) \sum_{\phi_m^t \in \Phi} p(\phi_m^t | g^t, s^t) \log \frac{p(\phi_m^t | g^t, s^t)}{p(\phi_m^t | s^t)} \quad (10)$$

Marginalizing over g^{t-1} the first term on the right hand side of Eq. 10 can be expressed as

$$\begin{aligned} \sum_{g^t \in \mathcal{G}} p(g^t | \phi_m^{0:t-1}, s^{0:t-1}, s^t) &= \sum_{g^t \in \mathcal{G}} \sum_{g^{t-1} \in \mathcal{G}} p(g^t, g^{t-1} | \phi_m^{0:t-1}, s^{0:t-1}, s^t) \\ &= \sum_{g^t \in \mathcal{G}} \sum_{g^{t-1} \in \mathcal{G}} p(g^t | g^{t-1}, \phi_m^{0:t-1}, s^{0:t-1}, s^t) p(g^{t-1} | \phi_m^{0:t-1}, s^{0:t-1}, s^t) \quad (11) \end{aligned}$$

and under the assumptions that the goal transition probability only depends on the previous goal, that the state at time t does not have an influence on the belief over g^{t-1} (violates causality), Eq. 11 can be simplified as

$$\sum_{g^t \in \mathcal{G}} p(g^t | \phi_m^{0:t-1}, s^{0:t-1}, s^t) = \sum_{g^{t-1} \in \mathcal{G}} p(g^{t-1} | \phi_m^{0:t-1}, s^{0:t-1}) \sum_{g^t \in \mathcal{G}} p(g^t | g^{t-1}). \quad (12)$$

By combining Eq. 12 with Eq. 10 we finally have

$$I(\phi_m^t; g^t | \phi_m^{0:t-1}, s^{0:t-1}, s^t) =$$
$$\sum_{g^{t-1} \in \mathcal{G}} p(g^{t-1} | \phi_m^{0:t-1}, s^{0:t-1}) \sum_{g^t \in \mathcal{G}} p(g^t | g^{t-1}) \sum_{\phi_m^t \in \Phi} p(\phi_m^t | g^t, s^t) \log \frac{p(\phi_m^t | g^t, s^t)}{p(\phi_m^t | s^t)}$$

$$= \sum_{g^{t-1} \in \mathcal{G}} b_g^{t-1} \sum_{g^t \in \mathcal{G}} p(g^t | g^{t-1}) \sum_{\phi_m^t \in \Phi} p(\phi_m^t | g^t, s^t) \log \frac{p(\phi_m^t | g^t, s^t)}{p(\phi_m^t | s^t)}. \quad (13)$$

Note that the quantity on the right hand side of Eq. 13 is the expectation of the Kullback-Leibler divergence, D_{KL}, between $p(\phi_m^t | g^t, s^t)$ and $p(\phi_m^t | s^t)$ and therefore Eq. 13 is equivalent to

$$I(\phi_m^t; g^t | \phi_m^{0:t-1}, s^{0:t-1}, s^t) = \mathbb{E}_{g^{t-1} \sim b_g^{t-1}, g^t \sim p(g^t | g^{t-1})} D_{\text{KL}} \left[p(\phi_m^t | g^t, s^t) \middle| p(\phi_m^t | s^t) \right] \quad (14)$$

and can be estimated using Monte Carlo techniques by generating samples according to the generative model in Fig. 1.

We define the full disambiguation metric $D : \mathcal{S} \to \mathbb{R}$ as

$$D(s) = I(\phi_m^t; g^t | \phi_m^{0:t-1}, s^{0:t-1}, s) - \lambda \sum_{g^{t-1} \in \mathcal{G}} b_g^{t-1} \cdot \left\| g^{t-1} - \Psi_q(s) \right\|. \quad (15)$$

The first term on the right-hand side is the conditional mutual information described earlier and the second term can be interpreted as a regularization term with λ being the regularization coefficient. The regularization term helps the optimizer to navigate an ill-defined optimization landscape which can occur if the mutual information term is identical for all states in the optimization domain. In our implementation, $\mathcal{Q}$ is the space of robot position; accordingly, the regularization term $\left\| g^{t-1} - \Psi_q(s) \right\|$ is the distance from the robot position q to the goal position g^{t-1}. Hence, the optimizer favors disambiguating states that are closer to the goal region. Note that $D(s)$ is a conditional metric that depends on the history of interface actions and states.

The maximally disambiguating state s^* is the optimizer of Eq.15 and is given by

$$s^* = \underset{s \in \mathcal{S}}{\operatorname{argmax}} D(s) \quad (16)$$

Simulation-based Validation of $D(s)$: We compared how well $D(s)$ was able to match an intuitive ground truth for what a maximally good disambiguating state should be. For a local neighborhood of each $s \in \mathcal{S}$, we computed s^* using Eq. 16 and compared it to a ground truth that was computed by picking the state in the same neighborhood that provided maximum expected difference between first and second maxima of b_g^t over a single timestep. We ran a total of 100 simulations for $|\mathcal{G}|$ ranging from 3 to 30 and $|\mathcal{S}| = 200$, and the match percentage was 100%.

Algorithm 1 Turn-Taking Interaction with Active Intent Disambiguation

1: **if** human-turn **then**
2: $a_h^t \sim p(a|s^t, g^t)$ ▷ task-level action [human]
3: $\phi_i^t \sim p(\phi_i^t|a^t)$ ▷ intended interface action [human]
4: $\phi_m^t \sim p(\phi_m^t|\phi_i^t)$ ▷ measured interface action [human]
5: $u_h^t = \Upsilon(\phi_m^t)$ ▷ control command corresponding to ϕ_m^t
6: Update b_g^t using Eq. 4 ▷ Bayesian belief update
7: $u_a^t = \Xi(\text{argmax}_g\, b_g^t)$ ▷ control signal to achieve inferred goal, g'
8: $u_f^t = \alpha \cdot u_a^t + (1-\alpha) \cdot u_h^t$ ▷ shared autonomy via control blending
9: $s^{t+1} \sim \mathcal{T}^u(s^t, u_f^t)$ ▷ state transition using blended control signal
10: **if** autonomy-turn **then**
11: **if** $H(b_g^t) > \kappa$ **then** ▷ if not confident, nudge into disambiguating state
12: Compute s^* using Eq. 16
13: $s^{t+1} = s^*$
14: **else** ▷ if confident, nudge towards inferred goal
15: $s^{t+1} = (\beta \cdot g' + (1-\beta) \cdot \Psi_q(s^t), \Psi_m(s^t))$

4 Shared Control via Contextual Nudges

We propose a turn-taking based interaction protocol outlined in Algorithm 1. The general idea is that at any time t, after having observed $\phi_m^{0:t-1}$ and $s^{0:t-1}$, the autonomous agent can choose to nudge the robot into s^*, from which subsequent actions executed by the human will extract maximum information regarding g^t. By doing so, the agent implicitly *helps itself* to provide accurate assistance in the future.

Task execution begins with the human. Lines 2-4 show how a human generates task-level actions (a_h^t) and utilizes a control interface to generate interface-level actions (ϕ_i^t and ϕ_m^t) which then get converted to low-level robot control commands (u_h^t) via a transformation function denoted as Υ (Line 5). The model of human behavior maintained by the autonomous agent assumes that the human seeks to minimize the distance travelled and the number of mode switches executed during teleoperation. The autonomous agent utilizes a goal-dependent policy (Ξ) to generate the autonomous command denoted as u_a^t (Line 7). The final low-level control command issued to the robot, denoted as u_f^t, is due to the blending-based shared control that is available during the human's turn (Line 8). The blending factor α is a strictly non-decreasing piecewise linear function of the probability $p(g')$ associated with the inferred goal g' and is given by

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

with $\rho_i \in [0,1]\ \forall\ i \in [1,2,3]$ and $\rho_2 > \rho_1$. In our implementation, we empirically set $\rho_1 = \frac{1.1}{n_g}, \rho_2 = \frac{1.2}{n_g}$ and $\rho_3 = 0.8$, where $n_g = |\mathcal{G}|$. Note that higher confidence in prediction results in higher values of α. The arbitrated command u_f^t then results in robot state transition according to a transition function $\mathcal{T}^u$.

During the autonomous agent's turn (Line 10), the agent does one of two things depending on the confidence of its prediction of the intended goal—as measured by the

entropy $H(\cdot)$ of the belief distribution over goals. Strategy 1: If prediction confidence is low (Line 11), then the agent nudges the robot into a maximally disambiguating state (s^*) in the local neighborhood of the current state (Lines 12-13). Note that the change in state could result in (a) a mode switch, (b) a change in position, or (c) both. In cases (b) or (c), the path is determined by the goal-dependent autonomous policy (Ξ). Strategy 2: If prediction confidence is high, then the agent moves the robot towards the inferred goal location by a distance determined by β (Line 15), leaving the active mode unchanged ($\Psi_m(s^t)$). The interaction protocol is such that the autonomous agent's turn cannot be interrupted by the human and is complete only when the target state is achieved. After the autonomous agent successfully moves the robot to the target state, the turn is handed over back to the human and the task execution continues.

With this strategy, the autonomous agent can continue to contribute to task progress, in addition to providing control-blending based assistance during the human's turn. Robot motion and mode switches executed by the autonomous agent during its turn can also provide valuable information *to* the human about the assistance strategies used by the autonomous agent, therefore potentially improving transparency and cooperation.

5 Experimental Design

Each experimental study session consists of four phases: Phase 1: Training and data collection to model $p(\phi_i^t|a^t)$. Phase 2: Training and data collection to model $p(\phi_m^t|\phi_i^t)$. Phase 3: Familiarization with teleoperation, control blending and the turn-taking based interaction protocol. Phase 4: Algorithm evaluation. In Phase 4, the subjects perform a navigation task using a 3D point robot using the sip-and-puff interface towards pre-defined goals under two experimental conditions. In total, we collected 432 trials (216 per evaluation condition). We conducted a human subject study (n = 9) to evaluate our turn-taking based interaction protocol that deploys the disambiguation algorithm. All participants gave their informed, signed consent to participate in the experiment which was approved by Northwestern University's Institutional Review Board.

5.1 Experimental Setup

We design a simulated navigation environment (Fig. 2) in which subjects operate a 1D sip-and-puff (SNP) device to (a) perform mode switches and (b) control a 3 Degrees-of-Freedom (DoF) point robot's motion along two translational (x, y) and one rotational (θ) dimensions, one at a time towards a pre-defined goal, g_{true} (shown in red). We opt for the SNP, as it is one of the most information-limited interfaces used by people with severe motor impairments. For this environment, modes

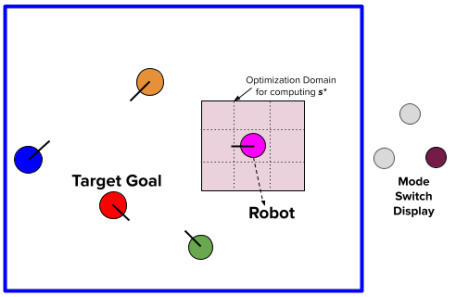


Fig. 2: Simulated navigation environment. The mode switch display highlights the current active mode. Clockwise and counter-clockwise mode switches are possible.

$\mathcal{M} = \{\textit{Horizontal}, \textit{Vertical}, \textit{Rotational}\}$. In order to compute s^* the continuous 3D robot state space is discretized into a $10\times10\times8$ grid that represents 10×10 x-y grid locations and eight discrete orientations for each cell. Note that this discretization is for the computation of $D(s)$ only; the robot positions, velocities, and goal positions are all continuous-valued. The autonomous control policy is generated using the algorithm developed in [11] and operates in the full 3D space.

In order to facilitate seamless turn-taking between the subject and the autonomous agent, a text display presents the subject with information regarding the state of interaction and the environment's boundary transitions from blue to red over a fixed time period ($\sim$3-4s) to remind the subjects that they should hand over the turn. At the beginning of the subject's turn they may wait any amount of time before issuing a first command, and use this time for planning or simply to rest. After this, during execution, handover to the autonomous agent is triggered by not issuing any commands for $\sim$1.5-2s, at any time—and thus whenever the *human* deems it appropriate.

5.2 Training protocol

Learning personalized distributions: Data collection for learning the model for the interface operation is done according to the procedures laid out in [6]. For estimating $p(\phi_i|a)$ users are shown a graphical depiction of a and are instructed to select the correct ϕ_i. For $p(\phi_m|\phi_i)$, the users are shown an interface-level action on the screen and asked to generate the same action using the interface.

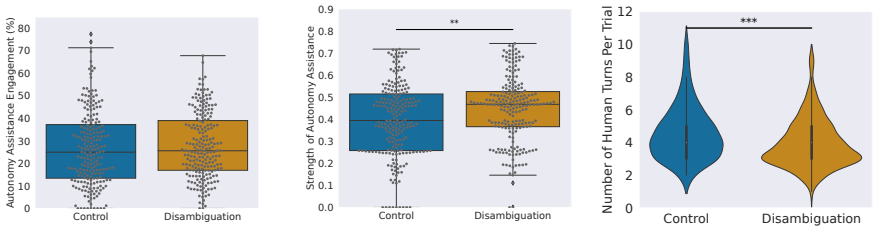
Familiarization with environment and robot control: Participants first are trained on the physical mechanism of operating the interface. Subsequently, they become familiarized with the environment and gain practice in both robot teleoperation as well as in interacting with the autonomous agent via control blending during the turn-taking process.

5.3 Algorithm Evaluation

In the evaluation task, the subject controls the motion of a 3-DoF point robot to reach a 3-D goal (Fig. 2). The number of the goals vary from three to four. For each trial, the starting position of the robot is randomized and diametrically opposite from the goal region. A trial always starts and ends with the subject's turn. Subjects perform the evaluation task under two conditions: (a) *Disambiguation* and (b) *Control*. The presentation order of conditions was random and balanced across subjects.

Disambiguation Condition: During the autonomous agent's turn, the procedure outlined in Lines 10-15 of Algorithm 1 is activated with the constraint that the optimization domain for computing s^* is a local neighborhood grid of fixed size 3×3 centered around the state s^t at the beginning of the autonomous agent's turn. λ is set to be 1.0.

Control Condition: During the autonomous agent's turn, the robot is moved towards the goal with the highest probability by issuing autonomous control commands. If multiple goals are tied for highest probability, then the mean of all those tied goals is the target state. The distance moved is sampled randomly from within the fixed size local neighborhood (3×3 grid) used in the *Disambiguation* condition.



(a) (b) (c)

Fig. 3: (a) Percentage of time that autonomy assistance is engaged during a trial. (b) Strength of autonomy assistance (as measured by the blending factor α) during a trial. (c) Number of human turns per trial. Box plots show median and quartiles. The black dots represent the individual data points.

A trial is deemed successful if the robot's pose coincides with the red goal (both position and orientation) within a predefined threshold. Subjects perform six blocks of eight trials each. After each block, the subjects are asked to respond to a NASA-TLX questionnaire and a post-task survey in which they are queried about their subjective evaluation of how well the autonomous agent is able to assist them during the task. We further evaluate the effectiveness of the disambiguation algorithm according to the following metrics.

Assistance Engagement: Fraction of time the autonomous agent activates assistance towards the true goal ($\alpha > 0$ and $g' \equiv g_{true}$) during the human's turn in a trial.

Strength of Assistance: Average value of the blending factor α over all timesteps when both goal inference is correct and blending assistance is active.

Number of Mode Switches: The number of human-initiated mode switches; a significant contributor to the cognitive and physical effort required for task execution.

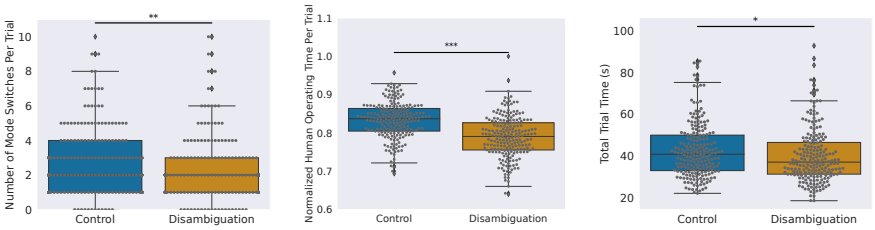
Human Effort: Fraction of total trial time that a subject spent operating the robot in order to accomplish the task.

Task Completion Time: Total time taken to complete the task successfully.

6 Results

We analyze group performances using the non-parametric Kruskal-Wallis test and perform the Conover's test post-hoc pairwise comparisons to find the strength of significance. For all figures, $* : p < 0.05$, $** : p < 0.01$, and $*** : p < 0.001$.

Objective Task Metrics: Fig. 3a shows the fraction of time the autonomous agent activates blending assistance towards the correct goal during the human's turn. For any time t if $\alpha > 0$ and $g' \equiv g_{true}$ then assistance is activated. We do not observe any statistically significant difference between the two experimental conditions. However, in Fig. 3b the strength of assistance offered by the autonomous agent towards the correct



(a) (b) (c)

Fig. 4: (a) Number of human-initiated mode switches during a trial. (b) Fraction of total trial time that subjects spent operating the robot. (c) Total task completion time per trial.

goal, as measured by the average value of the blending factor α over a trial, is higher for *Disambiguation* compared to the *Control* condition ($p < 0.01$). Since α is an non-decreasing function of the associated probability, higher values of α imply that $p(g')$ is higher as well. This indicates that in the *Disambiguation* condition, the autonomous agent is able to be more confident in its prediction of the true goal and therefore provide stronger assistance (greater control contribution by the autonomous agent to the overall control signal) without overtly engaging with the human for more time.

A statistically significant decrease in the number of mode switches furthermore is observed between the *Disambiguation* and *Control* conditions (Fig. 4a). In Figure 4b we observe a statistically significant decrease in the amount of time the subjects spent operating the robot in the *Disambiguation* condition, measured as a fraction of total trial time. Overall, task completion time is lower under the *Disambiguation* condition (Fig. 4c). We also observe that subjects perform fewer turns under the *Disambiguation* condition as compared to the *Control* condition (Fig. 3c).

Lastly, the overall task success for the *Disambiguation* condition (93.05%) was slightly higher than the *Control* condition (90.2%).

Figure 5 illustrates a trial during which the goal probability associated with the true goal (in red) dramatically jumped (twice) during the human's turn immediately following the robot being nudged into state s^* during the autonomous agent's turn.

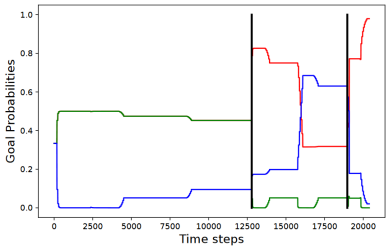


Fig. 5: Evolution of three goal probabilities (colored lines) for a *Disambiguation* trial. The black vertical line indicates the start of the human turn after the autonomous agent nudged the robot into state s^*. The evolution of green and red probabilities were identical prior to this timestep.

Fewer mode switches, fewer number of turns, and faster trial times likely correlate with less human effort. In the *Disambiguation* condition, the subject is able to execute actions that are maximally informative about the true goals as the agent nudges the robot into maximally disambiguating states. Task completion time is shorter because the autonomous agent

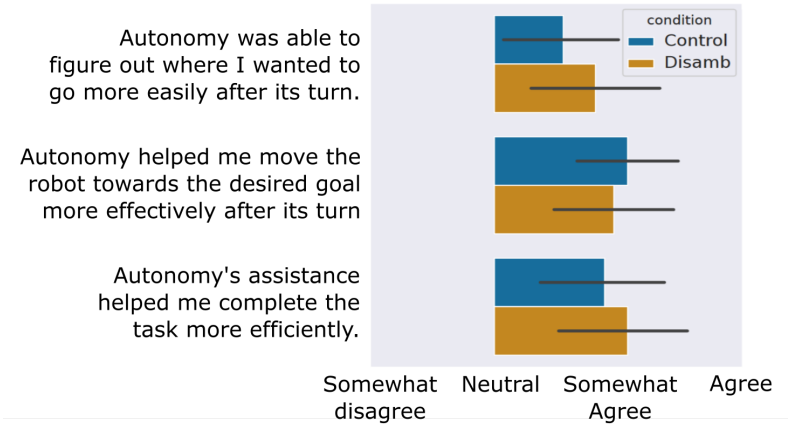


Fig. 6: User response to post-task survey questions. Mean with standard deviation.

is able to infer human intent with more confidence and therefore provide stronger and more accurate assistance towards the correct goal. Manual mode switches become unnecessary once control blending is activated by the autonomous agent, and the robot is also able to move in all three dimensions simultaneously.

Subjective Task Metrics: We use the raw NASA-TLX as a subjective measure of perceived workload [9]. Although the mean score for the *Disambiguation* condition (32.75) is slightly lower than the *Control* (34.45) condition, we do not observe a statistically significant difference. We evaluate user preferences and acceptance of our shared-control assistive paradigms using a questionnaire (Fig. 6). The statements are rated on a 7-point Likert scale from strongly disagree (1) to strongly agree (7). Overall, the subjects rate the *Disambiguation* condition higher than the *Control* condition when it comes to the agent's ability to *figure out* the human's intended goal faster. However, subject perception is that in the *Control* condition the autonomous agent is more effective in helping them *move* towards the desired goal.

7 Discussion

The computation of the disambiguation metric critically depends on whether it can be empirically estimated from a generative model. Our results indicate that despite having an approximate model for human behavior (one in which the human is assumed to minimize path distance to goal and the number of mode switches), the autonomous agent was able to leverage the model and successfully compute reasonably good disambiguating states. Performance would likely improve further if more accurate models of human behavior—learned from large amounts of data using state-of-the art machine learning techniques—are used in conjunction with the proposed disambiguation algorithm.

Although turn-taking allows the user to observe the autonomous agent's actions and acquire a mental model of the autonomous agent's policy, sufficient training and

priming is important so that the subject's expectation of the autonomous agent's policy is close to its true policy. With more training and practice, the human-autonomy team can achieve common ground faster and the human likely will be incentivized to work in a cooperative manner and *leverage* the assistance offered by the autonomous agent.

In Eq. 16, the domain over which the optimization occurs can vary depending on whether the autonomous agent wants to move the robot in a small neighborhood from the current position or whether it plans to execute large scale motions. In the small neighborhood condition, the idea is that the autonomous agent's actions would be interpreted by the user as small nudges as opposed to a complete takeover of user control. Disambiguation over the entire state space might, however, reduce human effort significantly by reducing the overall number of turns and also the fraction of total time the human operates the device. It might be beneficial to allow the user to pick the optimization domain depending on their preference. Yet another algorithmic modification would be to reason over larger time horizons; but this would incur a higher computational cost.

Prior task structure can be leveraged to simplify the computation of the disambiguation metric. For example, in an assistive robotic arm, although the full task space is six dimensional, prior constraints on task execution (such as the subject having to reach for an object before grasping it) might allow for disambiguation to be computed within a lower dimensional (such as 3D translation) space. Disambiguation is especially useful during the earlier parts of task execution when inference uncertainty is the highest; and in a typical manipulation task, translational reaching motion happens earlier.

8 Conclusion

In this paper we have presented a novel interface-aware intent disambiguation algorithm grounded in information-theoretic principles. The primary goal of this algorithm is to elicit maximally informative control signals from the user by placing them in states that have the highest disambiguation capabilities as determined by the metric. The paper also introduces a turn-taking based human-autonomy interaction protocol in which the autonomous agent utilizes the proposed disambiguation metric to extract information-rich actions from the human when uncertain about its prediction of human intent. The efficacy of both the proposed algorithm and protocol was evaluated via a 9 person human subject study. The results indicated that the disambiguation system resulted in a statistically significant decrease in task effort in terms of the number of manual mode switches executed and the fraction of time the users spent operating the robot.

References

1. Atanasov, N., Le Ny, J., Daniilidis, K., Pappas, G.J.: Information acquisition with sensing robots: Algorithms and error bounds. In: *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)* (2014)
2. Barfoot, T.D.: *State estimation for robotics*. Cambridge University Press (2017)
3. Brooks, C., Szafir, D.: Balanced information gathering and goal-oriented actions in shared autonomy. In: *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. pp. 85–94. IEEE (2019)

4. Callaway, F., Hardy, M., Griffiths, T.: Optimal nudging. In: *CogSci* (2020)
5. Ewert, B., Loer, K., Thomann, E.: Beyond nudge: advancing the state-of-the-art of behavioural public policy and administration. *Policy & Politics* 49(1), 3–23 (2021)
6. Gopinath, D., Javaremi, M.N., Argall, B.: Customized handling of unintended interface operation in assistive robots. In: *2021 IEEE International Conference on Robotics and Automation (ICRA)*. pp. 10406–10412 (2021)
7. Gopinath, D.E., Argall, B.D.: Active intent disambiguation for shared control robots. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 28(6), 1497–1506 (2020)
8. Gopinath, D.E., Argall, B.D.: Mode switch assistance to maximize human intent disambiguation. In: *Robotics: Science and Systems* (2017)
9. Hart, S.G.: NASA-task load index (NASA-TLX); 20 years later. In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (2006)
10. Herlant, L.V., Holladay, R.M., Srinivasa, S.S.: 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)
11. Huber, L., Billard, A., Slotine, J.J.: Avoidance of convex and concave obstacles with convergence ensured through contraction. *IEEE Robotics and Automation Letters* 4(2), 1462–1469 (2019)
12. Javdani, S., Admoni, H., Pellegrinelli, S., Srinivasa, S.S., Bagnell, J.A.: Shared autonomy via hindsight optimization for teleoperation and teaming. *arXiv preprint arXiv:1706.00155* (2017)
13. Johnson, E.J., Shu, S.B., Dellaert, B.G., Fox, C., Goldstein, D.G., Häubl, G., Larrick, R.P., Payne, J.W., Peters, E., Schkade, D., et al.: Beyond nudges: Tools of a choice architecture. *Marketing Letters* 23(2), 487–504 (2012)
14. Lala, D., Inoue, K., Kawahara, T.: Smooth turn-taking by a robot using an online continuous model to generate turn-taking cues. In: *2019 International Conference on Multimodal Interaction*. pp. 226–234 (2019)
15. LaPlante, M.P., et al.: Assistive technology devices and home accessibility features: prevalence, payment, need, and trends. *Advance Data from Vital and Health Statistics* (1992)
16. Miller, L.M., Silverman, Y., MacIver, M.A., Murphey, T.D.: Ergodic exploration of distributed information. *IEEE Transactions on Robotics* 32(1), 36–52 (2016)
17. Robotics, C.S.S., Nørskov, M., et al.: Nudging by social robots. *Culturally Sustainable Social Robotics: Proceedings of Robophilosophy 2020* 335, 337 (2021)
18. Sadigh, D., Sastry, S.S., Seshia, S.A., Dragan, A.: Information gathering actions over human internal state. In: *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. pp. 66–73. IEEE (2016)
19. Skantze, G.: Turn-taking in conversational systems and human-robot interaction: a review. *Computer Speech & Language* 67, 101178 (2021)
20. Soleiman, P., Moradi, H., Mahmoudi, M., Teymouri, M., Pouretemad, H.R.: Teaching turn-taking skills to children with autism using a parrot-like robot. *arXiv preprint arXiv:2101.12273* (2021)
21. Thaler, R.H.: From cashews to nudges: The evolution of behavioral economics. *American Economic Review* 108(6), 1265–87 (2018)
22. Thaler, R.H., Sunstein, C.R.: Nudge: Improving decisions about health, wealth, and happiness (2008)
23. Thomaz, A.L., Chao, C.: Turn-taking based on information flow for fluent human-robot interaction. *AI Magazine* 32(4), 53–63 (2011)