Towards an Information Theoretic Analysis of Human-Autonomy Interaction in Shared Autonomy

Thesis proposal

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

March 10, 2019

Abstract

Human-Robot Interaction (HRI) in the context of shared autonomy is a rich and complex phenomenon. The effectiveness and usefulness of shared-control human-machine systems critically depends on the fluency and efficacy of human-robot interaction. Efficient HRI can lead to an improvement in joint task performance with higher user satisfaction and enhanced trust, all of which are desired characteristics of a joint human-machine system. From an engineer/system designer's perspective, in order to achieve optimal performance the design of autonomy should adequately taken into account the richness, subtleties and complexity of the interaction between the human and the machine.

In this thesis proposal, I plan to propose a mathematical framework for human-robot interaction in the context of shared autonomy that utilizes 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, design of autonomy can be thought of as appropriately timed *interventions* at specific parts of model, with an intention to alter the bi-directional information flow between the human and machine. Using the proposed mathematical model, I will research three important problems that arise in HRI namely, a) **learning** b) **inference** and c) **joint task performance**. More specifically, I will focus on the information theoretic analysis of how each of the above mentioned phenomena unfolds during task execution. The eventual goal is to utilize the proposed mathematical framework to inform the design of autonomy that will help *facilitate human learning*, *improve inference accuracy* and *enhance task performance*.

Contents

1 Introduction

Robots are ubiquitous in the modern-day society and have revolutionized the relationship between man 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 [42], entertainment [23] and home robotics [18].

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 [47]. 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 [46].

The standard usage of these assistive machines, however, still relies on manual teleoperation by the human typically enacted through a control interface such as a joystick or a switch-based headarray; that is, in such scenarios robots are not endowed with any intelligence and are treated as *passive* machines that function as extensions of human motor abilities [56]. However, one of the most difficult conundrums is that greater the motor impairment of the user, the more limited the interfaces that are available for them to use. 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 [53]. In such cases, *robot autonomy*, the ability of 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 while interacting with humans in real-world scenarios [30]. However, in literature 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 [22]. In such cases, the introduction of *shared autonomy* seeks to find a middle ground between full teleoperation and autonomy by offloading only some aspects of task execution to the autonomy [65, 12].

In a shared autonomy system, the task responsibility is split between the user and autonomy with the aim of reducing human effort in accomplishing a task. Human-Robot Interation (HRI) in the context of shared autonomy is a rich and complex phenomena. The effectiveness and usefulness of shared-control human-machine systems critically depends on the quality and efficiency of human-robot interaction. 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 [28]. From an engineering perspective, design of appropriate kinds of autonomous behaviors for a shared-control system, therefore, needs to take into account the dynamics of human-robot interaction during the course of task execution [29].

Current research approaches for design of shared autonomy systems rely on various kinds of mathematical models to solve different aspects of HRI as independent subproblems and therefore suffer from generalization across tasks, robotic platforms and user types. For my thesis, I am motivated by the desire to develop a *unified* mathematical framework to analyze different aspects of HRI 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 plan to propose a mathematical framework that models HRI in the context of shared autonomy utilizing ideas from *probabilistic graphical models* [40] and *information theory* [11]. More specifically, the interaction will be modeled as *coupled perception-action loops* unfolding in time using *causal Bayesian Networks* [52]. The nodes in the network will 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, emphasis will be placed on analyzing the *information flow* between the nodes in the network [3]. Within this proposed framework of causal Bayesian networks, design of autonomy can be thought of as appropriately timed *interventions* that have the potential to alter bidirectional information flow between human and autonomy. Our hypothesis is that *information flow* is a more fundamental and low-level descriptor of interaction dynamics and joint system performance that system designers should focus on when designing autonomous behaviors. Using the proposed model, I intend to address three main subproblems relevant to shared autonomy namely, *learning*, *inference* and *task performance*.

The first research question (**RQ1**) that I will address in my work is *how can autonomy help humans learn robot dynamics better*. When a human interacts with a machine in a shared autonomy setting, both parties are continually learning about each others' intentions, plans and actions [31]. For example, for novice users familiarity with the device and knowledge about the dynamics of the control interface and the robot increase with extensive training and practice [50]. The initial forward (and inverse) dynamics model that the user maintains internally at the beginning of task execution might be drastically different from the true underlying system dynamics. Due to learning effects, 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 sample the state and action space in an efficient and exhaustive manner and therefore can erroneously extrapolate the dynamics between different regions of the workspace. Therefore, autonomy can play the role of a *teacher* and help the human in skill acquisition and provide appropriate guidance during the learning process. Potentially, this can have a significant impact in the design of training procedures for new users of assistive robots.

Inherent limitations of the control interface and motor impairments can possibly put an upper bound to skill level that can be acquired. In such scenarios, the need for autonomy for task execution becomes inevitable. However, any successful assistive robotic system needs to have a good idea of the user's needs and intentions. That is, *user intent inference* is a necessary and crucial component to ensure proper assistance [64]. Therefore, the second research question (**RQ2**) that I will address in my thesis is *how can autonomy assistance be designed so that inference becomes more accurate*. Typically, the user's internal state (desires, goals and intentions) is latent (if not fully, partially) from autonomy's perspective [35]. In a shared control setting inference is not a unidirectional phenomena. For example, from the users' perspective the internal logic with which autonomy helps them is not always explicitly known and therefore needs to be inferred as well. User satisfaction and acceptance heavily depends on the user's understanding of how the autonomy works. In this thesis, I plan to utilize the proposed mathematical model to reason about and shape the information flow from the user's internal state to autonomy to improve the inference accuracy.

In addition to facilitating learning (**RQ1**), and improving inference accuracy (**RQ2**), autonomy has to work in conjunction with the human to perform the task optimally. Therefore, the third and final research question (**RQ3**) that I hope to tackle in this thesis is *how to design autonomy assistance to ensure optimal task performance*. Typically, both subjective (user satisfaction, acceptance, trust) and objective metrics (task completion time, number of mode switches) equally inform the optimality criteria [22]. Rather than focusing on the above-mentioned metrics independently, in this thesis work I will focus on optimal bidirectional information flow between the human and autonomy. The hypothesis is that optimization of information flow between the autonomy and human will likely result in better communication of latent internal states thereby leading to a common ground for joint task execution. This will likely lead to enhanced cooperation and mutual understanding as a result of which the desired outcomes (better task performance, improved user satisfaction) will likely naturally emerge.

In summary, in this proposed thesis I intend to develop a mathematical framework to model HRI within shared autonomy and plan to research solutions to the questions presented above (**RQ1**, **RQ2** and **RQ3**). I am motivated by the need to develop a unified theoretical framework for shared autonomy. This work will be the first to treat information content and flow as the key components in understanding the dynamics of interaction between human and autonomy. More importantly, this work proposes a fundamentally different way of thinking about autonomy; one in which *autonomy is a exogenous intervention that alters the information flow in a coupled perception-action loop to bring about desired outcomes*.

2 Human-Robot Interaction in Shared Autonomy

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

2.1 Mathematical Models for Shared Autonomy

In shared autonomy, complementary abilities of humans and robots 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 autonomy is. In a recent survey paper on shared control Abbink et al. defines shared control 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, they also propose that in a shared autonomy 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, HRI in this context is a complex phenomena 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 mathematical models.

Researchers develop mathematical frameworks 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 present a discuss each of the above-mentioned categories in greater detail.

2.1.1 Models for Human Behavior

Teamwork in a shared autonomy system is enhanced when the team members understand each other's intentions, desires and goals. However, there are particular challenges that arise in human-robot interaction due to the differences in the mental and physical capabilities of humans and robots [27]. Robots can deal with such challenges by maintaining models of human cognition and behavior [33] spanning different timescales and levels.

Hiatt et al. in [27] utilize 'Marr's levels of analysis' [45] to categorize models for human behavior into three distinct categories namely: computational, algorithmic and implementational. According to the authors, 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 assumption 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, Dragan et al. assumes a framework in which the human is treated as a optimal agent that noisily optimizes a goal-dependent cost function [15]. 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 [63] and also because they provide a principled approach to how agents ought to behave. The 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 [67] and inverse optimal control [16]. In our work on human-driven customization of shared autonomy levels [22], we utilize ideas from optimal control theory to model human behavior in which we assume that humans are acting optimal with respect to an *unknown* cost function. We make no assumptions regarding the nature of 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 (encoded as a cost function) and therefore by having the human optimize the control allocation parameters directly, the user will tune them to his liking and preference. Similarly, our 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. 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 [46] and assistive robotics [19]. More recently, data-driven approaches based on Koopman operators have also been utilized to learn models of joint human-machine systems [7]. Koopman operator based approaches scale well to high-dimensional spaces as the computational complexity does not grow with the number of data points.

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 [35]. 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) [58]. For example, POMDP based models are used for assessing the human partner's trust in the autonomous partner [8]. POMDPs can also be used by autonomy in which the human intent are treated as latent states. By performing online inference on these latent states, autonomy will have the ability to incorporate human intentions (desired goals) into its own decision making process thereby implicitly taking into account user preferences. 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 give a state (policy) [14]. In our work, we have relied on heuristic approaches based on simple directedness and proximity based confidence functions to estimate reaching intent in a table-top manipulation tasks [22]. In order to incorporate information from past states, in my previous work, I have also developed a framework for intent recognition based on ideas from *Dynamic Field Theory* [55] in which the time-evolution of the probability distribution over intent is specified as a constrained dynamical system [21]. The activation function that drives the dynamical system plays a role akin to that of 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 inference of latent states and actions is made possible by learning specific proposal distributions[25] to be used within an Reversible Jump Markov Chain Monte Carlo (RJMCMC) sampling process [24].

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 the interdependencies between them at same as well as different time slices[49]. 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[59]. However, structure of DBNs need to pre-specified by a domain expert and learning the parameters requires large amounts of data.

2.1.2 Models for Policy Generation

Mathematical models are also widely used for generating/learning policies responsible for generating autonomy's control actions. In addition to successful task completion, autonomy would benefit from the 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, in imitation learning (IL) paradigms can be utilized to develop end-to-end systems that can directly map a high dimensional state to actions [4]. 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 [67]. Policies that optimize long-term accumulated reward (solution of the forward reinforcement learning (RL) problem) improves the robustness and generalization capabilities of the autonomous partner. 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 [38]. 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 [66].

Planning-based approaches such as probabilistic road maps [34] and RRT [41] 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. HRI in shared autonomy can become more seamless if the robot 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 [13]. 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 [57]. 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 [48]. In my work with assistive robotic manipulators for table-top manipulation [22], 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 [36]. More recent work from utilized 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 [54]. 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 [44].

2.1.3 Models for Control Allocation

In a shared control task responsibility is split between the human and autonomy. Therefore, control allocation, that is, how exactly should control be arbitrated between the human and autonomous partner, needs to implemented appropriately for desired outcomes. Different approaches exist for sharing control 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 higher 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 [56] to a select a desired destination in the world or a laser pointer [9] to point to a desired goal and the autonomy can 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 reach target which then combined with a object recognition and motion planning modules can produce the desired robot trajectory [6].

Blending-based approaches seek to arbitrate between human and autonomy actions at the control signal level (or the policy level) directly. In [15] Dragan et al. introduce the policy blending formalism for shared control 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. 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 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 [22]. By casting shared control allocation in a broader theoretical framework, Trautman has proposed a mathematical model for probabilistic shared control in complex dynamic environments [62].

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, Trautman is able to recover linear blending as a special case of the more general framework and shows that in general, linear blending is suboptimal with respect to the joint metric of agreeability, safety and efficiency.

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, the 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 and has been shown to add to the cognitive and physical burden and affects task performance negatively [53]. 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 [26]. Machine learning techniques have been utilized to learn mappings from robot state to control modes preferred by human users [32]. 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 [20].

2.2 Desired Characteristics of HRI (Could have a better subsection title)

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 exactly 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 fundamental characteristics of any successful team? Researchers in the fields of cognitive psychology and philosophy have been long interested in these questions.

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 [61]. Research in developmental psychology have 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. Tomasello et al. in [61] claims 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. Bratman points out that *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*. Bratman proceeds to identify three important characteristics of shared cooperative activity namely a) mutual responsiveness b) commitment of joint activity and c) commitment to mutual support [5]. 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" [10]. The theory of common ground was originally developed to understand communication between people and its main assumption is that effective communication requires coordination that elies 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 human-robot interaction has been successful to a large extent and has contributed to advances in the 'science' of human-robot interaction. Kiesler in [37] shows how 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. In [39] Klein et al. 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 by Lyons et al. in [43]. Fluency is yet another important aspect of a successful collaboration. Hoffman et al. has proposed 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 [29]. They 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 human robot interaction is the turn-taking paradigm. Thomaz et al. propose a first-order Markov process based model to describe the turn-turning dynamics between a human and a social robot. One of their primary findings is that human turn-taking behavior is mainly influenced by information flow [60]. However, the authors don't present a quantitative treatment of the information flow. Drawing influences from research in the human factors community, Nikolaidis et al. have developed human-robot teaming models that are shared between the team members [51]. 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 space have also emphasized the need for user-centric development of shared autonomy 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-robot 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 [17].

2.3 Summary

Human-robot interaction is a highly interdisciplinary field drawing influences from a variety of domains such as computer science, machine learning, cognitive psychology and philosophy. From the above presented survey of related literature, we can see that researchers have explicitly made use of mathematical models to deal with a variety of problems that arise HRI such as modeling human behavior, generating appropriate 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 relates 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, desires and goals for enhancing the quality of HRI. These ideas provide a solid theoretical basis for engineering practical shared autonomy systems. However, to the best of my knowledge a thorough mathematical characterization of the above mentioned 'softer aspects of HRI' still do not exist. Currently, the computational approach that researchers adopt to solve human-robot interaction decomposes the large picture into smaller subproblem that are tackled independent of each other.

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), I strongly think that information theoretic analysis of the interactions can 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 HRI with an emphasis on information theoretic analysis of interactions between the nodes in the network. Within this framework *information flow and exchange* assumes a more foundational and fundamental role and autonomy's influence on the interaction will be treated as an external intervention that will drive the interaction dynamics to have desired characteristics. In the following chapter, I will introduce the causal Bayesian Network based mathematical framework for studying HRI 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 []. Within this framework, autonomy is thought of as appropriately timed exogenous interventions that can alter the information flow between the different nodes in the network.

3.1 Causal Bayesian Networks for HRI

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 autonomy setting, we use 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 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*. It is likely that their line of sight is obstructed due to occlusion from the robot and other factors and as a result, the human is only able to receive partial information regarding the true state of the environment. Typical robot teleoperation is enacted through a control interface such as a joystick, switch-based headarray or a sip-n-puff tube. How the human chooses to act (control the robot) 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, if we focus on how autonomy interacts with the robot, the autonomy receives partial information about the robot state through various types of sensors (joint encoders, RGBD cameras) and takes actions to control the robot using available information. This interaction between autonomy and environment can also be thought of a perception-action loop.

In shared autonomy, the human and autonomy interacts with the environment concurrently to accomplish tasks jointly. In addition to interacting with the environment, 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 *how* the autonomy tries to assist the human in the task, in which case, the human needs to infer the autonomy's decision-making logic to effectively cooperate and coordinate with the autonomy). This joint interaction can be captured by coupling the independent perception-action loops into one.

Causal Bayesian Networks (CBS) provide a systematic mathematical framework to model this coupled perception-action loop that unfolds in time. In a CBN, the nodes represent the variables of interest for the quantification of HAI and the edges represent the statistical dependence between the nodes.

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 CBN.

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

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

where $p(x)$ is the probability mass distribution and the summation extends over all possible states the random variable can assume. 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. 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 into account some context. 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 time-series processes of random variables $\{\dots, X_{n-1}, X_n, X_{n+1}, \dots\}$ with process realizations denoted by $\{\dots, x_{n-1}, x_n, x_{n+1}, \dots\}$, where n denotes the discrete countable time index. Let $\boldsymbol{X}_n^{(k)} = \{X_{n-k+1}, \dots, X_{n-1}, X_n\}$ denote the k consecutive variables of X which has realizations $\boldsymbol{x}_n^{(k)} = \{x_{n-k+1}, \dots, x_{n-1}, x_n\}$[1].

Now let us consider another random time-series process of random variables 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

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

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

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

This particular mathematical form of transfer entropy is important for computational purposes. As is obvious from the equations above 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.

[1] $\boldsymbol{x}_n^{(k)}$ 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.

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 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* [**?**]. 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 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 by Kraskov et. al [**?**]. Their approach combines various techniques that are designed to reduce the bias and variance errors that can occur due to small sample sizes. This approach relies on a nearest-neighbors approach which is effectively equivalent to dynamically changing the kernel width to the density of the samples.

4 Proposed Studies

4.1 Quantifying Transparency

Systematic manipulation of two factors that can direct affect the transparency of autonomy's intent when controlling the robot, namely, signal sparsity and directional perturbation. Signal sparisty can be controlled using a dropout factor. Directional perturbation is amount by which the straight line path is perturbed.

Could also test, Joint angle vs. Cartesian representation.

4.2 Inference

4.3 Learning

4.4 Task Performance

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] 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.

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

[6] 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.

[7] 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.

[8] 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.

[9] 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.

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

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

[12] 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.

[13] 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.

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

[15] 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.

[16] 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.

[17] 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*.

[18] 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.

[19] 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.

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

[21] 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.

[22] 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.

[23] 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.

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

[25] 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.

[26] 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.

[27] 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.

[28] 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.

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

[30] 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.

[31] Shuhei Ikemoto, Heni Ben Amor, Takashi Minato, Bernhard Jung, and Hiroshi Ishiguro. Physical human-robot interaction: Mutual learning and adaptation. *IEEE robotics & automation magazine*, 19(4):24–35, 2012.

[32] 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*.

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

[34] 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.

[35] 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.

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

[37] 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.

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

[39] 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.

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

[41] 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.

[42] 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.

[43] 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.

[44] 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.

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

[46] 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.

[47] 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.

[48] 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.

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

[50] 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.

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

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

[53] 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.

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

[55] 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.

[56] 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.

[57] 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.

[58] 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.

[59] 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.

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

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

[62] 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.

[63] 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.

[64] Zhikun Wang, Katharina Mülling, Marc Peter Deisenroth, Heni Ben Amor, David Vogt, Bernhard Schölkopf, and Jan Peters. Probabilistic movement modeling for intention inference in human–robot interaction. *The International Journal of Robotics Research*, 32(7):841–858, 2013.

[65] 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.

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

[67] 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.