Towards an Information Theoretic Framework for Human-Autonomy Interaction in Shared Control

Thesis proposal

Deepak Edakkattil Gopinath
Department of Mechanical Engineering
Northwestern University
deepakgopinath@u.northwestern.edu

October 25, 2019

Abstract

Human-Autonomy Interaction (HAI) in the context of shared-control robotics is a rich and complex phenomenon. The effectiveness and usefulness of shared-control human-machine systems critically depends on the fluency and efficacy of HAI. Efficient HAI can lead to an improvement in joint task performance with higher user satisfaction and enhanced trust, all of which are desired characteristics of a human-machine system. From an engineer's perspective, in order to achieve optimal performance the design of autonomous behaviors should adequately take into account the rich and complex interaction dynamics between the human and the machine.

In my thesis, I plan to propose a mathematical framework for the characterization of HAI in shared-control robotics utilizing ideas from *probabilistic graphical models* and *information theory*. More specifically, the interaction between human and autonomy will be modeled as coupled perception-action loops unfolding in time using *causal Bayesian Networks*. Within this framework of causal Bayesian Networks, autonomy's actions can be thought of as appropriately timed *interventions* at specific parts of model with specific desired outcomes one of which is to alter the bi-directional information flow between the human and machine thereby affecting the fluency, cooperation and transparency of HAI. Using the proposed mathematical model, I will research three important problems that arise in HAI namely, a) **quantification of transparency**, b) **autonomy-guided improvement of human skill acquisition** and c) **enhancement of overall task performance**. The eventual goal is to utilize the proposed mathematical framework to inform the design of autonomy that will help to *quantify transparency levels in a given interaction scenario*, *facilitate human skill acquisition*, and *enhance task performance*.

Contents

1 Introduction

Robots are ubiquitous in modern society and have revolutionized the relationship between human and machine. Compared to a few decades ago, in the present day, robots have transitioned out of the rigid, structured and specialized industrial environments to the more rich, complex and unpredictable day-to-day human environments and have impacted diverse domains of human endeavor such as healthcare [48], entertainment [26] and home robotics [21].

The impact is even more significant in the domain of assistive and rehabilitation robotics, in which the potential to drastically enhance the quality of life for people suffering from motor impairments as a result of spinal cord or brain injuries is immense [54]. Devices such as smart wheelchairs, exoskeletons and assistive robotic arms can help to promote independence, boost self-esteem and help to extend mobility and manipulation capabilities of motor-impaired individuals and can revolutionize how they interact with society [53].

The standard usage of these assistive machines still relies on manual teleoperation by the human, typically enacted through a control interface such as a joystick or a switch-based head array; that is, in such scenarios these robots are not endowed with any intelligence and are treated as extensions of human motor abilities [66]. However, one of the most difficult conundrums is that the greater the motor impairment of the user, the more limited the interfaces that are available for them to use (for example, switch-based head array and sip-n-puff tubes). As a result, control of these machines can become extremely difficult due to the low dimensionality, sparsity and bandwidth of the control interfaces and are further exacerbated by the inherent complexity in robot dynamics and the physical limitations of the users [60]. In such cases, *robot autonomy*, the intelligence that enables robots to accomplish a task independently without requiring explicit instructions from a human, holds considerable promise as a tool to offset (and in some cases restore) the above-mentioned limitations. Advances in the fields of machine learning and artificial intelligence have helped to endow these assistive machines with better decision making and prediction capabilities when interacting with humans in real-world scenarios [34]. However, there is a growing consensus that users of assistive technologies *do not* prefer to cede full control authority to the robotic partner during task execution [25]. In such cases, the introduction of *shared control* seeks to find a middle ground between full teleoperation and autonomy by offloading only some aspects of task execution to the autonomy [76, 14].

In a shared-control system, the task responsibility is split between the user and autonomy usually with the aim of reducing human effort in accomplishing a task. Human-Autonomy Interaction (HAI)[1] in the context of shared-control is a rich and complex phenomenon. The effectiveness and usefulness of shared-control human-machine systems critically depend on the quality and efficiency of HAI. That is, for robots and humans to work side-by-side and achieve joint goals and accomplish various tasks in a *coordinated* and *cooperative* manner, it is imperative that both parties understand each other, communicate and infer internal desires and intentions efficiently [32]. From an engineering perspective, design of appropriate kinds of autonomous behaviors for a shared-control system, therefore, will likely benefit from the knowledge of the *dynamics* of HAI during the

[1]Human-Autonomy Interaction is closely related to Human-Robot Interaction. However, in this proposal, we interpret the term 'robot' as an embodied physical entity that exists in the world and 'autonomy' as the intelligence that endows the robot with a sense of agency. In the context of shared-control, human and autonomy jointly control the robot to accomplish tasks in a coordinated manner. The underlying interaction is between human and autonomy, hence the term HAI.

course of task execution [33].

Current approaches for design of shared-control systems rely on various kinds of mathematical models to solve different aspects of HAI as independent subproblems and therefore suffer from generalization across tasks, robotic platforms and user types. In this proposal, I am motivated by the desire to develop a *unified* mathematical framework to analyze different aspects of HAI under a single umbrella in an attempt to shed light on the more *fundamental* and *low-level* characteristics of human-robot teaming.

To that end, I propose a mathematical model of HAI in the context of shared-control, utilizing ideas from *probabilistic graphical models* [44] and *information theory* [12]. More specifically, the interaction is modeled as *coupled perception-action loops* unfolding in time using *causal Bayesian Networks* [59]. The nodes in the network represent the different variables (both latent and observed) that are relevant for the model and the edges represent the probabilistic influence they have on each other. In an attempt to quantify the fluency, transparency and cooperation levels that characterize the interaction, the *information flow* between the nodes in the network is analyzed [3]. This model is utilized by the autonomy to reason about the environment (which includes the human) and make decisions that will affect the overall joint task performance. Within this proposed framework of causal Bayesian networks, autonomy's actions are interpreted as appropriately timed *interventions* at different nodes of the network. These interventions, among other objectives, can potentially alter the bidirectional information flow between human and autonomy, thereby affecting the fluency and transparency of HAI. The key insight is that *information flow* is a more fundamental and low-level descriptor of interaction dynamics and joint system performance, and should be one of the main areas of focus for system designers when designing autonomous behaviors. Using the proposed model, I intend to address three main subproblems relevant to shared autonomy: *quantification of transparency*, *facilitation of human skill acquisition* and *enhancement of overall joint task performance*.

The first research question (**RQ1**) that I will address in my work concerns the *quantification of transparency* of HAI. 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 legible bi-directional communication of internal goals and intentions which in turn facilitates cooperation and coordinated task execution [40]. The effectiveness and usefulness of shared-control systems critically depend on the fluency and transparency of HAI. In **RQ1**, I will utilize the notion of information flow in Bayesian Networks as measured by *transfer entropy* to characterize autonomy-to-human and human-to-autonomy transparency directly from sensor measurements.

The second research question (**RQ2**) that I will address in my thesis is *how can autonomy help humans acquire robot teleoperation skills more effectively*. A higher skill level helps the users to express their intentions more clearly, thereby improving *human-to-autonomy* transparency. This in turn can help the autonomy infer the users' intentions more accurately and thereby provide appropriate kinds of assistance in a timely fashion. I am motivated by ideas from *curriculum learning* [4] and seek to develop learning schemes that guide the user to spend more time and gain more practice in those regions of the state space where robot control is difficult. In other words, autonomy can play the role of an *informative teacher* and help the human in skill acquisition and provide appropriate guidance during the training/practice phase. Potentially, this can have a significant impact in the design of training procedures for new users of assistive robots, resulting in better experiment designs and improving the baseline above which the assistance provided by autonomy can become more effective.

In my last research question (**RQ3**), I will focus on design of autonomous control policy that explicitly optimizes transparency of HAI in a shared-control setting, in addition to other task-related metrics. This will likely result in a common ground [40] for joint task execution, which will lead to enhanced cooperation, coordination and mutual trust. The hypothesis is that, as a result of these enhancements, the desired objective and subjective outcomes (such as improved task performance, enhanced inference accuracy and higher user satisfaction) will likely emerge.

In summary, this proposed thesis intends to develop an information-theoretic mathematical framework to characterize HAI within shared-control systems and plan to research solutions to the questions presented above (**RQ1**, **RQ2** and **RQ3**). This work will treat information content and flow as the key components in understanding the dynamics of interaction between human and autonomy during task execution and training. More importantly, this work proposes a fundamentally different way of thinking about autonomy's role; one in which *autonomy's actions are exogenous interventions that alter the information flow in a coupled perception-action loop to bring about desired outcomes during task execution and skill acquisition*.

2 Background Literature

In this chapter, I present a discussion of existing literature on mathematical approaches for modeling different aspects of HAI in a shared-control system. I also briefly discuss how research into softer aspects of HAI such as legibility, transparency, cooperation, attention and coordination in other domains such as cognitive psychology and philosophy can help to guide the design of effective human-robot teaming strategies.

2.1 Mathematical Models for Shared Control

In shared control, complementary abilities of humans and autonomy are leveraged to jointly accomplish various tasks such that the joint system is typically more capable than either the human or the machine on their own. However, there is no one singular definition of what shared control is. In a recent survey paper, shared control is defined as one in which '*...human(s) and robot(s) are interacting congruently in a perception-action cycle to perform a dynamic task, that either the human or the robot could execute individually under ideal circumstances*' [1]. More importantly, the authors also propose that in a shared-control system the human and robot actions should be linked by combining them into a final control action, plan or a decision and that each agent should directly perceive how its intent is influenced by the actions of other agent(s). Clearly, HAI in this context is a complex phenomenon and for robots to function properly alongside humans, it becomes imperative that they have the capability to predict their partners' actions and intentions effectively using accurate mathematical models.

Researchers develop mathematical models for various purposes; for example, to model human behavior, to learn autonomous policies from human demonstrations, to recognize objects in the environment and generate waypoints for navigation and grasp poses for manipulation, to determine control allocation and decide how the human and autonomy control commands should be arbitrated to produce the final control command. In the following subsections I discuss each of the above-mentioned categories in greater detail.

2.1.1 Models for Human Behavior

Teamwork in a shared-control system is enhanced when the team members understand each other's intentions and goals. However, there are particular challenges that arise in HAI due to the differences in the mental and physical capabilities of humans and robots [31]. Robots can deal with such challenges by maintaining models of human cognition and behavior [36] spanning different timescales and levels.

In [31] the authors utilize 'Marr's levels of analysis' [52] to categorize models for human behavior into three distinct categories namely: computational, algorithmic and implementational. Categorization of human models using Marr's level of analysis clarifies what aspects of human behavior is being modeled. Computational level techniques are ideal for scenarios that benefit from the knowledge of normative behavior that humans are ought to exhibit. These models typically rely on simplistic assumptions of perfectly rational behavior and treat human idiosyncrasies and deviations from the norm as observational noise. Algorithmic level analysis, on the other hand, seeks to delve into the processing constraints that agents have and how they lead to systematic errors thereby providing better insight into *why* agents deviate from normative behavior. However, the algorithmic models typically work well over shorter timescales and therefore are not suited for modeling human behavior that last over longer timescales.

Within the computational category, one of the most common methodologies is to adopt simple probabilistic models that attempt to model very low-level short-time horizon behaviors (such as reaching motions performed by humans during a manipulation task). For example, in [17] a framework in which the human is treated as an optimal agent that noisily optimizes a goal-dependent cost function is used. This model is used by the robot to infer user's intent in which the predicted goal is the one with the lowest cost given the user's control input. Optimality principles are particularly attractive because of their success in the domain of motor control [75] and also because they provide a principled approach to how agents ought to behave. This type of framework, however, requires well-defined cost functions that provide succinct description of the task at hand. Cost functions can either be hard-coded under assumptions of rationality or hand-designed by domain experts (for example, minimization of distance to goal) or can be learned from human demonstrations using techniques such as inverse reinforcement learning [79] and inverse optimal control [18]. Data-driven approaches utilizing conventional machine learning algorithms are also successfully used to recognize human behavior in a wide variety of domains such as social robotics [53] and assistive robotics [22]. More recently, data-driven approaches based on Koopman operators have also been utilized to learn models of joint human-machine systems [8]. Koopman operator based approaches scale well to high-dimensional spaces as the computational complexity does not grow with the number of data points.

In my work on human-driven customization of shared autonomy levels [25], we utilize ideas from optimal control theory to model human behavior under the assumption that humans act optimally with respect to an *unknown* (to autonomy) cost function. In this paper, we make no assumptions regarding the nature of this cost function, but instead provides the human the capability to customize/optimize the control allocation parameters directly. The key insight in this work is that the human has an internal representation of what s/he wants (possibly encoded as a cost function that the human is trying to optimize) and therefore by having the human optimize the control allocation parameters directly, the user will tune them to their personal liking and preference. Similarly, my work on mode switch assistance for intent disambiguation implicitly assumes that the model for human teleoperation of the robot is one in which the human is optimizing for shortest path to goal.

In general, computational approaches are well suited for situations in which one is primarily concerned about what the human is doing without necessarily reasoning about the underlying causes for the behavior.

In the algorithmic category, Hidden Markov Models (HMMs) and other Markov-based approaches are common choices for modeling human behavior [38]. HMMs are powerful due to their ability to express latent variables and can be used for efficient online inference of hidden states given a set of observations. An extension of HMMs that can be useful for modeling human-robot collaboration is the Partially Observable Markov Decision Process (POMDP) [68]. For example, POMDP based models are used for assessing the human partner's trust in the autonomous partner [9]. POMDPs can also be used by autonomy in which human intent is treated as a latent state. By performing online inference on these latent states, autonomy can incorporate human intentions into its own decision making process thereby implicitly taking into account user preferences and goals. In general, inference over latent variables in a POMDP framework is performed via Bayesian methods in which a prior distribution over the latent variables is updated via Bayes Theorem upon receiving new evidence. The choice of likelihood function, typically encodes the human's actions/preferences givem a state [16]. A generalization of HMMs, known as Dynamic Bayesian Networks (DBNs) have also been used successfully to model human-robot collaboration. DBNs are attractive due to their ability to handle multi-variate, mixed-observability variables and to represent temporal interdependencies [56]. DBNs have been successfully utilized to model human beliefs, desires and learning which can then be used for intent prediction and understanding cognitive phenomena such as concept learning [69]. However, structure of DBNs need to pre-specified by a domain expert and learning the parameters requires large amounts of data.

In my prior work, I have relied on heuristic approaches based on simple directedness and proximity-based confidence functions to estimate reaching intent in table-top manipulation tasks [25]. In order to effectively incorporate information from past states, I have also developed a field-theoretic framework for belief update utilizing ideas from *Dynamic Field Theory* [63] in which the time-evolution of the probability distribution over intent/goals is specified as a constrained dynamical system evolving in continuous time [24]. The activation function that drives the dynamical system is analogous to the likelihood function in Bayesian approaches in that it encodes the human's preferences. Additionally, in the domain of autonomous driving, I have developed a framework to learn agent models from partial observations that will help to predict their actions from noisy observations and deduce missing information about their environment. The probabilistic models for partially observed agents are learned via IRL techniques through which we recover the agent's policy and reward structure. Within the model, efficient imputation of latent states and actions is made possible by learning specific proposal distributions [28] to be used within an Reversible Jump Markov Chain Monte Carlo (RJMCMC) sampling process [27].

2.1.2 Models for Policy Generation

Mathematical models are also widely used for learning policies responsible for generating autonomy's control actions. In addition to successful task completion, autonomy would benefit from its actions to be legible, natural, and safe. *Learning from Demonstration* (LfD) [2] provides a framework to learn autonomous policies directly from user-provided demonstration data. For example, imitation learning (IL) paradigms can be utilized to develop end-to-end systems that can directly map a high dimensional states to actions [5]. A more generalizable approach is to cast the problem with the framework of *inverse reinforcement learning* (IRL) in which the goal of the algorithm is to recover the user's reward function [79]. Policies that optimize long-term accumulated reward (solution to the forward reinforcement learning (RL) problem) improves the robustness and generalization capabilities of the autonomous agent. Closely related to the RL approach is to utilize control-theoretic framework to derive optimal policies for a given task. Standard optimal control theory techniques are model-based, that is, they presume the existence of a dynamics model and solve for the optimal policy with respect to a specified cost function [42]. State-of-the-art RL techniques, on the other hand, can be model-free and can resort to sampling-based techniques to derive optimal policies [77].

Planning-based approaches such as probabilistic road maps [37] and RRT [47] are also widely used for generating motion plans (for both robotic manipulators as well as mobile robots). In addition to task accomplishment, in order to enhance user experience robot motion plans typically need to possess various other desired characteristics. HAI in shared-control can become more seamless if the autonomy is able to make its intentions legible to the user. To that end researchers have attempted to mathematically define legibility and predictability of robot motion [15]. Similarly, safety is of paramount importance when robots work in close proximity to humans. Therefore behaviors such as obstacle avoidance are incorporated into navigation plans [67]. To this end, the on-board perception system typically relies on object detection and recognition models to identify objects of interest and obstacles thereby characterizing favorable and unfavorable parts of the state space of the robot during navigation [55].

In my work with assistive robotic manipulators for table-top manipulation [25], I have utilized potential fields defined in the full six degrees-of-freedom Cartesian space (task space) to generate the robot control commands. Obstacle avoidance is implemented using velocity-dependent repellers and the intended goal is modeled as the attractor. Potential fields are computationally lightweight and produces more intuitive trajectories that correspond to straight line paths in Euclidean space [39]. More recent work have incorporated ideas from differential geometry to treat obstacles as local deformation in the geometry of the workspace and aims to derive motion policies directly in a curved Riemannian workspace [62]. The effect of the obstacle is to curve the geodesics (straight line paths) around itself as determined by the local curvature induced by the obstacle [51].

2.1.3 Models for Control Allocation

In a shared-control setting task responsibility is split between the human and autonomy. Therefore, control arbitration between the human and the autonomous partner needs to implemented appropriately for desired outcomes. Different approaches exist for control allocation between the human and autonomy that depend on the application domain, user preferences and robot platform. Broadly speaking, shared-control approaches can be broadly classified into two main categories: hierarchical and blending-based approaches.

In the hierarchical paradigm, control allocation typically occurs at the task level in which high-level task goals are entrusted with the human and the autonomy takes care of the low-level control of the robot. For example, in the assistive domain, smart wheelchair users can use a click-based interface [66] to select a desired destination in the world or a laser pointer [10] to point to a desired goal. The autonomy can then generate the global as well as local plans utilizing any state-of-the-art motion planners. In the domain of table-top manipulation users can use natural language to specify the desired grasp or manipulation target which then combined with object recognition and motion planning modules can produce the desired robot trajectory [7].

Blending-based approaches seek to arbitrate between human and autonomy actions at the control signal level (or the policy level) directly. In [17] a policy-blending formalism for shared control is presented in which the authors propose that "arbitration should be contextual and the should depend on the robot's confidence in itself and in the user, as well as on the particulars of the user". A commonly used arbitration scheme is one in which the final control command issued to the robot is a linear combination of human and autonomy control commands. The blending parameter can be fixed or can be a parameterized function of the context and the autonomy's confidence in its prediction of the user's intent. By casting shared control allocation in a broader theoretical framework, a mathematical model for probabilistic shared control in complex dynamic environments was proposed in [74]. In this work, the interactive relationship between the human, autonomy and the environment is modeled as a undirected graphical model. The paper also introduces the notion of an *interaction* function between the operator and the autonomy that captures the "agreeability" between the human and autonomy. For specific forms of the interaction function, linear blending is recovered as a special case of the more general framework. The paper also shows that in general, linear blending is suboptimal with respect to the joint metric of agreeability, safety and efficiency.

In our work on human-in-the-loop customization of control allocation parameters, we utilize a blending-based shared control in which the linear blending factor is a function of the probability of the predicted goal (agnostic to the type of intent inference algorithm used). We assume a piecewise-linear function (with three tunable parameters) and under the assumption that the human is optimizing an unknown cost function, we develop an iterative procedure with which the user is able to tune the arbitration function parameters to the his/her own satisfaction and preference [25].

In the domain of assistive robotics, there exists yet another particularly challenging problem. The standard usage of these assistive machines relies on manual teleoperation typically enacted through a control interface such as a joystick. However, greater the motor impairment of the user, the more limited are the interfaces available for them to use. These interfaces (for example, sip-and-puffs and switch-based head arrays) are low-dimensional, discrete interfaces that can only operate in subsets of the entire control space (referred to as *control modes*). The dimensionality mismatch between the interface and the robot's controllable degrees-of-freedom (DoF) necessitates the user to switch between control modes during teleoperation which has been shown to increase the cognitive and physical burden and affects task performance negatively [60]. In order to offset the drop in performance due to shifting focus (also known as task switching) from the task at hand due to switching between different control modes various mode switch assistance paradigms have been proposed. A simple time-optimal mode switching scheme has shown to improve task performance [30]. Machine learning techniques have been utilized to learn mappings from robot state to control modes preferred by human users [35]. Robot operation in certain control modes can also help the autonomy to infer the user's intent more accurately and confidently, especially in scenarios where autonomy's inference of user intent is exclusively informed by human's control commands issued via the limited control interfaces.

To that end, in our work, on mode switch assistance for intent disambiguation we developed a disambiguation metric to characterize the intent disambiguation capabilities of a control dimension/control mode. By having the user operate the robot in the disambiguating control mode, the control commands become more *intent-expressive* and as a result autonomy is able to step in and provide appropriate assistance. In this work, we also introduce the notion of *inverse legibility*, in which the roles are switched and the human-generated actions *help the robot* to infer human intent confidently and accurately [23].

2.2 Characteristics of Ideal HAI

Humans are highly adept at working together in teams and are able to effectively coordinate and cooperate with their teammates in a seamless, natural and fluid manner in a joint task setting. How are humans able to interact efficiently in a team setting? What do humans prioritize, besides successful task accomplishment when working alongside others? What are some of the fundamental characteristics of a successful team? Researchers in the fields of cognitive psychology and social sciences have been long been interested in these questions [29, 45, 61].

It has been widely established that one of the most essential components for successful teamwork is the need for efficient communication between the members of the team and the need for a shared context and intentionality between the team members [73]. Research in developmental psychology has established that humans develop the ability to share goals and intentions with others very early in life and is widely exhibited in the context of cooperative activity. [73] discusses that human ability to share intentions arise from two main capabilities. First, humans have the capability to infer the latent internal states of other agents from observations of their behavior. Second, humans also have the motivation to share mental states which forms the basis of cooperative activity. *Shared cooperative activity* is an important and recognizable characteristic of everyday human life. Some of the examples from everyday life include, playing a sport together, performing a musical piece with an orchestra *et cetera* [6]. In [6] the author identifies three important characteristics of shared cooperative activity namely a) mutual responsiveness b) commitment of joint activity and c) commitment to mutual support. A related concept is how individuals partaking in a joint activity develop what is known as *common ground*—"that is, the knowledge, beliefs and suppositions they believe they share about the activity" [11]. The theory of common ground was originally developed to understand communication between people and its main assumption is that effective communication requires coordination that relies on shared knowledge to reach mutual understanding. Common ground and mutual understanding between partners typically increase over time as a result of learning from continued interactions and can lead more seamless interaction.

Application of social, psychological and cognitive theories of interactions to the domain of HAI has been successful to a large extent and has contributed to advances in the 'science' of human-autonomy interaction. The theory of 'common ground' can be applied to scenarios in which robots interact with people in public spaces and reinforces the need for shared mental models to form common ground [40]. In [43] the authors outline ten challenges that exist for making an autonomous partner a "team player", one of which is the need for autonomous partners to be "..reasonably predictable and reasonably able to predict others' actions". This points to the need for transparent, legible and predictable robot actions that will likely have a positive impact on the users' perceived trust, satisfaction and comfort level. The need for transparency in communication in order to promote shared awareness and intent has also been studied in [50]. Fluency is yet another important aspect of a successful collaboration. A model for joint human robot action in which the robot makes anticipatory decisions based on the confidence of their validity and relative risk in an attempt to improve the fluency of human-robot interaction is proposed in [33]. The authors make the observation that fluency is not necessarily always related to task efficiency and therefore requires more rigorous quantitative treatment and to that end propose three different fluency metrics as well. Another common paradigm for formalizing HAI is *turn-taking*. A first-order Markov process based model to describe the turn-turning dynamics between a human and a social robot is proposed in [71]. One of their primary findings is that human turn-taking behavior is mainly influenced by information flow. However, the authors do not present a quantitative treatment of the information flow. Drawing influences from research in the human factors community, human-robot teaming models that are shared between the team members have been proposed in [58]. This work is an attempt to build shared mental models that have shown to improve team performance and task efficiency. In their work, the robot derives a team mental model described as a POMDP from observations of coordinated team work performed by two or more expert humans.

In recent years, researchers from the design community have also emphasized the need for user-centric development of shared-control systems. This is particularly important in the domain of assistive robotics in which the transition of the technology into the real world depends a great deal on user satisfaction and acceptance. For assistive technologies to be successfully adopted it is clear that various research components must come together in a seamless manner. These components range from design of software and hardware modules, appropriate human-autonomy interaction schemes and other technical aspects such as efficient algorithms, sensor technologies and control interface design. In our work, we have recognized this need for convergence of research directions, in which end-users play an active role in the iterative design process [19].

2.3 Summary

HAI research is a highly interdisciplinary field drawing influences from a variety of domains such as computer science, machine learning, cognitive psychology and philosophy. From the survey of related literature presented in this chapter, we can see that mathematical models have been explicitly used to deal with a variety of problems that arise in HAI such as modeling human behavior, generating control policies for autonomy, for determining the type and level of assistance and to perceive and making sense of the environment. On the other hand, adoption of ideas from cognitive psychology provides a prescriptive framework that can guide engineers to design autonomous behavior with desirable characteristics.

Some of the recurring themes that are emphasized throughout literature relate to the need for shared mental models, efficient and transparent communication, mutual responsiveness, having the need for common ground and the ability to understand each other's intentions and goals in order to enhance the quality of HAI. These ideas provide a solid theoretical basis for the design of practical shared autonomy systems.

Due to the paramount importance of efficient coordination between a human and a robot in a shared autonomy system, typically facilitated by the exchange large amounts of relevant information (via different modalities), information-theoretic analysis of these interactions can potentially shed light on the efficiency, transparency and legibility of communication. To that end, in this thesis I propose a mathematical framework based on causal Bayesian Networks to analyse HAI with an emphasis on information-theoretic analysis of interactions between the nodes in the network. Within this framework *information flow and exchange* assume a more foundational and fundamental role . Autonomy's influence on the interaction will be treated as interventions at specific nodes of the network that will drive the interaction dynamics to have desired outcomes. In the following chapter, I will introduce the causal Bayesian Network based mathematical framework for studying HAI and a primer on some of the relevant concepts from information theory.

3 Proposed Modeling Framework for HAI in Shared Autonomy

In this chapter, I introduce our model for HAI for information-theoretic analysis represented as a Causal Bayesian Network (CBN). The dynamical exchange between the human and the autonomous partner is modeled as a coupled perception-action loop. The nodes in the CBN represent the various variables (both latent and observed) that are relevant for HAI and the edges represent the probabilistic influence they have on each other [72]. Within this framework, autonomy's actions are interpreted as exogenous interventions that alter the information flow in this coupled perception-action loop to bring about desired outcomes during task execution and learning.

3.1 Causal Bayesian Networks for HAI

In order to motivate how the notion of perception-action loops can be used to describe the interaction between human and autonomy in a shared-control setting, we describe a concrete example.

Consider a scenario in which a motor-impaired human and an autonomous partner jointly control an assistive robotic arm to perform table-top manipulation. For simplicity, we assume that shared control is achieved via a control blending paradigm in which the human's control command and the autonomy's control command are arbitrated in some fashion to produce a final control command that is issued to the robot controllers. We assume that the assistive robotic arm is mounted on a wheelchair in which the person is seated. From the seated position, the person is able to observe and *perceive* various aspects of the environment such as the positions and orientations of the various table-top objects, the shape and color of the objects, the pose of the robot's end-effector *et cetera*. The person's line of sight could be obstructed due to occlusion from the robot and other factors and as a result the human only receives partial information regarding the true state of the environment. Typical robot teleoperation is enacted through a control interface such as a joystick. The actions performed by the human may depend on a variety of factors such as the partial observations of the environment, internal goals and desires, task specifications, constraints of the control interface and so on and so forth. Upon taking an action the environment state evolves due to the inherent stochastic dynamics and the process repeats in time. In essence, the interaction between the human and the environment can be thought of as a **perception-action loop unfolding in time**. On the other hand, in a similar fashion the autonomy also receives partial information about the robot state through various types of sensors (joint encoders, RGBD cameras) and takes different types of actions to control the robot using available information. This interaction between autonomy and environment can also be thought of as another perception-action loop.

Perception-action cycle is considered to be the fundamental logic of the central nervous system, in which perception and action processes are closely interlinked [13]. Perception leads to action, and action leads to perception. In the context of HAI, as the perception-action loop unfolds in time, the human and autonomy interact with each other via explicit and implicit exchange of information. Both agents continually infer one another's latent states and actions. For example, in situations where the human goal is not explicitly specified the autonomy has to infer the human's internal goal from sensor data. Similarly, the human might not be fully aware of the autonomy's collaboration strategy and will have to infer the autonomy's decision-making logic to effectively cooperate and coordinate with the autonomy. This joint interaction can be modeled as the *coupled*

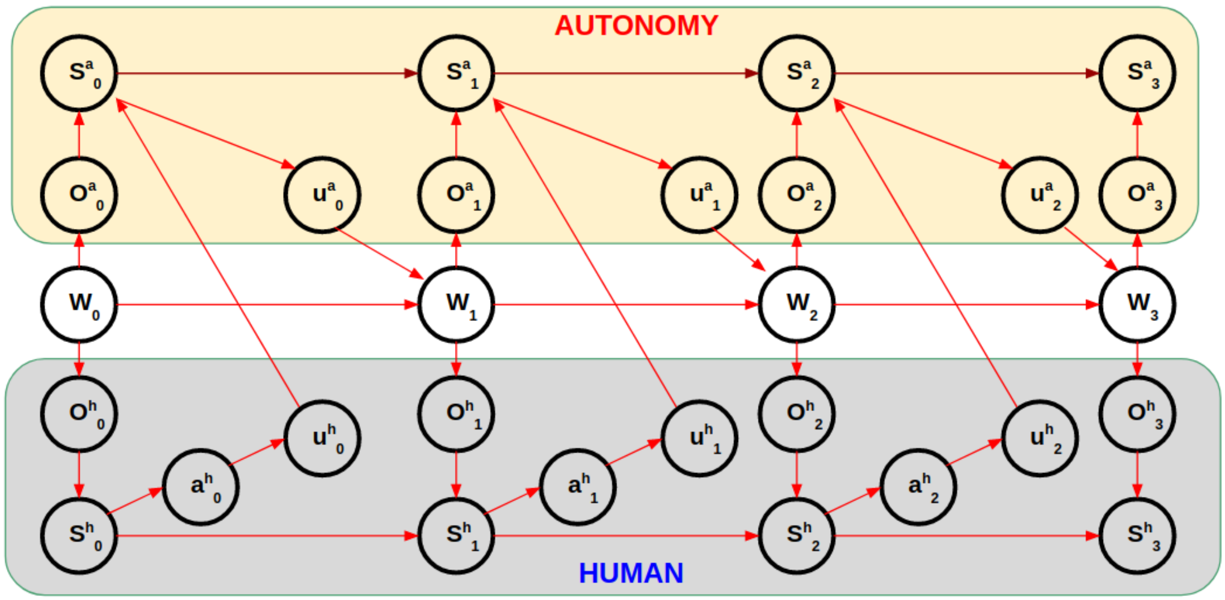

Figure 1: Our model of Human-Autonomy Interaction as a coupled perception-action loop that unfolds in time represented as a Causal Bayesian Network. The nodes represent various relevant variables that interact with each other at discrete time steps. $\mathbf{w}_t$ refers to the world state (includes the robot). $\mathbf{s}_t^h$ and $\mathbf{s}_t^a$ denote the internal state of the human and autonomy respectively. $\mathbf{o}_t^h$ and $\mathbf{o}_t^a$ refer to the noisy observations of the true robot state that are accessible to the human and autonomy respectively. $\mathbf{a}_t^h$ represents the action taken by the human and $\mathbf{u}_t^h$ denotes the human control command as filtered through a control interface such as a joystick. $\mathbf{u}_t^a$ represents the autonomy's control command. The evolution of robot state is governed by the stochastic dynamics of the environment. Note that, this CBN represents one of the many different ways in which autonomy and human interact.

perception-action loops (Figure 1). Within this coupled system, the environment of one of the agents subsumes the other agent(s).

Causal Bayesian Networks (CBN) provide a systematic mathematical framework to model the time dynamics of a coupled perception-action loop. The nodes of the network represent the relevant variables pertaining to both human and autonomy (latent and observed), and the edges represent the probabilistic influence they have on each other [59]. Once modeled as a CBN, the human-autonomy system lends itself perfectly to information-theoretic analysis by which we can quantify the *information dynamics* that unfolds during HAI in a mathematically concrete fashion. Once information flow is quantified, design of autonomy can be accomplished with an aim to *shape* the information flow in the joint system towards desired specifications to achieve desired outcomes. By characterizing the information dynamics in HAI using the proposed framework, we seek to:

1. Quantify colloquial notions of transparency, cooperation, coordination etc., that are relevant to HAI using information-theoretic notions of directed information flow and transfer entropy [64]. These measures can be computed directly from the statistics of data without the need for domain- and task-specific metrics and models. Information-theoretic measures such as transfer entropy and predictive information 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 a complex system [3].

2. Develop a framework that will provide a systematic and principled approach to design of autonomy, in which autonomy's actions are interpreted as appropriately timed *interventions* with an aim to modulate the bi-directional information flow between the human and autonomy thereby facilitating faster inference, skill acquisition and enhanced task performance.

3. Bring different aspects of HAI, such as inference, skill acquisition, task performance, transparency and cooperation under a single umbrella in order to shed light on the more fundamental and low-level descriptors and characteristics of human-autonomy teaming.

3.2 Primer on Information Theory

In this subsection we describe some of the fundamental information-theoretic quantities that are essential for the quantification of information flow between nodes of a causal Bayesian Network.

3.2.1 Entropy and Mutual Information

The most fundamental quantity in information theory is *entropy*. For a discrete random variable X the entropy, $H(X)$ is given by

$$H(X) = -\sum_{x\in\Omega_x} p(x)\log_2 p(x)$$

where $p(x)$ is the probability mass distribution and the summation extends over all possible states the random variable can assume (Ω_x is the state space of X). Entropy can be interpreted as the average uncertainty in the value of a sample of a variable. The above definition of entropy for a single random variable can be extended to two variables in a natural way. For random variables X and Y the *joint* entropy is defined as

$$H(X,Y) = -\sum_{x\in\Omega_x}\sum_{y\in\Omega_y} p(x,y)\log_2 p(x,y)$$

where $p(x,y)$ is the joint probability distribution and the summation is over all possible values that (x,y) can acquire. Ω_x and Ω_y denote the state space of the X and Y respectively. This definition can be extended in a similar fashion to an arbitrary number of variables.

Closely related is also the idea of *conditional* entropy, which is the entropy of a random variable after we have taken some context into account. The conditional entropy of random variable X given Y is defined as

$$H(X|Y) = -\sum_{x\in\Omega_x}\sum_{y\in\Omega_y} p(x,y)\log_2 p(x|y)$$

Yet another important information-theoretic quantity of interest is *mutual information*. Mutual information is the amount of information *shared* between two random variables X and Y and can be interpreted as the statistical dependence between them. The mutual information $I(X;Y)$ is defined as

$$\begin{aligned} I(X;Y) &= \sum_{x\in\Omega_x}\sum_{y\in\Omega_y} p(x,y)\log_2\frac{p(x,y)}{p(x)p(y)} \\ &= H(X) - H(X|Y) = H(Y) - H(Y|X) \end{aligned}$$

and can be interpreted as the KL divergence of the product of the marginal distributions from the joint distribution. Furthermore, mutual information is symmetric in its arguments.

Not surprisingly, the *conditional mutual information*, an information measure crucial for the computation of transfer entropy, is the shared information between two random variables X and Y in the context of a third random variable Z. It is given by

$$\begin{aligned} I(X;Y|Z) &= \sum_{x\in\Omega_x}\sum_{y\in\Omega_y}\sum_{z\in\Omega_z} p(x,y,z)\log_2\frac{p(x,y,z)p(z)}{p(x,z)p(y,z)} \\ &= \sum_{x\in\Omega_x}\sum_{y\in\Omega_y}\sum_{z\in\Omega_z} p(x,y,z)\log_2\frac{p(x|y,z)}{p(x|z)} \\ &= H(X|Z) - H(X|Y,Z). \end{aligned}$$

3.2.2 Transfer Entropy

All the above mentioned measures deal with static random variables. If we want to investigate dynamics of random time-series processes, transition probabilities need to be considered. Let X be a random process denoted by $\{\ldots, X_{n-1}, X_n, X_{n+1}, \ldots\}$ with specific instantiations of the random process denoted by $\{\ldots, x_{n-1}, x_n, x_{n+1}, \ldots\}$, where n denotes a discrete countable index (in many cases, the index is discretized time). Let $\boldsymbol{X}_n^{(k)} = \{X_{n-k+1}, \ldots, X_{n-1}, X_n\}$ denote the k consecutive variables of X which has specific instantiations denoted as $\boldsymbol{x}_n^{(k)} = \{x_{n-k+1}, \ldots, x_{n-1}, x_n\}$.[2] Now let us consider another random process denoted by Y. We shall refer to Y as the source process and X as the target process.

Transfer entropy captures the notion of information transfer between Y and X, as the amount of information that the source process provides about X_{n+1} (the target's next state) after considering the target's past states. Therefore, transfer entropy is a directional measure and is asymmetric with respect to the two random processes X and Y. The transfer entropy is defined as

$$TE^{k,l}_{Y\to X}(n) = \sum p(x_{n+1}, \boldsymbol{x}_n^{(k)}, \boldsymbol{y}_n^{(l)})\log\frac{p(x_{n+1}|\boldsymbol{x}_n^{(k)}, \boldsymbol{y}_n^{(l)})}{p(x_{n+1}|\boldsymbol{x}_n^{(k)})}$$

where k and l are the embedding dimensions of the target and the source respectively and the summations is over all possible joint configurations for $x_{n+1}, \boldsymbol{x}_n^{(k)}$ and $\boldsymbol{y}_n^{(l)}$. Transfer entropy can be expressed in terms of conditional mutual information between the variables as

$$TE^{k,l}_{Y\to X}(n) = I(X_{n+1}; \boldsymbol{Y}_n^{(l)}|\boldsymbol{X}_n^{(k)}) \tag{1}$$

[2] $\boldsymbol{x}_n^{(k)}$ are the Takens' *embedding vectors* whose *embedding* dimension is k and represents the *state* of the k^{th} order Markov process. This is due to the Taken's delay embedding theorem which allows for the reconstruction of the underlying state representation of a dynamical system from the time series data.

where I is the mutual information.

This particular mathematical form of transfer entropy is important for computational purposes. As is obvious from the equations abov,e transfer entropy values depend on the embedding dimensions. Embedding dimensions are selected to ensure that active information storage (information contained in the past states of the target) is eliminated properly and is not counted in the transfer entropy computation.

3.2.3 Estimation from Data

Although the mathematical expressions for all the information-theoretic measures discussed so far in this paper are relatively straightforward and interpretable there are various computational issues that arise in practice. First of all, the probability densities contained in each of the measures need to be empirically estimated from a finite number of data samples obtained from the time-series of the random process of interest. Any such estimator is prone to bias and variance due to the limited number of samples available. This problem is exacerbated for continuous valued random variables.

One approach to estimate the relevant probability densities is to use *kernel estimators* [65]. The joint probability densities are estimated using a *kernel function* denoted by Θ which measures the 'closeness' of pairs of samples. For example, we can estimate the joint density of two variables as

$$\hat{p}_r(x_n, y_n) = \frac{1}{N} \sum_{n'=1}^{N} \Theta\left(\left\| \begin{bmatrix} x_n - x_{n'} \\ y_n - y_{n'} \end{bmatrix} \right\| - r \right)$$

where N is the total number of samples, r is the kernel width and Θ is a step kernel such that $\Theta(x > 0) = 0$ and $\Theta(x \leq 0) = 1$ and $\|\cdot\|$ is the maximum norm distance function otherwise known as the Chebyshev distance. Kernel estimators are model-free and therefore can be utilized to measure nonlinear relationships.

An improvement upon the kernel-based estimation approach for mutual information was proposed in [46]. This approach combines various techniques that are designed to reduce the bias and variance errors that can occur due to small sample sizes and relies on a nearest-neighbors approach which is effectively equivalent to dynamically changing the kernel width to the density of the samples.

In the following chapter, I will describe the studies that will be conducted as a part of this thesis.

4 Proposed Contributions

4.1 Information Theoretic Quantification of Transparency

4.1.1 Introduction

Transparency in HAI supports efficient, flexible and coordinated interaction and facilitates higher overall team performance [49]. One of the interpretations of transparency concerns the *observability* and *predictability* of an agent's behavior [20]. During HAI in a shared-control system, the agents participating in joint task execution continually perform different types of actions that affect the environment. Some actions are more transparent than others—that is, some actions make the agent's underlying decision-making logic, internal states and goals more clear to an observer than others [70].

Although transparency has been recognized as a critical component for successful HAI, mathematical quantification of transparency is still an open question. One of the main goals of this study is to quantify transparency in information-theoretic terms directly from sensor data and validate it against humans' perception of transparent behavior. This characterization will help in the design of control policies which reason explicitly about softer aspects of human-robot teaming and will likely result in better user satisfaction and acceptance. More specifically, we will use the notion of multivariate transfer entropy to characterize the information flow between the nodes of the Bayesian network shown in Figure 1. Transfer entropy is an information-theoretic metric that aims to capture the *directed information flow/transfer* from a source random process to a target random process [64]. Higher transfer entropy implies that the knowledge of the source process's past state(s) improves the predictability of the target process's future state.

In any task in which the agents have well-defined goals or shared goals, the improvement of transparency (and *consequently* the predictability) of one of the agent's actions can result in increased predictability of other agents' actions as well. That is, under rationality assumptions, if one agent makes an attempt to be more transparent about its internal state and intentions, the other agent(s) will likely make use of available information from transparent behavior to act more predictably. Higher transparency implies higher predictability of actions and vice versa.

4.1.2 Contributions

The main contributions of this study will be:

1. Introduction of a theoretical framework for HAI based on causal Bayesian Networks in order to reason about information flow between nodes of the network to characterize fluency, coordination and transparency of HAI.

2. An information-theoretic characterization of transparency in human-autonomy interaction directly from sensor data utilizing the notion of *transfer entropy*.

4.1.3 Approach

In order to investigate the potential usefulness of information-theoretic ideas to characterize transparency levels in HAI, in this study we exclusively focus on *autonomy-to-human transparency*. In

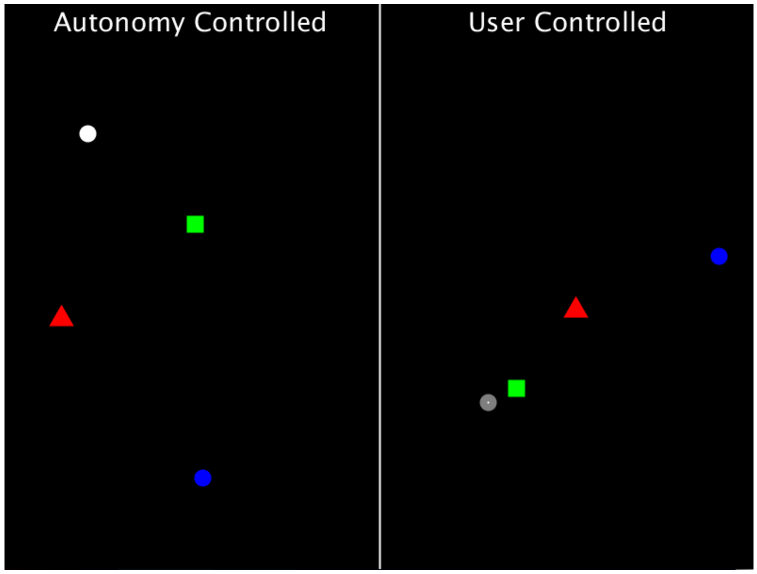

Figure 2: Simulated 2D point robot environment. Left: The autonomy controlled robot is represented as a white solid circle. The goals are represented in different colors and shapes (green square, red triangle and blue circle). Right: The human controlled robot is represented as a solid gray circle. The different goals are represented in different colors and shapes (green square, red triangle and blue circle). For each trial, the autonomy controlled robot will move towards one of the two goals (unknown to the user). The user has to infer what the autonomy's goal and teleoperate the human-controlled robot to the similar goal.

information-theoretic terms, we define autonomy-to-human transparency at time t as

$$\text{transparency}_{A\to H}(t) = TE^{k,l}_{A\to H}(t) \tag{2}$$

where k and l are the embedding dimensions of the source process (A) and the target process (H) respectively where TE is defined as in Equation 1. It has to be noted that autonomy could use different modalities to improve transparency. However, in this study we focus only on robot motion, that is, the autonomy uses *robot motion* as a way to implicitly communicate its intentions and goals. We carefully manipulate two independent factors that directly affect robot motion (and consequently transparency) and observe how they affect the predictability of human actions.

Note that transfer entropy computation is carried out between two time series signals. In the proposed definition of transparency in Equation 2, A and H refer to different time series signals corresponding to sensor measurements associated with the autonomy and human. These will be dependent on the task and modality used for expressing intent. For example, if the autonomy only uses *robot motion* as a way to communicate its intent we can use the time series associated with robot trajectory, end-effector velocity, directedness towards the end-goal etc., as the source process and time series associated with the user control commands as the target process.

4.1.4 Experimental Design

The experiment consists of a human subject teleoperating a point robot in 2D Cartesian space (Figure 2, Right) while observing the behavior of another autonomous 2D point robot in a reaching task (Figure 2, Left). The goals are represented in different colors and shapes and the number of goals could vary from two to four in each trial. Note that in Figure 2 the goal locations differ between the autonomy's display window (left) versus the human display window (right), in order to ensure that the subject does not simply mimic the autonomy controlled robot motion. In each trial the autonomy has an internal goal (a reaching target) unknown to the subject. In each trial, the subject has to infer the autonomy's goal and teleoperate the robot towards the correspond goal in the right half as quickly as possible.

Transparency of autonomy's actions directly depends on the legibility of robot motion and therefore maximum transparency is achieved when the autonomy drives the robot along a legible path towards the goal without any interruptions.

In order to modulate the transparency of autonomy's actions we systematically manipulate two independent factors:

1. **Signal dropout**: This factor controls the percentage of the total trial time that the robot is stalled due to zero control commands due to which the robot motion becomes discontinuous.

2. **Direction of motion**: Any deviation from a legible path can potentially affect the transparency of autonomy's actions. Non-legible paths will be created by using higher order Bezier curves.

These independent factors were chosen so that they directly affect the legibility of motion thereby affecting autonomy-to-human transparency.

At the end of each trial, subjective evaluation of transparency of autonomy's actions will also be collected via questionnaires so that correlation analysis can be performed between the transfer entropy-based transparency measure and the subject's perception of transparency.

Before the testing phase, subjects will undergo a short training phase during which they will observe unimpeded legible robot motion towards various targets on the table. This is done to ensure that the human has a sufficiently good model of how the autonomy will accomplish the reaching task and can be used to establish a baseline for what constitutes transparent behavior. The core hypothesis we seek to test in this study is that *in trials with higher transparency, the user actions (control commands issued via the interface) will be more predictable and therefore transfer entropy from robot state to user command will be higher.*

4.2 Autonomy Guided Skill Acquisition in Humans

4.2.1 Introduction

In the context of shared-control assistive robotics, the users typically maintain some amount of robot control at all times. The primary function of the assistance provided by the autonomy is to supplement or in some cases complement the users' capabilities in order to improve overall task performance. If the users are not adept at controlling the robot on their own (after considering the constraints they have due to motor impairments or control interface), it can affect their capability to express their intent properly which in turn can negatively impact *human-to-autonomy* transparency. That is, it is likely beneficial for the overall system that the users have a high level of skill in teleoperating a robot. However, high-dimensional robot control using low-dimensional control interfaces is challenging. Users gain experience and progressively become better at robot control via trial and error and practice.

Our previous studies [25, 23] have shown us that when humans learn to control a robotic arm using a standard control interface, they are faced with different types of learning problems, some of which are,

1. They need to learn how to physically activate the control interface. For example, how to ‘press’ a button, how to ‘push’ a joystick *et cetera*.
2. They need to learn what the control mappings are. For example, what is the direction of robot end-effector motion upon pushing the joystick forward? Learning control mappings involve implicit learning of the reference coordinate systems as well as the associated coordinate transformations.
3. They need to learn a forward model that predicts the next state of the robot given the current state and action taken and an inverse model that can generate an action given the current state and a desired next state. This is essential if the robot is to be used for performing specific tasks.

Robot control skill acquisition that happens via self-guided training process can potentially have the following issues:

1. During the training phase users might not necessarily explore different parts of the robot workspace in a proper manner. As a result, the idiosyncracies of robot kinematics (such as nonlinearities and singularities) will not be understood properly. Nonlinearities dominate the kinematics typically near the edges of the workspace or near self-collision configurations.
2. It is also likely that during self-guided training users do not experience what we define as *cognitively hard states*. We define *cognitively hard states* as those in which the inverse control problem takes more cognitive and computational resources. The difficulty likely arises due to the complexity of coordinate transformations involved and the nonlinearities in kinematics.

For example, familiarity with the robot and knowledge about the dynamics of the control interface and the robot increase gradually with training and practice [57]. The initial forward (and inverse) kinematics (or dynamics) model that the user maintains internally could be drastically different from the true underlying robot kinematics (or dynamics). As a result of extensive practice, the internal model will likely become closer to the true model. However, the learning strategies that humans adopt need not always be optimal, for example, users might not explore the state and action space in an efficient and exhaustive manner and therefore can erroneously extrapolate the learned internal model between different regions of the workspace.

In this study, I will explore *how autonomy can help humans in skill acquisition, specifically robot teleoperation using low-dimensional control interfaces*. I am inspired by ideas in *curriculum learning* in which a learner (an animal or an artificial machine learning system) learns about a training distribution or a hypothesis by exploring a set of examples following a systematic curriculum, typically by experiencing ‘easy’ examples first followed by the harder ones.

4.2.2 Contributions

The main contributions of this study will be

1. Utilization of information density maps and ergodicity with respect to these maps to characterize how users explore the state space during training and task execution to reveal the difficulties faced by users during robot teleoperation.

2. An algorithm to generate an iterative training curriculum in order to make the training phase more efficient for better experiments and faster skill acquisition.

3. Introducing the role of autonomy as an informative teacher (or a coach) that can help the users with maximal skill acquisition.

4.2.3 Approach

In this framework, we view the process of learning to control a robot as a special case of motor learning in which there is a flow of information from the unknown function (nonlinear kinematic function that governs the time evolution of the robot's kinematic state) to the user's internal representation of the function. The goal of the algorithm is to present the user with examples such the user gain maximal information regarding the underlying unknown function (or in other words, will maximally reduce the entropy). The algorithm will drive the user's training trajectory towards a) those parts of the state-space that are typically not explored during self-guided training, b) to those regions which users typically find to be cognitively harder during task execution and finally c) to those regions which are accessed by expert users during task execution. That is, the goal of autonomy is to utilize a data-driven approach to identify those regions of the state space where the humans require more practice to improve their skill in teleoperating the robot.

The first step towards the design of a teaching curriculum is to quantify how self-guided learning and subsequent task execution proceeds and identify room for improvement. To this end, our first goal is to understand the temporal and spatial statistics of how different subjects navigate the robot's workspace during self-guided training and task execution. Additionally, we also want a novice user to gradually learn how an expert user typically goes about task execution.

First, we will use a collection of observations (of state-space trajectories) from multiple instances of self-guided training from different users to generate a *training-phase trajectory set* denoted as ϕ_u^{train}. Second, a *task-phase trajectory set* will be generated from data collected from the subjects during task execution (denoted as ϕ_u^{task}). 'Cognitively hard states' will correspond to regions with higher density as it is likely that users will have spent more time in such states due to increased cognitive load. Third, we will also use information maps (ϕ_{expert}^{task}) generated from trajectory data collected from expert users during task execution as well. ϕ_{expert}^{task} will be used a baseline to evaluate how different the user is from an expert.

If U denotes the set of points (or continuous regions) that constitute the entire task space, then the set given by $U \setminus \phi_u^{train}$ corresponds to those regions of the task space that the users neglected during self-guided training. Likewise, $\phi_u^{task} \cup \phi_{expert}^{task}$ captures those regions visited by the user during task execution (with higher density of points in cognitively hard states) and those regions where the user needs to practice to become an expert.

The *combined trajectory set* given by $(U \setminus \phi_u^{train}) \cup \phi_u^{task} \cup \phi_{expert}^{task}$ corresponds to all regions of the workspace where the subject needs more practice during training. The autonomy will sample different 'practice zones' from this combined trajectory set and will nudge the user to practice robot control in and around, these sampled practice zones. The *ergodicity* of the user's training trajectory with respect to the combined probability density can be used as a proxy to the amount of practice that the user had.[3] By having the user spend more time practicing robot control in state space regions that are typically not explored during self-guided training and are cognitively hard (as informed by the combined information density map), the hypothesis is that the users will gain a better understanding of robot kinematics and control throughout the workspace and thereby will be able to perform a wider variety of tasks more skillfully which in turn will improve human-to-autonomy transparency.

4.2.4 Experiment Design

The information density maps described in the previous section will be generated offline by combining state-space trajectory data during training and testing from multiple subjects. These information density maps encode the average behavior of how humans typically navigate the task space during self-guided training and task execution.

For the study, the subjects will be divided into two groups, A and B. Subjects in Group A will undergo a training phase in which the practice zones will be randomly sampled from a uniform density map. The practice zones for Group B will be sampled according to the combined information density map as described in the previous section.

After the training phase, the subjects will undergo a testing phase during which they will perform a number of predefined reaching tasks that will require complex control of the robotic arm.

The main idea is that *subjects in Group B will outperform (with respect to to various objective task-related metrics such as task completion time, number of mode switches performed et cetera) subjects in Group A during the testing phase due to being more skilled in robot control as a result of extensive practice according to the training curriculum chosen by the autonomy.*

4.3 Optimizing Task Performance

4.3.1 Introduction

Design of autonomy for a shared-control system ideally should focus on HAI that improves task performance and safety concurrently with an increase of user satisfaction and acceptance. One of the most critical factors that contributes to overall performance of a human-robot team is transparency of interaction [40]. That is, knowledge of *why*, *how* and *what* each other agent is doing can lead to better interaction.

Typical design of autonomy for shared-control systems aims to improve various objective aspects of joint task performance, such as task completion times [30], energy consumption [78] and inference accuracy [36] [35]. Optimization-based techniques are used to derive autonomy's policy in which the cost functions that capture desired task behavior are pre-specified by the system designer [36]. However, in the domain of assistive shared-control, subjective metrics such as user satisfaction, comfort and trust are also of paramount importance for successful adoption of these technologies [41]. Determining the exact mathematical structure for the cost function that incorporates these subjective metrics is likely an intractable problem [25].

In this study, we seek to concurrently improve different aspects of HAI (both subjective and objective task metrics) by the optimization of a more low-level aspect of interaction which is that of *transparency* in addition to objective task-specific metrics. The hypothesis is that optimization

[3] The combined probability density can be estimated using various kernel density estimation techniques such as Parzen Window method or Gaussian Mixture Models *et cetera.*

of bi-directional transparency will likely result in better communication of latent internal states. By leveraging the fact that human actions can be influenced by robot actions, robot policy can be designed in such a way that it drives the joint human-autonomy system to states with higher transparency. This will result in a common ground [40] for joint task execution, which will lead to enhanced cooperation, coordination and mutual trust. As a result of these enhancements, the desired objective and subjective outcomes will naturally emerge.

4.3.2 Approach

In this study the generation of autonomy policy is framed as an optimal control problem. More formally, the goal of this optimal control problem is to determine an autonomy control signal $\boldsymbol{u}_r^t \;\; \forall \;\; t \in [0, T]$ such that the quantity $\int\limits_0^T transparency_{A \to H}(t) - \int\limits_0^T cost(t)$ is maximized, where $cost(t)$ encodes the objective task-specific cost that the autonomy is trying to minimize (for example, path distance to goal, time to goal, number of mode switches *et cetera*). The solution to this optimal control problem is an autonomy signal that will simultaneously tries to maximize autonomy-to-human transparency and minimize task-specific cost. This optimization procedure will be implemented using an iterative Model Predictive Control (MPC) framework. A pre-requisite to perform MPC is well-defined models.

First, we need a robot dynamics/kinematics model given by $\boldsymbol{x}_r^{t+1} = f_{\boldsymbol{\Theta}}(\boldsymbol{x}_r^t, \boldsymbol{u}_f^t)$, where $\boldsymbol{x}_r^t$ is the robot state and $\boldsymbol{u}_f^t$ refers to the final control command issued to the robot. For simple robots, the kinematics could possibly be expressed as analytical functions. For more complex robots, the kinematic models could be learned directly from data using different techniques such as deep neural networks or Koopman Operators.

Second, we need a model for the human policy. In this work we are specifically interested in a model for human actions in the presence of autonomy. This model can be thought of an inverse controller that the human implements to generate actions in a given context and can be written as $\boldsymbol{u}_h^t = p_{\boldsymbol{\eta}}(\boldsymbol{x}_r^t, \boldsymbol{u}_r^t, g^t)$ (and possibly history of these variables as well) where $\boldsymbol{\eta}$ represents the model parameters, $\boldsymbol{u}_h^t$ is the human control command, $\boldsymbol{u}_r^t$ is the autonomy control command and g^t is the human's intended goal at time t. Similar to the robot kinematics model, the human action prediction model can also be learned offline using data-driven methods.

Lastly, we need an arbitration function that will determine how the user control command ($\boldsymbol{u}_h^t$) and autonomy control command ($\boldsymbol{u}_r^t$) will be combined. The arbitration function β will generate a final control command $\boldsymbol{u}_f^t = \beta_{\boldsymbol{\alpha}}(\boldsymbol{u}_r^t, \boldsymbol{u}_h^t)$, where $\boldsymbol{\alpha}$ represents the arbitration function parameters. $\boldsymbol{\alpha}$ could either be fixed or could be optimized for concurrently with the main objective.

4.3.3 Contributions

The contributions of this study is two-fold:

1. Development of an integrated pipeline in which autonomy explicitly reasons about the control policy deployed by humans when teleoperating robots in the presence of autonomy,

2. Development of a control policy that optimizes transparency in HAI in addition to task-specific objective metrics.

4.3.4 Experiment Design

The robot kinematics model (f_{Θ}) and the human action prediction model ($p_{\boldsymbol{\eta}}$) will be trained offline prior to individual subject sessions. Model performance will be evaluated using standard cross validation techniques that are widely employed in the domain of machine learning.

Each subject session will consist of a training phase and a testing phase. During the training phase, the subjects will undergo a training procedure as outlined in Section 4.2.3 during which the subject will get accustomed to the control interface, the control mappings and gain sufficient skill in robot control. During each session, subjects will control the assistive robotic arm to perform simple reaching tasks in a shared-control setting. Two conditions will be tested: (condition A) in which the autonomy signal is generated using a simple distance based optimizer (straight line potential field) and (condition B) in which the optimization procedure described in the previous subsection will be used to generate the autonomy signal.

We seek to investigate how task related objective metrics as well as subjective metrics differ between the conditions A and B. The main hypothesis is that *in trials under condition B the perceived transparency levels will be higher and as a result subjective metrics such as user satisfaction will likely be higher.* In order to evaluate whether the optimization procedure resulted in generating autonomous control signals that increased transparency we will also collect data regarding perceived transparency levels by having the subjects fill out a survey after every trial.

5 Timeline

5.1 Plan for Completion of Research

Timeline	Work	Progress
May 2019	RAL-IROS Submission	Ongoing
November 2019	Thesis Proposal	Ongoing
September 2019 - Jan 2020	Section 4.1 Paper Submission	Ongoing
Jan 2020 - April 2020	Section 4.2 Conference Submission	Planning
April 2020 - December 2020	Section 4.3 Conference + Journal Submission	Planning
January 2021	Begin Thesis writing	
April 2021	Submit Draft of thesis	
May 2021	Revise and Defend!	

References

[1] David A Abbink, Tom Carlson, Mark Mulder, Joost CF de Winter, Farzad Aminravan, Tricia L Gibo, and Erwin R Boer. A topology of shared control systemsfinding common ground in diversity. *IEEE Transactions on Human-Machine Systems*, (99):1–17, 2018.

[2] Brenna D Argall, Sonia Chernova, Manuela Veloso, and Brett Browning. A survey of robot learning from demonstration. *Robotics and Autonomous Systems*, 57(5):469–483, 2009.

[3] Nihat Ay and Daniel Polani. Information flows in causal networks. *Advances in complex systems*, 11(01):17–41, 2008.

[4] Yoshua Bengio, Jérôme Louradour, Ronan Collobert, and Jason Weston. Curriculum learning. In *Proceedings of the 26th annual international conference on machine learning*, pages 41–48. ACM, 2009.

[5] Mariusz Bojarski, Davide Del Testa, Daniel Dworakowski, Bernhard Firner, Beat Flepp, Prasoon Goyal, Lawrence D Jackel, Mathew Monfort, Urs Muller, Jiakai Zhang, et al. End to end learning for self-driving cars. *arXiv preprint arXiv:1604.07316*, 2016.

[6] Michael E Bratman. Shared cooperative activity. *The philosophical review*, 101(2):327–341, 1992.

[7] Alexander Broad, Jacob Arkin, Nathan Ratliff, Thomas Howard, Brenna Argall, and Distributed Correspondence Graph. Towards real-time natural language corrections for assistive robots. In *RSS Workshop on Model Learning for Human-Robot Communication*, 2016.

[8] Alexander Broad, Todd Murphey, and Brenna Argall. Learning models for shared control of human-machine systems with unknown dynamics. *arXiv preprint arXiv:1808.08268*, 2018.

[9] Min Chen, Stefanos Nikolaidis, Harold Soh, David Hsu, and Siddhartha Srinivasa. Planning with trust for human-robot collaboration. In *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, pages 307–315. ACM, 2018.

[10] Young Sang Choi, Cressel D. Anderson, Jonathan D. Glass, and Charles C. Kemp. Laser pointers and a touch screen: intuitive interfaces for autonomous mobile manipulation for the motor impaired. In *Proceedings of the International SIGACCESS Conference on Computers and Accessibility*, 2008.

[11] HH Clark. Using language. new york, ny, us, 1996.

[12] Thomas M Cover and Joy A Thomas. *Elements of information theory*. John Wiley & Sons, 2012.

[13] Vassilis Cutsuridis. Cognitive models of the perception-action cycle: A view from the brain. In *The 2013 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8. IEEE, 2013.

[14] Eric Demeester, Alexander Hüntemann, Dirk Vanhooydonck, Gerolf Vanacker, Hendrik Van Brussel, and Marnix Nuttin. User-adapted plan recognition and user-adapted shared control: A bayesian approach to semi-autonomous wheelchair driving. *Autonomous Robots*, 24(2):193–211, 2008.

[15] Anca D. Dragan, Kenton CT. Lee, and Siddhartha S. Srinivasa. Legibility and predictability of robot motion. In *Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 2013.

[16] Anca D. Dragan and Siddhartha S. Srinivasa. *Formalizing assistive teleoperation*. MIT Press, 2012.

[17] Anca D. Dragan and Siddhartha S. Srinivasa. A policy-blending formalism for shared control. *The International Journal of Robotics Research*, 32(7):790–805, 2013.

[18] Krishnamurthy Dvijotham and Emanuel Todorov. Inverse optimal control with linearly-solvable mdps. In *Proceedings of the 27th International Conference on Machine Learning (ICML-10)*, pages 335–342, 2010.

[19] Gopinath Deepak Egli, Pascal Arturo and Brenna Argall. A call for convergence of research directions in assistive robotics. In *RSS Workshop on Socially and Physically Assistive Robotics for Humanity, Ann Arbor, Michigan, USA*.

[20] Mica R Endsley. From here to autonomy: lessons learned from human–automation research. *Human Factors*, 59(1):5–27, 2017.

[21] David Fischinger, Peter Einramhof, Konstantinos Papoutsakis, Walter Wohlkinger, Peter Mayer, Paul Panek, Stefan Hofmann, Tobias Koertner, Astrid Weiss, Antonis Argyros, et al. Hobbit, a care robot supporting independent living at home: First prototype and lessons learned. *Robotics and Autonomous Systems*, 75:60–78, 2016.

[22] Aditya Goil, Matthew Derry, and Brenna D Argall. Using machine learning to blend human and robot controls for assisted wheelchair navigation. In *Rehabilitation Robotics (ICORR), 2013 IEEE International Conference on*, pages 1–6. IEEE, 2013.

[23] Deepak Gopinath and Brenna Argall. Mode switch assistance to maximize human intent disambiguation. In *Robotics: Science and Systems*, 2017.

[24] Deepak Gopinath and Brenna D Argall. Dynamic neural fields for short-term behavior recognition from motion cues. In *Proceedings of Human-Robot Interaction 2018 Workshop on Social Human-Robot Interaction of Human-care Service Robots (HRI), Chicago, Illinois*. ACM, 2018.

[25] Deepak Gopinath, Siddarth Jain, and Brenna D. Argall. Human-in-the-loop optimization of shared autonomy in assistive robotics. *IEEE Robotics and Automation Letters*, 2(1):247–254, 2017.

[26] Deepak Gopinath and Gil Weinberg. A generative physical model approach for enhancing the stroke palette for robotic drummers. *Robotics and Autonomous Systems*, 86:207–215, 2016.

[27] Peter J Green. Reversible jump markov chain monte carlo computation and bayesian model determination. *Biometrika*, 82(4):711–732, 1995.

[28] Shixiang Shane Gu, Zoubin Ghahramani, and Richard E Turner. Neural adaptive sequential monte carlo. In *Advances in Neural Information Processing Systems*, pages 2629–2637, 2015.

[29] Stephen J Guastello and Denise D Guastello. Origins of coordination and team effectiveness: A perspective from game theory and nonlinear dynamics. *Journal of applied psychology*, 83(3):423, 1998.

[30] Laura V. Herlant, Rachel M. Holladay, and Siddhartha S. Srinivasa. Assistive teleoperation of robot arms via automatic time-optimal mode switching. In *Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 2016.

[31] Laura M Hiatt, Cody Narber, Esube Bekele, Sangeet S Khemlani, and J Gregory Trafton. Human modeling for human–robot collaboration. *The International Journal of Robotics Research*, 36(5-7):580–596, 2017.

[32] Jean-Michel Hoc. Towards a cognitive approach to human–machine cooperation in dynamic situations. *International journal of human-computer studies*, 54(4):509–540, 2001.

[33] Guy Hoffman and Cynthia Breazeal. Cost-based anticipatory action selection for human–robot fluency. *IEEE transactions on robotics*, 23(5):952–961, 2007.

[34] Chien-Ming Huang, Sean Andrist, Allison Sauppé, and Bilge Mutlu. Using gaze patterns to predict task intent in collaboration. *Frontiers in psychology*, 6:1049, 2015.

[35] Siddarth Jain and Brenna Argall. Robot learning to switch control modes for assistive teleoperation. In *RSS Workshop on Planning for Human-Robot Interaction: Shared Autonomy and Collaborative Robotics*.

[36] Shervin Javdani, Siddhartha S Srinivasa, and J Andrew Bagnell. Shared autonomy via hindsight optimization. *arXiv preprint arXiv:1503.07619*, 2015.

[37] Lydia E Kavraki, Mihail N Kolountzakis, and J-C Latombe. Analysis of probabilistic roadmaps for path planning. In *Robotics and Automation, 1996. Proceedings., 1996 IEEE International Conference on*, volume 4, pages 3020–3025. IEEE, 1996.

[38] Richard Kelley, Alireza Tavakkoli, Christopher King, Monica Nicolescu, Mircea Nicolescu, and George Bebis. Understanding human intentions via hidden markov models in autonomous mobile robots. In *Proceedings of the 3rd ACM/IEEE International Conference on Human Robot Interaction*, pages 367–374. ACM, 2008.

[39] Oussama Khatib. Real-time obstacle avoidance for manipulators and mobile robots. *The International Journal of Robotics Research*, 5(1):90–98, 1986.

[40] Sara Kiesler. Fostering common ground in human-robot interaction. In *Robot and Human Interactive Communication, 2005. ROMAN 2005. IEEE International Workshop on*, pages 729–734. IEEE, 2005.

[41] Dae-Jin Kim, Rebekah Hazlett-Knudsen, Heather Culver-Godfrey, Greta Rucks, Tara Cunningham, David Portee, John Bricout, Zhao Wang, and Aman Behal. How autonomy impacts performance and satisfaction: Results from a study with spinal cord injured subjects using an assistive robot. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 42(1):2–14, 2012.

[42] Donald E Kirk. *Optimal control theory: an introduction*. Springer, 1970.

[43] Glen Klien, David D Woods, Jeffrey M Bradshaw, Robert R Hoffman, and Paul J Feltovich. Ten challenges for making automation a" team player" in joint human-agent activity. *IEEE Intelligent Systems*, 19(6):91–95, 2004.

[44] Daphne Koller, Nir Friedman, and Francis Bach. *Probabilistic graphical models: principles and techniques*. MIT press, 2009.

[45] Steve WJ Kozlowski and Georgia T Chao. The dynamics of emergence: Cognition and cohesion in work teams. *Managerial and Decision Economics*, 33(5-6):335–354, 2012.

[46] Alexander Kraskov, Harald Stögbauer, and Peter Grassberger. Estimating mutual information. *Physical review E*, 69(6):066138, 2004.

[47] James J Kuffner and Steven M LaValle. Rrt-connect: An efficient approach to single-query path planning. In *Robotics and Automation, 2000. Proceedings. ICRA'00. IEEE International Conference on*, volume 2, pages 995–1001. IEEE, 2000.

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

[49] Joseph B Lyons. Being transparent about transparency: A model for human-robot interaction. In *2013 AAAI Spring Symposium Series*, 2013.

[50] Joseph B Lyons and Paul R Havig. Transparency in a human-machine context: Approaches for fostering shared awareness/intent. In *International Conference on Virtual, Augmented and Mixed Reality*, pages 181–190. Springer, 2014.

[51] Jim Mainprice, Nathan Ratliff, and Stefan Schaal. Warping the workspace geometry with electric potentials for motion optimization of manipulation tasks. In *Intelligent Robots and Systems (IROS), 2016 IEEE/RSJ International Conference on*, pages 3156–3163. IEEE, 2016.

[52] David Marr. Vision: A computational investigation into the human representation and processing of visual information. mit press. *Cambridge, Massachusetts*, 1982.

[53] Maja J Matarić, Jon Eriksson, David J Feil-Seifer, and Carolee J Winstein. Socially assistive robotics for post-stroke rehabilitation. *Journal of NeuroEngineering and Rehabilitation*, 4(1):5, 2007.

[54] Katharina Muelling, Arun Venkatraman, Jean-Sebastien Valois, John E Downey, Jeffrey Weiss, Shervin Javdani, Martial Hebert, Andrew B Schwartz, Jennifer L Collinger, and J Andrew Bagnell. Autonomy infused teleoperation with application to brain computer interface controlled manipulation. *Autonomous Robots*, pages 1–22, 2017.

[55] Urs Muller, Jan Ben, Eric Cosatto, Beat Flepp, and Yann L Cun. Off-road obstacle avoidance through end-to-end learning. In *Advances in neural information processing systems*, pages 739–746, 2006.

[56] Kevin Patrick Murphy and Stuart Russell. Dynamic bayesian networks: representation, inference and learning. 2002.

[57] Ferdinando A MussaIvaldi and Emilio Bizzi. Motor learning through the combination of primitives. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 2000.

[58] Stefanos Nikolaidis and Julie Shah. Human-robot teaming using shared mental models. *ACM/IEEE HRI*, 2012.

[59] Judea Pearl. *Causality*. Cambridge university press, 2009.

[60] Patrick M Pilarski, Michael R Dawson, Thomas Degris, Jason P Carey, and Richard S Sutton. Dynamic switching and real-time machine learning for improved human control of assistive biomedical robots. In *Proceedings of the IEEE RAS & EMBS International Conference on Biomedical Robotics and Biomechatronics (BioRob)* , pages 296–302. IEEE, 2012.

[61] David G Rand and Martin A Nowak. Human cooperation. *Trends in cognitive sciences*, 17(8):413–425, 2013.

[62] Nathan D Ratliff, Jan Issac, and Daniel Kappler. Riemannian motion policies. *arXiv preprint arXiv:1801.02854*, 2018.

[63] Gregor Schöner, Michael Dose, and Christoph Engels. Dynamics of behavior: Theory and applications for autonomous robot architectures. *Robotics and Autonomous Systems*, 16(2-4):213–245, 1995.

[64] Thomas Schreiber. Measuring information transfer. *Physical review letters*, 85(2):461, 2000.

[65] Bernard W Silverman. *Density estimation for statistics and data analysis*, volume 26. CRC press, 1986.

[66] Tyler Simpson, Colin Broughton, Michel JA Gauthier, and Arthur Prochazka. Tooth-click control of a hands-free computer interface. *IEEE Transactions on Biomedical Engineering*, 55(8):2050–2056, 2008.

[67] Justin G Storms and Dawn M Tilbury. Blending of human and obstacle avoidance control for a high speed mobile robot. In *American Control Conference (ACC)*, pages 3488–3493. IEEE, 2014.

[68] Tarek Taha, Jaime Valls Miró, and Gamini Dissanayake. A pomdp framework for modelling human interaction with assistive robots. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, pages 544–549. IEEE, 2011.

[69] Karim A Tahboub. Intelligent human-machine interaction based on dynamic bayesian networks probabilistic intention recognition. *Journal of Intelligent and Robotic Systems*, 45(1):31–52, 2006.

[70] Andreas Theodorou, Robert H Wortham, and Joanna J Bryson. Designing and implementing transparency for real time inspection of autonomous robots. *Connection Science*, 29(3):230–241, 2017.

[71] Andrea L Thomaz and Crystal Chao. Turn-taking based on information flow for fluent human-robot interaction. *AI Magazine*, 32(4):53–63, 2011.

[72] Naftali Tishby and Daniel Polani. Information theory of decisions and actions. In *Perception-action cycle*, pages 601–636. Springer, 2011.

[73] Michael Tomasello and Malinda Carpenter. Shared intentionality. *Developmental Science*, 10(1):121–125, 2007.

[74] Pete Trautman. Assistive planning in complex, dynamic environments: a probabilistic approach. In *Systems, Man, and Cybernetics (SMC), 2015 IEEE International Conference on*, pages 3072–3078. IEEE, 2015.

[75] Yoji Uno, Mitsuo Kawato, and Rika Suzuki. Formation and control of optimal trajectory in human multijoint arm movement. *Biological cybernetics*, 61(2):89–101, 1989.

[76] Glenn Wasson, Pradip Sheth, Majd Alwan, Kevin Granata, Alexandre Ledoux, and Cunjun Huang. User intent in a shared control framework for pedestrian mobility aids. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2003.

[77] Christopher JCH Watkins and Peter Dayan. Q-learning. *Machine learning*, 8(3-4):279–292, 1992.

[78] Juanjuan Zhang, Pieter Fiers, Kirby A Witte, Rachel W Jackson, Katherine L Poggensee, Christopher G Atkeson, and Steven H Collins. Human-in-the-loop optimization of exoskeleton assistance during walking. *Science*, 356(6344):1280—1284, 2017.

[79] Brian D Ziebart, Andrew L Maas, J Andrew Bagnell, and Anind K Dey. Maximum entropy inverse reinforcement learning. In *AAAI*, volume 8, pages 1433–1438. Chicago, IL, USA, 2008.