

Efficient and Trustworthy Social Navigation via Explicit and Implicit Robot–Human Communication

Yuhang Che , Allison M. Okamura , *Fellow, IEEE*, and Dorsa Sadigh

Abstract—In this article, we present a planning framework that uses a combination of implicit (robot motion) and explicit (visual/audio/haptic feedback) communication during mobile robot navigation. First, we developed a model that approximates both continuous movements and discrete behavior modes in human navigation, considering the effects of implicit and explicit communication on human decision-making. The model approximates the human as an optimal agent, with a reward function obtained through inverse reinforcement learning. Second, a planner uses this model to generate communicative actions that maximize the robot’s transparency and efficiency. We implemented the planner on a mobile robot, using a wearable haptic device for explicit communication. In a user study of an indoor human–robot pair orthogonal crossing situation, the robot is able to actively communicate its intent to users in order to avoid collisions and facilitate efficient trajectories. Results show that the planner generated plans that are easier to understand, reduce users’ effort, and increase users’ trust of the robot, compared to simply performing collision avoidance. The key contribution of this article is the integration and analysis of explicit communication (together with implicit communication) for social navigation.

Index Terms—Haptics and haptic interfaces, human-centered robotics, motion and path planning, social human–robot interaction.

I. INTRODUCTION

MOBILE robots are entering human environments, with applications ranging from delivery and support in warehouses, to home and social services. In this work, we focus on social navigation, in which the movements and decisions of robots and humans affect each other. Researchers have studied various methods for generating robot motions that can avoid collision and comply to social rules during navigation [1]–[7].

Manuscript received July 18, 2019; accepted December 22, 2019. Date of publication January 23, 2020; date of current version June 4, 2020. This work was supported by the Stanford AI Lab, Stanford School of Medicine and Stanford Graduate School of Business under the Human-Centered AI Seed Grant Program. This article was recommended for publication by Associate Editor S. Calinon and Editor E. Yoshida upon evaluation of the reviewers’ comments. (Corresponding author: Yuhang Che.)

Y. Che is with the Waymo LLC, Mountain View, CA 94043 USA (e-mail: yuhangc@waymo.com).

A. M. Okamura is with the Department of Mechanical Engineering, Stanford University, Stanford, CA 94305 USA (e-mail: aokamura@stanford.edu).

D. Sadigh is with the Department of Computer Science, Stanford University, Stanford, CA 94305 USA (e-mail: dorsa@cs.stanford.edu).

This article has supplementary downloadable multimedia material available at <http://ieeexplore.ieee.org> provided by the authors. This includes a video, which shows the participant and robot movements, which should clarify the nature of the interactions under the different conditions. This material is 31.1 MB in size.

Color versions of one or more of the figures in this article are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TRO.2020.2964824



Fig. 1. Social navigation scenario in which the robot communicates its intent (to yield to the human) to the human both implicitly and explicitly. Implicit communication is achieved via robot motion (slowing down and stopping), and explicit communication is achieved through a wearable haptic interface.

These methods have shown great potential toward bringing intelligent robots into human environments. However, there is an inherent limitation—robot motion alone does not capture all aspects of social interaction. In scenarios such as social navigation, there are many ways that humans interact with each other. For example, by establishing eye contact, two people often assume that they will both take actions to avoid each other (people often act more conservatively if they think others are not paying attention). In addition, there are scenarios where conflicts occur and require more involved communication. One common example is people choosing to move to the same side when trying to avoid each other, and then moving to the other side simultaneously again. In this case, a person usually has to stop and gesture the other person to go first. Social robots need to be able to handle such conflicts in a similar manner by incorporating different ways of communication.

The aforementioned limitation motivates us to explore richer interaction and communication schemes for robot social navigation. Specifically, we propose to consider both *implicit communication* and *explicit communication* (see Fig. 1). In human–robot interaction literature, there are various definitions of implicit and explicit communication [8]–[10]. Our definition of implicit communication closely follows that of [10]—the information from implicit communication needs to be extracted by leveraging the context and common ground. Conversely, we define explicit communication as actions that convey information in a deliberate and nonambiguous manner.

Examples of implicit communication include eye contact and body language. In social navigation, people often encode their intentions into their motion, which serves as implicit communication [11]. Similarly, a robot's motion carries information about its intent. The notion of motion as implicit communication is closely related to the idea of legibility [12], which measures how well a robot's motion (and behavior, more generally speaking) expresses its intent. In this work, we develop an algorithm that enables the emergence of legible motions in social navigation.

Given our definition, explicit communication can take many forms, including spoken language, visual signals (symbols and colors), and audio signals (sirens), as long as the meanings of these signals are established beforehand. In this work, we focus on the use of haptic feedback because it is immediate and targeted. In addition, haptic feedback is advantageous when users' visual and auditory channels are already overloaded. However, the algorithms developed in this work are not limited to haptic communication and can be directly applied to other forms of explicit communication.

Both implicit and explicit communication during navigation will affect how people move. Thus, we propose that the ability to understand and predict navigation behaviors of humans would be beneficial to planning communicative actions. We developed a data-driven, probabilistic model to predict human behaviors, including both continuous movements and discrete behavior modes, given robot actions. Leveraging the human model, we develop an algorithm that plans for a robot's motion and communication through haptic feedback. Our approach relies on the assumption that users are cooperative—they will not intentionally thwart the robot. In environments in which humans and robots work together, such as offices, homes, hospitals, and warehouses, this assumption is reasonable. We leave scenarios where humans act adversarially to future work. We also assume that users are equipped with wearable haptic interfaces, and the robot can send feedback via these interfaces (can be replaced with sound or visual cues if appropriate). In this work, we focus on the interaction between one robot and one human user.

The main contributions of this article are as follows.

- 1) A predictive model of human navigation behavior in response to both implicit and explicit robot communication. The model is described in Section IV, and the framework in which the model is applied is described in Section III.
- 2) An interactive planning algorithm based on the human model that enables a robot to proactively communicate through either implicit or explicit communication, with the goal of efficient and transparent interaction (see Section V).
- 3) Implementation of the proposed algorithm on a physical mobile robot platform, and analysis and verification of the algorithm with user studies. We first evaluated and tuned the system in simulation (see Section VI) and then, performed experiments (see Section VII).

II. BACKGROUND

A. Social Navigation

In traditional robot motion planning frameworks, humans are usually treated as obstacles that the robot should avoid [13].

However, in social navigation scenarios, humans are different from most other obstacles—they have a purpose, follow social rules, and react to other agents in the environment, including the robot [14]. Researchers have explored various methods for social-aware motion planning.

Understanding and modeling reactive behaviors of humans is necessary for planning socially acceptable actions. A popular framework to model human navigation in response to other agents is the social force model (SFM) [15]–[17], which frames pedestrians as affected by interactive forces that drive them toward their goals and away from each other. SFM has been applied to human–robot interaction and social navigation scenarios [2], [18]–[20]. Some work in the field of multiagent planning also consider the reactive nature of moving agents. For example, optimal reciprocal collision avoidance (ORCA) [21] assumes that each agent takes some responsibility of resolving pairwise conflicts, which result in global collision free paths. Guy *et al.* [22] extended ORCA to model human navigation by introducing additional parameters such as reaction time and personal space. These reaction-based methods are popular for crowd simulation because of the high computational efficiency. However, they tend to be less accurate in predicting individual trajectories, and do not capture the stochasticity of human behavior.

Researchers have applied probabilistic methods to modeling human behavior and planning for social navigation. The interactive Gaussian process model was proposed in [23]; it takes into account the variability of human behavior and mutual interactions. Trautman *et al.* [1] used this model to navigate a robot through a crowded cafeteria. Inverse reinforcement learning (IRL) [24]–[27] is another framework that allows a probabilistic representation of human behavior. IRL was applied in [4]–[6], and [28] to learn pedestrian models and plan for socially compliant navigation. In general, these approaches aim to generate “humanlike” actions, with the underlying assumption that the robot should behave similarly to a human in order to be socially acceptable. We also use IRL as part of our model to predict humans' responses. Our work differs from previous approaches in that we model jointly the effects of both explicit communication and robot motions on human navigation behavior.

Recently, deep neural networks also gained attention in robot navigation applications. Chen *et al.* [7] incorporated social norms and multiagent reactions into a deep reinforcement learning framework to enable robot navigation in a pedestrian-rich environment. Tai *et al.* [29] applied a generative adversarial imitation learning strategy to generate socially compliant navigation from raw sensor data. Deep learning frameworks have the advantage that they often impose less assumptions on human behavior. However, it is unclear whether they generate motions that are natural and transparent. In addition, the amount of data required for training these neural networks can be large and expensive to obtain. Our approach is more data efficient and explainable by leveraging modeling techniques.

B. Expressive Motion in Robotics

Besides social navigation, researchers have investigated the problem of planning interactive and communicative motions

in the field of computational human–robot interaction [30]. Dragan *et al.* [31] formalized the idea of legibility and proposed methods to generate legible motions for robot manipulators [12]. Legible motions are motions that express the robot’s intent to observers, and can be different from predictable motions [31]. Sadigh *et al.* [32] modeled a human–robot system jointly as a dynamical system, and proposed methods to plan motions that actively influence human behaviors and further actively gather information of internal state of humans [33], [34]. Similar ideas were explored in human–robot collaboration scenarios. Bestick *et al.* [35] developed a method to enable a robot to purposefully influence the choices of a person in handover tasks to help the person avoid suboptimal grasps. Nikolaidis *et al.* [36] studied human–robot mutual adaptation, and proposed a decision framework to allow the robot to guide participants toward a better way of completing the task. In our work, we consider robot motions that can inform or affect people as a form of implicit communication, as people often have to infer about the robot’s intent with context.

C. Nonverbal Explicit Communication Methods

In this article, we define actions that communicate information deliberately and unambiguously as explicit communication. The most common form of explicit communication is verbal communication. Nonverbal communication can be made explicit, such as pointing gestures [37], [38], head nod [39], and visual displays [40].

Haptic feedback has been used as an explicit communication mechanism in human–robot interaction. Scheggi *et al.* [41], [42] designed a vibrotactile haptic bracelet to guide the user along trajectories that are feasible for human leader–robot follower formation tasks. The bracelet can display three distinct signals that inform human to slow down and turn toward left or right. Sieber *et al.* [43] used haptic feedback to assist a user teleoperating a team of robots manipulating an object by communicating the status (stable versus in transient) of the robot team. Here, we use haptic feedback to explicitly convey the robot’s discrete intent to the user in collision avoidance scenarios during navigation.

Nonverbal communication, such as gaze, are considered as implicit if it indirectly conveys information [44]–[46]. In this work, we use the motion of the mobile robot, which does not impose any constraint on the appearance of the robot and can naturally be incorporated in social navigation.

III. PLANNING FRAMEWORK FOR SOCIALLY AWARE ROBOT NAVIGATION

Humans have an impressive ability to navigate seamlessly around other people. They walk, and still walk smoothly and efficiently when faced with a large group of people walking toward or around them in busy streets, concerts, cafeterias, conferences, or office spaces. However, robots lack the same *elegance* while navigating around people. Our goal is to design mobile robots that navigate in a similar fashion to humans in spaces shared by humans and robots. *Our insight for social navigation is that mobile robots—similar to humans—need to proactively plan for interactions with people.* In this section, we provide a high-level outline of our interaction framework,

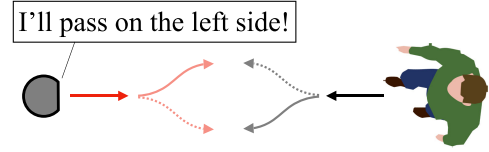


Fig. 2. Example scenario where the robot needs to plan for appropriate movement and communication.

and we will provide a more detailed description in Sections IV and V.

In this work, we focus on a scenario in which one robot interacts with one human. To illustrate this, we use the running example illustrated in Fig. 2—a robot and a human need to pass each other either on the left side or on the right side. To achieve this pass smoothly, the robot and human must understand each other’s intent and coordinate their movements. The objective of our planning framework is to generate appropriate explicit communication and robot motions (which also serve as implicit communication) to facilitate human–robot interactions in such scenarios. The explicit communication consists of a finite number of discrete signals, for example, expressing a plan to pass on the left or the right side. The robot motions consist of continuous actions such as the robot’s linear and angular velocities.

We model the joint human–robot system as a fully observable interactive dynamical system in which the actions of the human and the robot can influence each other. In addition, we consider the stochasticity of actions of the human and the effect of communication in our planning framework.

Let s_h^t and s_r^t denote the physical state of the human and the robot at a certain time step t , respectively. The states are continuous, and include positions and velocities of each agent. The robot can apply continuous control inputs u_r^t that change its state through the following dynamics:

$$s_r^t = \mathcal{F}_r(s_r^{t-1}, u_r^t). \quad (1)$$

Similarly, the state of the human is directly changed by her action u_h^t

$$s_h^t = \mathcal{F}_h(s_h^{t-1}, u_h^t). \quad (2)$$

We let $\mathbf{u}_r = [u_r^1, \dots, u_r^N]^\top$ and $\mathbf{u}_h = [u_h^1, \dots, u_h^N]^\top$ denote the sequence of robot and human actions over a finite horizon N ($[\cdot]^\top$ denotes a vector). $\xi_r^{1:N} = [(s_r^1, u_r^1), \dots, (s_r^N, u_r^N)]^\top$ and $\xi_h^{1:N} = [(s_h^1, u_h^1), \dots, (s_h^N, u_h^N)]^\top$ are then the state-action trajectories of the robot and the human over the horizon N , respectively. We use u_c to refer to the explicit communication actions. Each of the actions u_r^i and u_h^i are vectors with appropriate dimensions. In this work, we focus on a 1-D communication action u_c , i.e., $u_c \in \mathcal{U}_c \cup \text{“none.”}$

We define an overall reward function for the robot $R_r(\xi_r^{1:N}, \xi_h^{1:N}, u_c)$. We let this reward function depend on both the robot’s trajectory and the trajectory of the human. This reward function encodes properties such as efficiency (e.g., robot’s velocity, distance to goal), safety (distance to human), and other metrics. In this work, we use an optimization-based planning algorithm, where the robot will optimize the expected sum of its reward function R_r . Equation (3) represents this expected

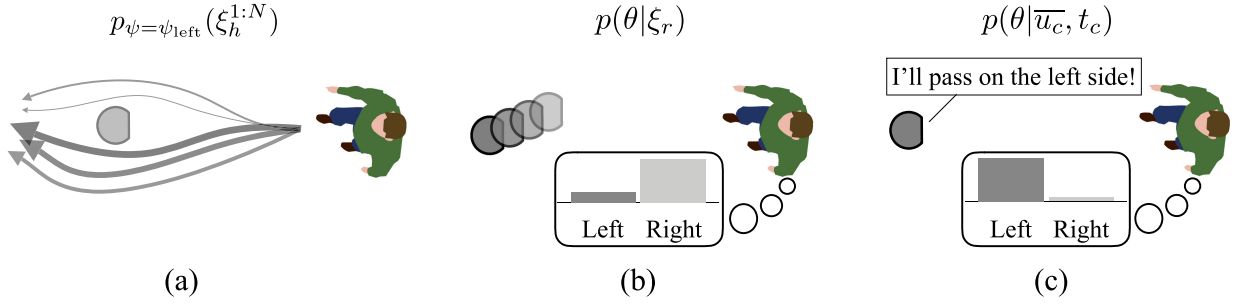


Fig. 3. Overview of different components of the model. (a) We model the distribution of human trajectories given their behavior mode ψ . Each behavior mode is associated with an underlying intent. In this example, trajectories that pass the robot on the left have much higher probabilities (thicker lines), because the underlying intent is to pass on the left. (b) Robot's implicit communication (motion) affects the belief of the human over the robot's intent, and therefore, affects the behavior mode of the human. (c) Robot's explicit communication also has strong effects on the belief and behavior mode of the human.

reward, which is computed over predicted trajectories $\xi_h^{1:N}$ of the human. We emphasize that the distribution of $\xi_h^{1:N}$ is affected by the robot actions. In the next section, we present a hybrid model that predicts this distribution.

We will discuss our planning algorithm in Section V. We will use model predictive control (MPC) [47] algorithm, where at every time step, the robot computes a finite horizon sequence of actions to maximize its expected reward

$$\mathbf{u}_r^*, u_c^* = \arg \max_{\mathbf{u}_r, u_c} \mathbb{E}_{\xi_h^{1:N}} [R_r(\xi_r^{1:N}, \xi_h^{1:N}, u_c)]. \quad (3)$$

In an MPC scheme, at each time step, the robot only executes the first control in the optimal sequence \mathbf{u}_r^* , and then, replans at the next time step. The robot also plans on an explicit communication $u_c^* \in \mathcal{U}_c \cup \text{"none."}$ The robot decides to either communicate ($u_c^* \in \mathcal{U}_c$), or does not provide any communication ($u_c^* = \text{"none"}$). When computing the plan at each time step, we assume that explicit communication can only happen at the beginning of the planning horizon. This is because only the first set of actions in the planned sequence is executed, and a new plan will be generated in the next time step.

The key insight of this framework is that the robot is aware of the stochasticity of human behavior and the influences of communication on humans' actions. As a result, the robot will proactively communicate its intent to avoid undesirable situations. Take the example scenario in Fig. 2—if the robot and the human choose to pass each other on different sides, a collision could happen. By expressing its intent explicitly, the robot minimizes the chance of a potential collision.

IV. HYBRID MODEL OF HUMAN NAVIGATION

In this section, we present our model of human navigation, focusing on how the robot's implicit and explicit communication affect behaviors of the human.

To compute the optimal robot actions using (3), the robot is required to predict human actions over a finite time horizon. Since human actions are stochastic, the robot needs to predict the distribution of \mathbf{u}_h . We assume the human dynamics \mathcal{F}_h are known. Therefore, the distribution of \mathbf{u}_h induces an equivalent distribution over human trajectories $\xi_h^{1:N}$ for a given \mathbf{u}_h and

initial state. Modeling this distribution exactly is quite challenging and may be computationally expensive. Our objective is to develop an approximate model that captures the interactive nature of human behavior, i.e., *how the human will react to the robot's movements and explicit communication*.

Fig. 3 illustrates the three most important components of our model. On a high level, we model human navigation behavior as a joint mixture of distribution over the trajectory, similar to [4]

$$p(\xi_h^{1:N}) = \sum_{\psi} p_{\psi}(\xi_h^{1:N}) p(\psi). \quad (4)$$

The first part of the mixture $p_{\psi}(\xi_h^{1:N})$ models the distribution of the continuous state-actions, given a specific behavior mode ψ . The second part $p(\psi)$ models the distribution of the discrete behavior mode ψ . The behavior mode reflects the underlying intent of the human, and is affected by both the robot's implicit communication (motion) ξ_r and explicit communication \overline{u}_c .

The idea of discrete behavior modes or classes of actions is sometimes associated with homotopy classes of trajectories [4], [48]. Here, we do not enforce the trajectories, given a specific value of ψ belongs to one homotopy class. Instead, we associate each value of ψ with an underlying reward function R_h^{ψ} that the human optimizes. With an inverse reinforcement learning framework, our model assigns high probabilities to trajectories that comply with the behavior mode (one homotopy that corresponds to the underlying intent of the human), and low but nonzero probabilities to trajectories in other homotopy classes. This representation is more generalizable than homotopy, as different behavior modes are not necessarily different in homotopy in other scenarios. For example, tailgating versus keeping safe distances are two behavior modes in driving, and they cannot be differentiated by homotopy. However, they can be characterized by different cost functions that drivers may optimize.

For computational efficiency, we use relatively simple models for human and robot dynamics. The human state $s_h^t = [x_h, y_h, \dot{x}_h, \dot{y}_h]^T$ consists of positions and velocities in the 2-D plane, and the robot state $s_r^t = [x_r, y_r, \theta_r]^T$ is simply its pose. We use a constant acceleration model for human dynamics. Thus, the human control input is $u_h^t = [\ddot{x}_h, \ddot{y}_h]^T$. For the robot dynamics, we use a differential drive kinematic model, with

control input $u_r^t = [v, \omega]^T$, consisting of linear and angular velocities.

The formulation in (4) describes our modeling approach in its most general form. In our specific setup, the distribution $p(\xi_h^{1:N})$ should be conditioned on the history of human and robot trajectories, history of explicit communication, and future robot actions over the prediction horizon. To numerically compute the distribution and use it for planning, we make the following two assumptions.

Assumption 1: The trajectory of the human $\xi_h^{1:N}$ over the prediction horizon depends on the environment E (including the goal of the human) and the future robot trajectory $\xi_r^{1:N}$

$$p_\psi(\xi_h^{1:N} | \xi_r^{1:N}, E). \quad (5)$$

Here, actions of the human depend on the robot's actions. However, the robot also needs to predict actions of the human and plan accordingly. This interdependency leads to the *infinite regress* problem, i.e., a sequence of reasonings about each others' actions that can never come to an end. To resolve this problem, we assume that the human has access to the future robot trajectory, and model the interaction between the human and the robot as a two-player Stackelberg game (leader-follower game). We argue that this assumption is reasonable—given relatively short planning horizon, humans are usually able to predict immediate actions of other agents. Since the focus of this work is the interaction between the robot and the human, the goals and environment are assumed to be known and fixed.

Assumption 2: We assume that the behavior mode is only affected by past information. With this, we can express the second term in (4) as

$$p(\psi | \xi_r^{-N:0}, \xi_h^{-N:0}, u_c^{-N:0}) \quad (6)$$

where $\xi_r^{-N:0}, \xi_h^{-N:0}$ are the past trajectories of the robot and the human and $u_c^{-N:0}$ is the past explicit communications over a horizon of length N . We further assume that only the most recent explicit communication actually affects the behavior mode

$$p(\psi | \xi_r^{-N:0}, \xi_h^{-N:0}, \bar{u}_c, t_c). \quad (7)$$

Here, \bar{u}_c represents the most recent explicit communication and t_c represents the time when \bar{u}_c is communicated. At each time step, the robot can decide to not perform any explicit communication ($u_c = \text{"none"}$). The most recent communication refers to the most recent $\bar{u}_c \neq \text{"none"}$, e.g., the last time haptic feedback was provided.

In the next two subsections, we discuss details of modeling humans' continuous navigation actions as in (5) and the humans' discrete behavior mode as in (7).

A. Modeling Continuous Navigation Actions of a Human

Inverse Reinforcement Learning: We employ a data-driven approach to model humans' continuous navigation actions. Specifically, we use maximum entropy IRL (MaxEnt IRL) [25], [27] to learn the distribution of human actions that matches a set of provided demonstrations in expectation. We assume the humans are approximately optimizing a reward function $R_h^\psi(\xi_h^{1:N}, \xi_r^{1:N}, E)$ for a given discrete class ψ . Under this

model, the probability of the actions \mathbf{u}_h is proportional to the exponential of the total reward of the trajectory (Z is a normalization vector)

$$\begin{aligned} p_\psi(\xi_h^{1:N} | \xi_r^{1:N}, E) &= \frac{1}{Z} \exp \left(R_h^\psi(\xi_h^{1:N}, \xi_r^{1:N}, E) \right) \\ &= \frac{1}{Z} \exp \left(\mathbf{w}_\psi \cdot \mathbf{f}(\xi_h^{1:N}, \xi_r^{1:N}, E) \right). \end{aligned} \quad (8)$$

We parameterize the human reward function as a linear combination of features \mathbf{f} that capture relevant properties of the navigation behavior. The features are weighted by \mathbf{w}_ψ , which can be learned by maximizing the overall likelihood of the provided demonstrations \mathcal{D}

$$\arg \max_{\mathbf{w}} \prod_{\mathbf{u}_h \in \mathcal{D}} p_{\mathbf{w}}(\xi_h^{1:N} | \xi_r^{1:N}, E). \quad (9)$$

The weight vector \mathbf{w}_ψ is a function of the behavior mode ψ . Given a selected set of features, such as distance to the other agents, heading, or velocity, we learn appropriate weights corresponding to the humans' reward functions for each discrete class ψ from collected demonstrations. We will discuss the specific features used in this work in Section VI.

B. Modeling Behavior Modes of a Human

When modeling the distribution of behavior modes of a human, we consider the effect of both implicit communication (via robot movements) and explicit communication (via haptic feedback). We assume that the human will infer the robot's intent, and act cooperatively during the interaction. Mathematically, this suggests that the distribution of behavior modes of the human is related to her belief over the robot's intent θ . In our particular social navigation problem, we assume this belief is equal to the probability of choosing a behavior mode

$$p(\psi | \xi_r^{-N:0}, \xi_h^{-N:0}, \bar{u}_c, t_c) = p(\theta | \xi_r^{-N:0}, \xi_h^{-N:0}, \bar{u}_c, t_c). \quad (10)$$

In other words, if the human believes that there is an 80% chance that the robot will yield priority, she will behave in "human priority" mode with the same probability of 0.8. Nonlinear relationships between $p(\psi)$ and $p(\theta)$ could also be used to model risk-averse (more willing to yield priority) and risk-seeking behaviors.

Using Bayes rule, we can transform (10) to

$$\begin{aligned} p(\theta | \xi_r^{-N:0}, \xi_h^{-N:0}, \bar{u}_c, t_c) \\ \propto p(\xi_r^{-N:0}, \bar{u}_c, t_c | \theta, \xi_h^{-N:0}) \cdot p(\theta | \xi_h^{-N:0}) \\ = p(\xi_r^{-N:0} | \theta, \xi_h^{-N:0}) \cdot p(\bar{u}_c, t_c | \theta) \cdot p(\theta). \end{aligned} \quad (11)$$

The second step in (11) assumes conditional independence of the robot trajectory $\xi_r^{-N:0}$ and explicit communication u_c . With this factorization, we can separately model the effect of robot motion (implicit communication) and explicit communication on behavior modes of the human. The last term $p(\theta)$ is the prior on the robot's intent, which should be decided based on the application. In our implementation, there is an equal chance that the robot communicates different intents. So, we choose a uniform prior.

The formulation casts the backward inference problem (from action to intent) into forward prediction problems (from intent to action). Next, we derive methods to compute each part of (11).

1) *Implicit Communication*: We assume that the human, in general, expects the robot to be rational and efficient according to some reward function. Applying the principle of maximum entropy, we model the human as expecting robot movements with probability

$$p(\xi_r^{-N:0}|\theta, \xi_h^{-N:0}) = \frac{\exp(R_r^\theta(\xi_r^{-N:0}, \xi_h^{-N:0}))}{\int_{\xi_r'} \exp(R_r^\theta(\xi_r', \xi_h^{-N:0})) d\xi_r'} \quad (12)$$

where $R_r^\theta(\xi_r^{-N:0}, \xi_h^{-N:0})$ is a reward function that the human expects the robot to optimize, given its intent θ . To compute the integration in the denominator of (12), we use the second-order Taylor series to approximate $R_r^\theta(\xi_r^{-N:0}, \xi_h^{-N:0})$.

2) *Explicit Communication*: The belief of the human over the robot's intent is strongly affected by explicit communication, because the intent is directly conveyed. However, the effect of explicit communication should decay over time, as the robot's intent may change, and only short-term intents are communicated. Inspired by a model of human short-term verbal retention [8], we propose

$$p(\bar{u}_c, t_c|\theta) = \frac{1}{Z} \begin{cases} A \exp(-\frac{t-t_c}{M}) + 1 & \bar{u}_c = \theta \\ 1 & \bar{u}_c \neq \theta \end{cases} \quad (13)$$

Here, A and M are parameters that determine the characteristic of the distribution and Z is a normalization factor. The explicit communication initially reflects the true intent with very high probability. However, the inference strength decays over time, and eventually, the communication becomes irrelevant to the robot's latest intent.

To clarify this model, let us consider a simpler version of (10) $P(\theta|\bar{u}_c, t_c)$ to infer the robot's intent with only explicit communication. Assume that $\theta \in \{\theta_1, \theta_2\}$ is binary, $\bar{u}_c = \theta_1$, and $t_c = 0$, and use Bayes rule again

$$\begin{aligned} p(\theta = \theta_1|\bar{u}_c, t_c) &= \frac{p(\bar{u}_c, t_c|\theta_1)p(\theta_1)}{p(\bar{u}_c, t_c|\theta_1)p(\theta_1) + p(\bar{u}_c, t_c|\theta_2)p(\theta_2)} \\ &= \frac{A \exp(-(t-t_c)/M) + 1}{A \exp(-(t-t_c)/M) + 2}. \end{aligned} \quad (14)$$

The change of $p(\theta = \theta_1|\bar{u}_c, t_c)$ over time is plotted in Fig. 4. Initially, the robot's intent can be inferred with high probability given the explicit communication. As time passes, the inference strength decreases and eventually, the belief over the robot's intent becomes uniform (0.5).

In the scenarios we consider, the human infers the robot's intent with both implicit and explicit communication. The combined effect can be computed with (11)–(13) given in this section.

V. PLANNING FOR COMMUNICATION

The general planning framework was described in Section III. In this section, we discuss the details of the algorithm, including

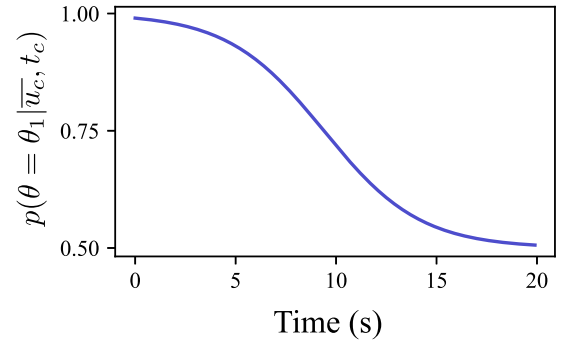


Fig. 4. Inference on the robot's intent based on explicit communication, assuming that $\bar{u}_c = \theta_1$ happened at $t_c = 0$.

the design of reward functions, derivation of the solution to the optimization, and an outline of our implementation.

A. Robot Reward Function

The overall reward R_r that the robot optimizes in (3) consists of four parts that quantify *robot efficiency*, *human comfort*, *safety*, and *reward of explicit communication*.

1) *Robot Efficiency*: The robot should get as close as possible to its target with minimum effort

$$R_{r,r}(\xi_r^{1:N}) = \|x_r^N - x_r^g\| + \gamma_r \sum_{t=1}^N \|u_r^t\|^2 \quad (15)$$

where γ_r is a weight that balances effort and goal. This component determines how fast the robot approaches the goal. x_r^N is the robot position at time step N and x_r^g is the goal position. $\|\cdot\|$ denotes Euclidean norm.

2) *Human Comfort*: We want the human to spend less effort to achieve her goal

$$R_{r,h}(\xi_h^{1:N}) = \|x_h^N - x_h^g\| + \gamma_h \sum_{t=1}^N \|u_h^t\|^2. \quad (16)$$

Similar to γ_r , γ_h balances between the human reaching her goal and her effort.

3) *Safety*: The robot should avoid collisions with the human

$$R_{r,safety}(\xi_r^{1:N}, \xi_h^{1:N}) = \sum_{t=1}^N \exp\left(-\frac{\|\mathbf{x}_{h,t} - \mathbf{x}_{r,t}\|^2}{D_r^2}\right). \quad (17)$$

4) *Explicit Communication*: Too much explicit communication will distract or annoy the human. Therefore, we set a constant reward for performing each explicit communication

$$R_{r,EC}(u_c) = \gamma_{EC} \cdot \mathbb{1}(u_c \neq \text{"none"}) \quad (18)$$

where $\mathbb{1}$ is the indicator function.

The components defined above are then combined linearly with weights c_i

$$R_r(\xi_r^{1:N}, \xi_h^{1:N}, u_c) = \sum_{i=0}^4 c_i \cdot R_{r,i}(\xi_r^{1:N}, \xi_h^{1:N}, u_c). \quad (19)$$

B. Human-Aware Planning

To solve for the optimal robot actions using (3), we need to compute the expectation of the robot reward over the distribution of \mathbf{u}_h . In the last section, we derived a hybrid model for human behavior. To speed up computation, here, we only consider the most likely trajectory, given each possible human intent. With this, the expectation in (3) can be expressed as

$$\mathbb{E}_{\xi_h^{1:N}} [R_r] = \sum_{\psi} p(\psi) \cdot R_r(\xi_r^{1:N}, \xi_h^{\psi}, u_c) \quad (20)$$

where ξ_h^{ψ} is the most likely human response, given ψ , and can be computed as

$$\xi_h^{\psi} = \arg \max_{\xi_h^{1:N}} \mathbf{w}_{\psi} \cdot \mathbf{f}(\xi_h^{1:N}, \xi_r^{1:N}). \quad (21)$$

The distribution $p(\psi)$ is given by (10) and (11).

We use gradient-based optimization to solve the optimal robot controls \mathbf{u}_r^* . This requires the gradient information of the expected reward in (20). Letting $\bar{R}_r = \mathbb{E}_{\xi_h^{1:N}} [R_r]$, we aim to find

$$\frac{\partial \bar{R}_r}{\partial \mathbf{u}_r} = \frac{\partial \bar{R}_r}{\partial \mathbf{u}_r} + \sum_{\psi} \frac{\partial \bar{R}_r}{\partial \mathbf{u}_h} \frac{\partial \mathbf{u}_h^{\psi}}{\partial \mathbf{u}_r} \quad (22)$$

where \mathbf{u}_h^{ψ} is the most likely human action. As we have a symbolic representation of R_r , both $\frac{\partial \bar{R}_r}{\partial \mathbf{u}_r}$ and $\frac{\partial \bar{R}_r}{\partial \mathbf{u}_h}$ can be computed analytically. Following the implicit differentiation method discussed in [32], we compute the last unknown term using the following expression:

$$\frac{\partial \mathbf{u}_h^{\psi}}{\partial \mathbf{u}_r} = \left[\frac{\partial^2 R_h^{\psi}}{\partial \mathbf{u}_h^2} \right]^{-1} \left[-\frac{\partial^2 R_h^{\psi}}{\partial \mathbf{u}_h \partial \mathbf{u}_r} \right]. \quad (23)$$

We have assumed a discrete set of communication actions $\mathcal{U}_c \cup \text{"none."}$ To obtain the optimal explicit communication u_c^* , we enumerate all possible communication actions $u_c \in \mathcal{U}_c \cup \text{"none"}$ and solve the optimization for each u_c . This can be done in parallel because the optimizations do not depend on each other.

The planning algorithm is outlined in Algorithm 1. At every time step, the algorithm first retrieves the states of the robot and the human, and performs a one-step prediction of robot and human states in order to compensate the time spent for planning. Then, it updates the belief over the behavior mode of the human ψ with new observations using (10) and (11). Before optimizing for robot actions, the algorithm needs to generate an initial guess for the initial state of the human. If there is a plan from the previous time step, the plan is used as the initial guess. When there is no previous plan (first time step, or first detection of human), we generate the initial guesses as follows. First, we compute the robot actions ignoring the human. We use a feedback control-based policy to steer the robot toward its goal [49]. We then compute the human actions using an attracting potential field at the goal position, and a repelling potential field centered at the robot. The algorithm then performs a set of optimizations, one for each explicit communication action. Taking advantage of modern multicore CPUs, we can run these

Algorithm 1: Outline of the Planning Algorithm.

Data: \mathcal{U}_c - set of explicit communicative actions

```

1 repeat
    /* get current+predicted states */
2    $s_r^t, u_r^t, s_h^t \leftarrow \text{GetTrackingInfo}()$ ;
3    $\hat{s}_r^t, \hat{s}_h^t \leftarrow \text{PredictOneStep}(s_r^t, s_h^t)$ ;
    /* update the hybrid model */
4    $p(\psi) \leftarrow \text{UpdateModel}(p(\psi), s_r^t, u_r^t, s_h^t, u_c^t, t_c)$ ;
    /* find the optimal actions */
5    $\mathbf{u}_h^0, \mathbf{u}_h^0 \leftarrow$ 
       $\text{GenerateInitGuess}(\hat{s}_r^t, \hat{s}_h^t, x_{g,r}, x_{g,h})$ ;
6   for  $u_c \in \mathcal{U}_c \cup \{\text{null}\}$  do in parallel
7      $\mathbf{u}_r(u_c), \bar{R}_r(u_c) \leftarrow$ 
       $\text{ComputePlan}(\mathbf{u}_r^0, \mathbf{u}_h^0, \hat{s}_r^t, \hat{s}_h^t, p(\psi), u_c)$ ;
8   end
9    $u_c^* \leftarrow \arg \max_{u_c} \bar{R}_r(u_c)$ ;
10   $\mathbf{u}_r^* = \mathbf{u}_r(u_c^*)$ ;
11 if  $u_c^* \neq \text{null}$  then
12    $\text{Communicate}(u_c^*)$ ;
13    $u_c' \leftarrow u_c^*$ ;
14    $t_c \leftarrow t$ ;
15 end
16  $\text{ExecuteControl}(\mathbf{u}_r^*)$ ;
17  $t \leftarrow t + 1$ ;
18 until  $\text{GoalReached}(x_r^t, x_{g,r}, \epsilon)$ ;
    /* execute the actions */

```

optimizations in parallel. Finally, the best explicit communication u_c^* and robot movements (implicit communication) \mathbf{u}_r^* are selected and executed.

We implemented the planning algorithm in C++, and used the software package NLOpt [50] to perform numerical optimization. In our implementation, we chose a planning horizon of $N = 6$, and a time step of 0.5 s.

VI. SIMULATION RESULTS

In order to validate the proposed modeling and planning framework, we first collected human demonstrations and conducted a simulation experiment. In this section, we describe the specific scenario for our experiment, validation of the human model, and simulation results.

A. Social Navigation Scenario

We consider the scenario shown in Fig. 1 for our simulation and experimental evaluation—a human and a robot move in (approximately) orthogonal directions and encounter each other. In this scenario, they have to coordinate with each other and decide who should pass first to avoid collision. We chose this scenario over the face-to-face encounter scenario because in the face-to-face encounter, there is a relatively strong social norm that people pass each other on the right-hand side. We prefer to start with a scenario with no such bias to allow us better focus on the effect of communication. The robot can explicitly

communicate its intent—to yield priority to the human (human priority) or not (robot priority). To numerically evaluate the model and compute the optimal plan, we need to define features leading to a learned reward function, which can then be used to model humans' continuous actions (8) as well as behavior modes (12) of the human.

We select features that are commonly used to characterize social navigation of pedestrians [51].

- 1) **Velocity:** Pedestrians tend to maintain a constant desired velocity. We use a feature that sums up the squared velocity along the trajectory

$$f_{\text{velocity}} = \sum_{t=1}^N \|\dot{x}_h^t\|^2. \quad (24)$$

Here, and in the following equations, $\|\cdot\|$ represents L2 norm.

- 2) **Acceleration:** To navigate efficiently, pedestrians usually avoid unnecessary accelerations

$$f_{\text{acceleration}} = \sum_{t=1}^N \|\ddot{x}_h^t\|^2 = \sum_{t=1}^N \|u_h^t\|^2. \quad (25)$$

- 3) **Distance to goal:** Pedestrians typically try to get to the target position as close as possible, given a time horizon

$$f_{\text{goal}} = \|x_h^N - x_{\text{goal}}\|. \quad (26)$$

- 4) **Avoiding static obstacles:** Pedestrians avoid static obstacles in the environment

$$f_{\text{obstacle}} = \sum_{t=1}^N \exp\left(-\frac{\|x_h^t - o_{\text{closest}}^t\|^2}{D_o^2}\right) \quad (27)$$

where $o_{\text{closest},t}$ is the position of the closest obstacle to the human at time t and D_o is a scaling factor.

- 5) **Collision avoidance with the robot:** Pedestrians tend to avoid each other in navigation. We assume that they avoid the robot in a similar fashion

$$f_{\text{avoidance}} = \sum_{t=1}^N \exp\left(-\frac{\|x_h^t - x_r^t\|^2}{D_r^2}\right). \quad (28)$$

- 6) **Avoiding the front side of the robot:** In addition to avoiding the robot, we observe that humans tend to not cut in front of the robot, especially when they think that the robot has priority. This behavior is captured by the feature

$$f_{\text{front}} = \sum_{t=1}^N \exp\left(-(x_h^t - x_f^t)^T \mathbf{D}_f^{-1} (x_h^t - x_f^t)\right) \quad (29)$$

where x_f^t is a position in front of the robot, and \mathbf{D}_f scales the Gaussian and aligns it with the robot's orientation.

The features (except for velocity and acceleration) are visualized in Fig. 5. Human and robot trajectories are from a demonstration we collected to train the IRL model. The figure shows that the human indeed avoided low-reward regions (cool color) and navigated to the high-reward region (warm color).

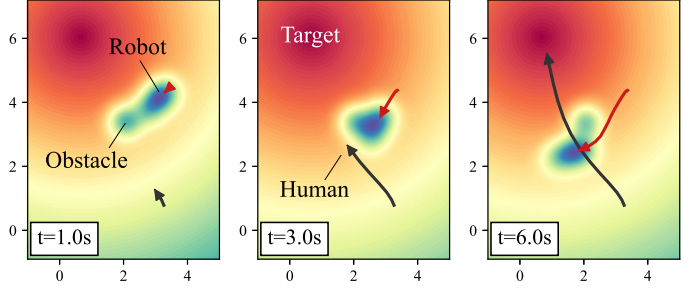


Fig. 5. Visualization of (a subset of) the features. Warm color indicates high reward and cool color indicates low reward. Human trajectory is in black and robot trajectory is in red. Arrows indicate the positions and moving directions at the specified time step.

Rewards for the Behavior Mode Model: In general, it is challenging to define the expected reward function of the human $R_h^\psi(\xi_r^{-N:0}, \xi_h^{-N:0})$, as it is not measurable. Here, we design two reward functions based on our observations. Data-driven methods, such as IRL, can potentially produce more accurate results. Our goal, however, is to derive an approximate model that can inform the planning algorithm. Therefore, we prefer to start with simpler designs.

1) **Human Priority** ($\psi = \psi_h$): When the robot's intent is to yield priority to the human, it should avoid moving directly toward the human and getting too close. We use a reward function

$$R_h^{\psi_h} = - \sum_{t=-N}^0 \frac{\hat{n}_t \cdot (\dot{x}_r^t - \dot{x}_h^t)}{\max(\|x_r^t - x_h^t\|^2, 1.0)} \quad (30)$$

where \hat{n}_t is a unit vector that points from the robot to the human: $\hat{n}_t = (x_r^t - x_h^t) / \|x_r^t - x_h^t\|$. Here, $t \leq 0$ means that the summation is over the previous time steps.

2) **Robot Priority** ($\psi = \psi_r$): When the robot has priority, the human would expect it to move with a desired velocity, regardless of the state of the human

$$R_h^{\psi_r} = - \sum_{t=-N}^0 \frac{|v_r^t - v_d|}{\max(\|x_r^t - x_h^t\|^2, 1.0)}. \quad (31)$$

Here, v_r^t is the robot's speed and v_d is the desired speed.

In (30) and (31), the reward at each time step is divided by $\max(\|x_r^t - x_h^t\|^2, 1.0)$ because when the robot is relatively far from the user, its actions carry less information about its intent. We use the $\max()$ operation to achieve numerically stable results.

B. Human Model Evaluation

1) **Data Collection for Learning Human Model:** We collected navigation demonstrations from four users. Each user was asked to walk back and forth between two target locations (setup is similar to that of Section VII). At the same time, a mobile robot (TurtleBot2, Open Source Robotics Foundation, Inc.) also moved between a set of target positions in the environment. The target positions were selected so that the robot would cross path

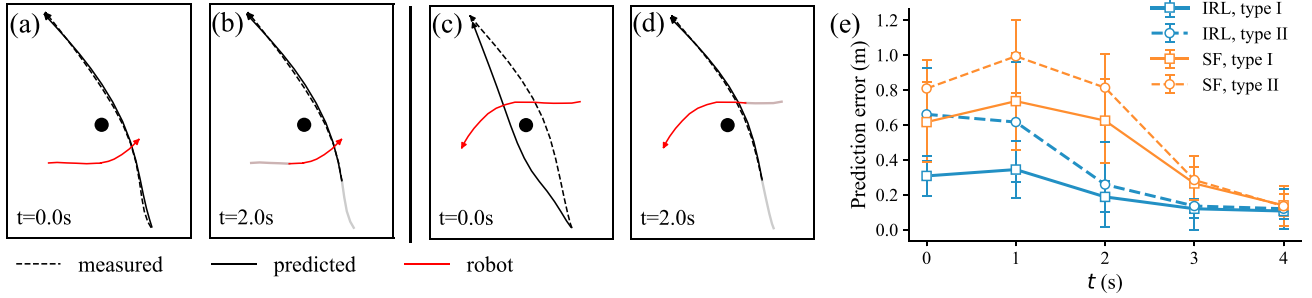


Fig. 6. Sample predictions using the learned reward functions and cross validation. (a) and (b) Example where the prediction matches the actual measurement. (c) and (d) Example where initially the prediction does not match the measurement. We use the model to repredict human actions at each time step, and the prediction starts to match measurement after 2 s. (e) Cross validation of prediction accuracy. Type I corresponds to cases where the model initially predicts the same homotopy class as the demonstration [(a), (b)], and type II corresponds to cases where the model initially predicts a different homotopy class [(c), (d)]. Blue indicates errors of the IRL model, and orange indicates errors of the SFM.

with the user. The user was allowed to walk freely as in a normal human environment, while avoiding collision with the robot.

We collected data for two scenarios—the human priority scenario and the robot priority scenario. When the user had priority, the robot would slow down and yield to the user upon encounter (we decreased the maximum allowable speed of the robot as it got close to the user). When the robot had priority, it would simply ignore the user. In each condition, the user was told the intent of the robot, so the user always knew the robot’s true intent. To distract the user from focusing solely on the robot, we asked the user to remember two numbers at one target location, and answer arithmetic problems using the two numbers at the other target location.

For each user, we recorded 64 trials for each scenario (a trial is moving from one target location to the other). We used 80% of the data to train the model, and the other 20% as the testing set.

2) Prediction With Learned Reward Function: Using MaxEnt IRL, we recover the human reward functions in human priority and robot priority scenarios. The reward functions can be used to compute the most likely continuous navigation actions using (21). This computes the actions over a fixed time horizon. To predict the entire trajectory, we use an MPC approach—we compute the most likely actions starting from time step k , predict the human state at $k + 1$, and recompute the most likely actions starting from $k + 1$.

Fig. 6 presents two prediction examples. In the first example [see Fig. 6(a) and (b)], the predicted trajectory matches the measured trajectory very well (same homotopy class, type I). In the second example [see Fig. 6(c) and (d)], the model initially predicts a trajectory that passes the obstacle on a different side (different homotopy class, type II). Although the prediction does not match the measured trajectory, it is still reasonable. In this example, both the predicted and measured trajectories pass behind the robot. Passing on either side of the obstacle has little effect on the reward and is almost equally likely. When predicting at $t = 2.0$ s, because the human already started to move toward one side of the obstacle, the model is able to generate a more accurate prediction.

The cross validation result is shown in Fig. 6(e). The prediction error is calculated as the average Euclidean distance

between the predicted and measured trajectories

$$e = \frac{1}{N'} \sum_{t=1}^{N'} \|x_{\text{pred}}^t - x_{\text{meas}}^t\| \quad (32)$$

where N' is the total number of time steps. We compute the prediction error separately for cases where the initial prediction is in the same homotopy class as the measurement (type I), and cases where the initial prediction and the measurement are in different homotopy classes (type II). It can be observed that the initial prediction error is much smaller for type I, but both types become more accurate as the human gets closer to the goal. Compared with a SFM we used previously [20], the model described here performed better for this scenario.

3) Evaluation of the Behavior Mode Model: In addition to continuous trajectories, our model can also predict behavior modes, or equivalently, the belief of the human over the robot’s intent. As it is impossible to measure this belief in an experiment, we aim to show that the prediction is reasonable with a few test scenarios.

Fig. 7(a) presents a scenario where the robot slowed down to let the user pass first. The scenario is illustrated with the three plots in the top row showing trajectories and the environment at different time steps. The data are from one demonstration we collected for training the IRL model. The plot in the bottom row shows the predicted belief $p(\theta = \theta_h)$ over time, given no explicit communication, and two different explicit communication actions at $t_c = 0.5$ s—communicating human priority ($u_c = \theta_h$) and communicating robot priority ($u_c = \theta_r$). In the case of no communication, the belief decreases initially, but rises up as the robot stops to yield to the user. This result matches the intuition that the user can infer the robot’s intent (human priority) to some extent by observing its movement (slowing down). When the robot communicates human priority at the beginning, the belief jumps to a high value and maintains this value afterward. An interesting test scenario is when the robot communicates robot priority, which is not its true intent. According to our model, initially, the belief drops, indicating that the user believes the explicit communication. However, when the robot slows down to let the user pass first, the belief starts to increase, indicating

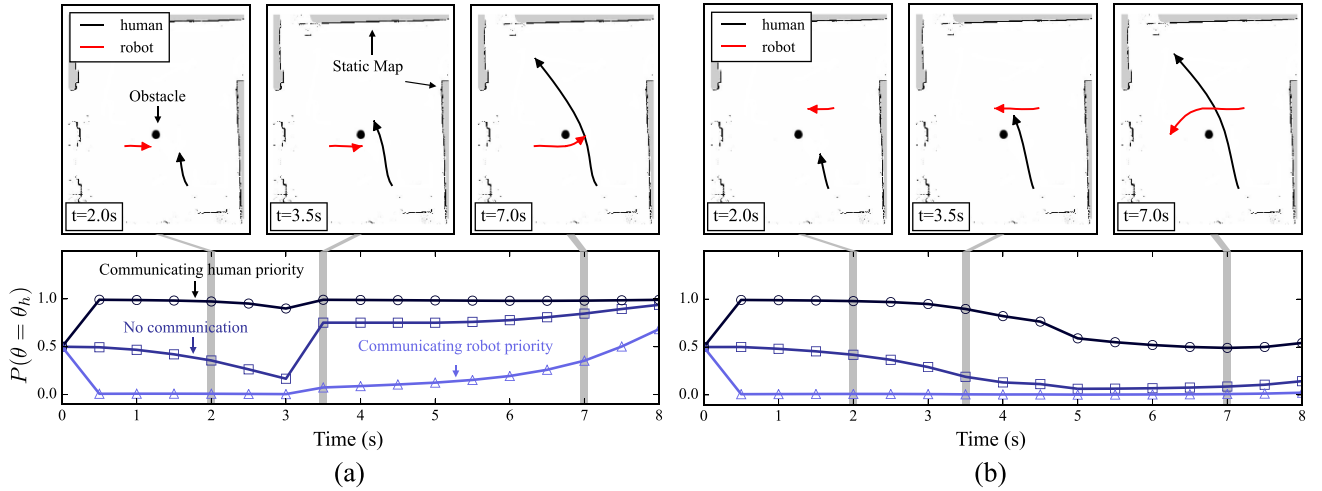


Fig. 7. Demonstrations of the behavior mode model in different scenarios. (a) A scenario where the robot slowed down to let the user pass first. Top plots show the trajectories of the robot and the user, and a map of the environment at three different time steps. Bottom plot shows the user's belief over the robot's intent (human priority) over time, given different explicit communication at $t = 0.5$ s, predicted by our model. (b) A scenario where the robot didn't slow down and passed first.

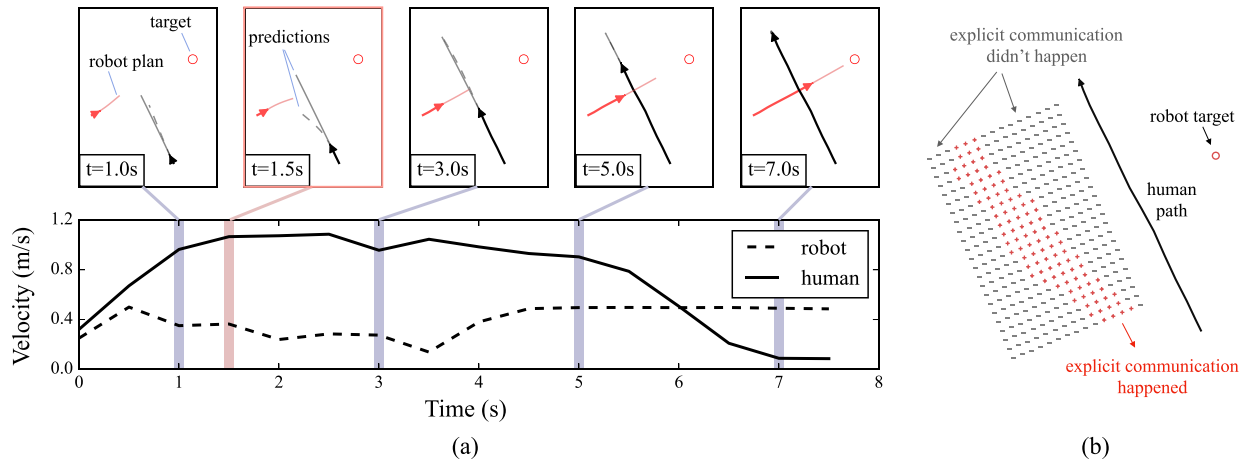


Fig. 8. (a) Simulated scenario using the proposed interactive planner. In this scenario, the robot slows down and explicitly communicates its intent (human priority) to the human. Top: Trajectory snippets at five individual time steps. Black and red arrows represent the positions and moving directions of the human and the robot at the given time step. Thin, solid red lines represent the robot's planned trajectories, thin solid/dashed black lines represent predicted human trajectories. Bottom: Velocity profile of the robot and the human. (b) Various starting positions of the robot versus whether explicit communication is used by the planner for the same scenario. Red region represents the starting positions where explicit communication is used.

that the user starts to believe otherwise as she observes the robot movements.

We also tested scenarios where the robot's intent is to give itself priority. An example is shown in Fig. 7(b). Similarly, we show the change in belief, given different explicit communication, in the bottom plot. These examples demonstrate that our model can capture the effect of both explicit and implicit communication, and the prediction matches our intuition.

C. Case Study in Simulation

Before conducting a user study with a physical mobile robot (described in Section VII), we first tested our algorithm in simulation. The purpose of the simulation is twofold—first, to validate that the proposed algorithm can work in idea setups, and

second, to select appropriate parameters that produce reasonable behaviors.

Fig. 8(a) illustrates a simulated scenario where the human and the robot move along orthogonal paths. Here, the simulated human user follows a predefined trajectory, and the trajectory is deterministic regardless of the robot's actions. While this is not realistic, the purpose is to test whether the planner can generate reasonable robot behaviors. We will further validate the effectiveness of the plan with real-world user studies. The subplots in the top row show the scenario at five different time steps, and the subplot in the bottom row shows the speed of the user and the robot over time. We can observe that the robot starts to slow down at $t = 1$ s, and explicitly communicates its intent to the user at $t = 1.5$ s. The robot continues moving slowly to allow the user to pass first, and then, speeds up toward its goal.

To understand why the planner does this, we visualize the generated plan and the predictions of human movements, as shown in Fig. 8(a). As the robot speeds up and approaches the intersection, the user becomes unsure of how to avoid the robot. This is reflected in the second subplot in the first row, where the user's possible future trajectories diverge. In this simulation, we set the reward function coefficients $c_h \gg c_r$, so that the planner cares more about the human user's efficiency and comfort. As a result, the planner explicitly communicates human priority to the user to minimize the chance that the user would slow down to yield to the robot.

We tested the planner with various robot starting positions and target positions. Fig. 8(b) presents the relationship of robot starting position and whether explicit communication is generated by the planner, for one specific target position. We can see that explicit communication is only used if the robot starts within the banded region in red. When the robot starts too close to the intersection (upper-right grey region), the planner predicts that the user will let the robot pass first as it is closer to the intersection. When the robot starts too far (bottom-left grey region), the planner predicts that the user will pass first. Explicit communication is used only when it becomes ambiguous who should pass first.

We also studied the effect of the coefficients c_h and c_r in the reward function in (19). Setting $c_h \gg c_r$ resulted in a submissive robot that would yield to the user when there was a potential collision. Conversely, setting $c_h \ll c_r$ resulted in aggressive robot behaviors. When $c_h \approx c_r$, the robot tends to be submissive. In our implementation, the robot is usually slower than the human, so it is often more efficient to let the user pass first. Based on the simulation results, we decided to experimentally test two sets of parameters: $c_h : c_r = 1 : 5$ as *robot-prioritized* scenario and $c_h : c_r = 2 : 1$ as *human-prioritized* scenario.

VII. EXPERIMENTAL RESULTS

Simulation results from the previous section have demonstrated that the planning framework can generate appropriate implicit and explicit communication actions. This section presents a real-world user study to further test the effectiveness of the proposed planning framework.

A. Experimental Setup

We again used a TurtleBot2 as the mobile robot platform. The robot was equipped with a Hokuyo URG-04LX-UG01 laser range finder and an ASUS Zenbook UX303UB as the onboard computer. The planning algorithm is implemented on a desktop computer with Intel Core i7 processor and 16 GB RAM. The two computers communicate with each other through a wireless network.

The onboard computer processes the laser range finder readings to localize the robot and estimate the position and velocity of nearby pedestrians. We localize the robot in a static map using laser-based Monte Carlo localization [52]. The robot detects and tracks pedestrians using a leg detection algorithm provided by robot operating system (ROS) [53]. The algorithm first attempts to segment leg objects from laser range finder

readings with a pretrained classifier, then pairs nearby legs and associates the detection with a tracker for each person. The localization and tracking results are sent to the desktop computer that runs our planning algorithm. The planning algorithm computes the plan and sends it back to the robot.

The explicit communication is displayed to the user via a wearable haptic interface that consists of a single vibration motor [Haptuator Mark II, shown in Fig. 9(d)]. The interface is capable of rendering distinct signals by modulating the vibration pattern. In this experiment, we display haptic cues with different vibration amplitudes and durations to indicate the robot's intent. Robot priority is represented by a single long vibration (duration 1.5 s, max current 250 mA), and human priority is represented by three short pulses (0.2 s vibration with 0.2 s pause in between, max current 150 mA). To train the user interpreting the haptic cues, we first display cues in random order (three trials for each vibration pattern) and explain the meaning. Then, we test the user with randomized cues. We found that all users can recognize each cue correctly. The training process is done for each user before the actual experiment.

The experiment field is illustrated in Fig. 9(a), which is a room of size 8×10 m. Two tables are placed at the two ends of the field as the targets for the user. To distract the user from focusing solely on the robot, we place questionnaires on the tables, and ask the user to remember and answer arithmetic questions. The entire experiment is recorded with an overhead camera (GoPro Hero 4, recording at 60 Hz), and we postprocess the video to extract the trajectories of the user and the robot. To facilitate tracking, we ask the user to wear a purple hat, and we attach an ArUco marker [54] on the top of the robot.

B. Experimental Design

We conducted a 3×1 within-subject experiment, manipulating the communication mode (baseline, implicit only, explicit + implicit) and measuring users' path lengths, whether they pass the robot in the front, and their trust in the robot. Details of the experiment design and procedure are described below.

1) *Manipulated Factors*: We manipulate one factor—the *communication mode*. The user's performance under each communication mode reflects the effectiveness of the communication method. In addition, the experimental trials consist of different *task priorities* and *robot starting positions*. We randomize the order of task priority so that the user does not know the robot's true intent beforehand. We also vary the starting position of the robot to create richer interaction scenarios and prevent the user from developing a fixed pattern (e.g., always pass in front of the robot regardless of communication). Overall, the experiment is divided into three sessions based on the communication mode.

- i) *explicit + implicit*: The robot communicates its intent to the user both explicitly (via haptic feedback) and implicitly (by changing speed and direction), using a model to predict the user's movement.
- ii) *implicit only*: The robot does not plan for explicit communication, but still changes speed and direction according to a model that predicts the user's movement.

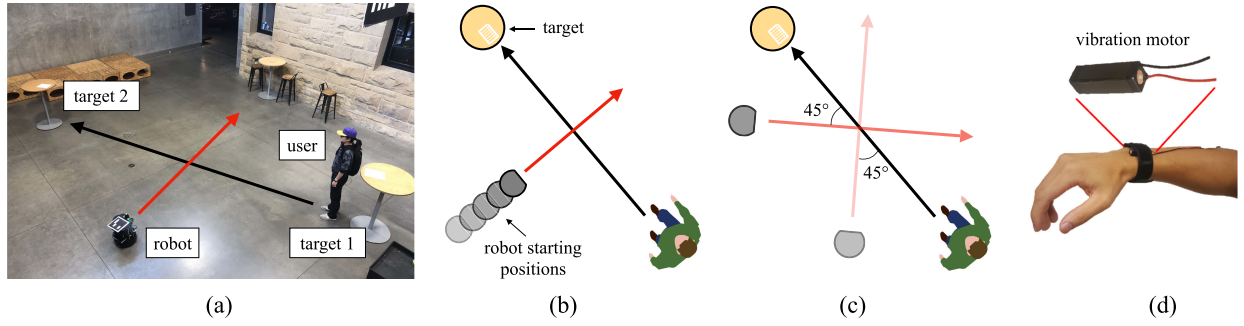


Fig. 9. (a) Experimental field. (b) and (c) Different starting and goal positions for the robot. (d) Wearable haptic interface and the vibration motor.

- iii) *baseline*: The robot simply performs collision avoidance with the user without predicting the user's movement.

Each session consists of 20 trials. In each trial, the human user and the robot both move from a target position to a goal position, as shown in Fig. 9(b). The robot is assigned one of the two task priorities in each trial—robot priority ($c_h : c_r = 1 : 5$) or human priority ($c_h : c_r = 2 : 1$). We also vary the starting position of the robot. The starting positions can be classified as *far* (four out of 20 trials), *close* (4/20), and *normal* (12/20), based on the distance to the user's starting position [see Fig. 9(b)]. There are equal numbers of high-priority trials and low-priority trials for each type of starting position. The order of the trials for each communication mode is pseudorandomized. We counterbalance the order of three communication modes among all users using a Latin square design. The first ten trials are treated as training trials and not included for analysis. In the training trials, we will reveal the robot's true intent to the user after each trial. We never tell the robot's intent to the user in the experimental trials. The purpose of the training trials is to familiarize the user with the experimental procedure and the robot's behavior. After the experiment, we ask the user to rate her trust in the robot on a scale from 1 to 7, 1 being the least trust and 7 being the most trust. The exact question we use in the survey is "please rate your overall trust in the robot in the first/second/third condition." There are seven choices under each question and we explain to the user the meaning of the scale.

In addition to the orthogonal encounter scenario, we also tested the planner in other experimental setups. Fig. 9(c) presents human-robot encounter at different angles ($+45^\circ$ and -45°). We found that the planner worked similarly in these scenarios.

2) *Dependent Measures*: We measure the user's path length for each trial. Path length is commonly used to estimate the user's efficiency and effort in navigation tasks. We also measure whether the user passes the robot in the front, which is an indicator of whether the user understands the robot's intent correctly (the user is expected to pass the robot in the front in human priority trials and from the behind in robot priority trials, if they understand the robot's intent correctly). This measurement reflects the transparency of the robot's behavior. Finally, we measure the user's trust in the robot. Trust is measured using a postexperiment survey as described previously. In the session where explicit communication is allowed, we also record whether explicit communication happens and the time at which it happens.

- 3) *Hypothesis*: We hypothesize as follows.

- i) Using *explicit + implicit* communication conveys the robot's intent more clearly than *implicit only* and *baseline*, such that users will elect to pass in front of or behind the robot as appropriate for a given priority.
- ii) The user's average path length is shorter when the robot plans for communication with the human model (*explicit + implicit* and *implicit only* modes).
- iii) The user's overall trust rating in the robot is higher when the robot plans for *explicit + implicit* communication, compared to the other two conditions.

4) *Subject Allocation*: A total of 12 people (seven males and five females, age from 23 to 33) participated in the experiment after giving informed consent, under a protocol that was approved by the Stanford University Institutional Review Board. We used a within-subject design and counterbalanced the order of the three sessions.

C. Analysis and Results

Fig. 10 and Table I summarize major results. We describe the analysis and results in detail in the following subsections.

1) *Understanding Robot Intent*: To characterize how often the user correctly understood the robot's intent, we compute the percentage of trials that the user passed in front of the robot for each communication mode and task priority. A one-way repeated measures ANOVA revealed a significant effect on percentage passing in the front for the communication mode ($F(2, 33) = 14.61, p < .001$ for human priority trials, $F(2, 33) = 5.75, p = .007$ for robot priority trials). We performed the statistical test separately for robot and human priority trials. We performed a *post hoc* analysis with Tukey HSD to determine pairwise differences. For human priority trials, results show that the baseline mode is significantly different from the implicit only mode ($p = .005$) and the implicit + explicit mode ($p < .001$). For robot priority trials, the implicit + explicit mode is significantly different from the implicit only mode ($p = 0.027$) and the baseline mode ($p = 0.01$). This supports Hypothesis I that the user can better understand the robot's intent with explicit communication.

2) *Path Length*: Path length is a commonly used metric to measure efficiency. We computed average path length (over all trials of each condition and robot intent) of each user. Similar to the metric of passing in the front, we performed a

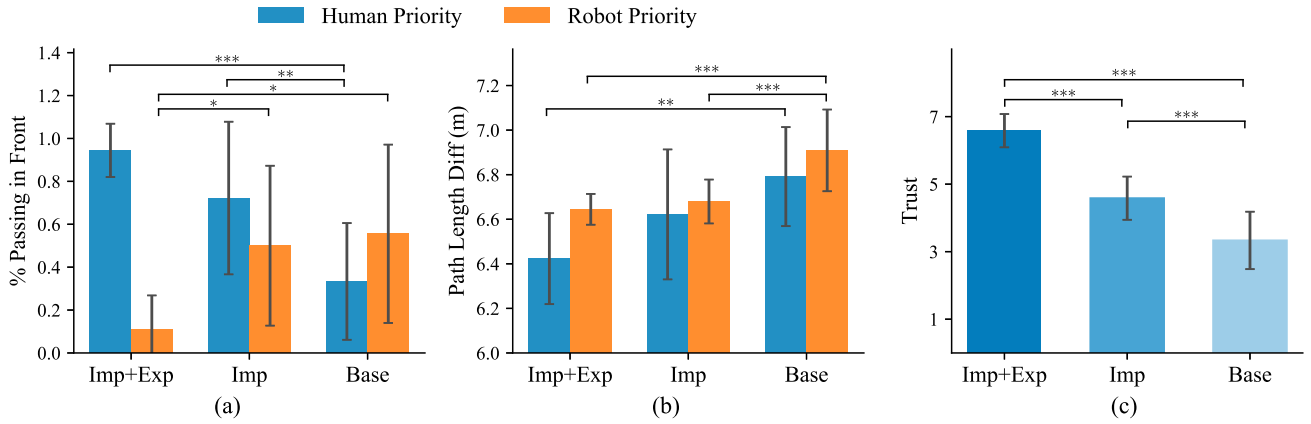


Fig. 10. Comparison of different metrics for three experimental conditions. (a) Percentage of time that the user passed in front of the robot. (b) Average path length. A base length (6.2 m) is subtracted from all the values for visualization purpose. (c) User's trust in the robot. Brackets indicate statistical significance (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$).

TABLE I
MEAN AND STANDARD DEVIATION OF MEASUREMENTS

Measure	Imp+Exp	Imp	Base
% Passing in Front	0.94 ± 0.12	0.72 ± 0.36	0.33 ± 0.27
	0.11 ± 0.15	0.50 ± 0.37	0.55 ± 0.41
Path Length	6.42 ± 0.20	6.62 ± 0.29	6.79 ± 0.22
	6.64 ± 0.07	6.68 ± 0.10	6.91 ± 0.18
Trust	6.58 ± 0.49	4.58 ± 0.64	3.33 ± 0.85

For the first two measurements, the first row is human priority and the second row is robot priority.

one-way repeated measure ANOVA analysis, and results show a significant effect ($F(2, 33) = 6.36, p = .005$ for human priority trials, $F(2, 33) = 14.18, p < .001$ for robot priority trials). *Post hoc* comparisons reveal that, for human priority trials, the average path length of the implicit + explicit mode is significantly shorter than the baseline mode ($p = .003$). For robot priority trials, the average path length of the baseline mode is significantly longer than the implicit only mode ($p < .001$) and implicit + explicit mode ($p < .001$). The result suggests that the user can navigate more efficiently when the robot plans for communication (Hypothesis II).

3) *Trust*: Past research has shown that the human would trust the robot more if it can express its intent via motion [55] or other forms of communication [40], and we expect similar results in our experiment. Here, users' trust in the robot was measured with a postexperiment survey, as explained in the previous subsection. A one-way repeated measure ANOVA analysis revealed a significant effect for communication mode ($F(2, 33) = 64.5, Fp < .001$). *Post hoc* Tukey HSD analysis showed that all three pairs are significantly different from each other ($p < .001$ for all), which supports Hypothesis III.

D. Effect of Communication

Fig. 10(a) shows that when there was both implicit and explicit communication, users were able to understand the robot's intent most of the time and acted accordingly—users passed in front of the robot 92% of the time in human priority trials, and only 9% of the time in robot priority trials. When there was only

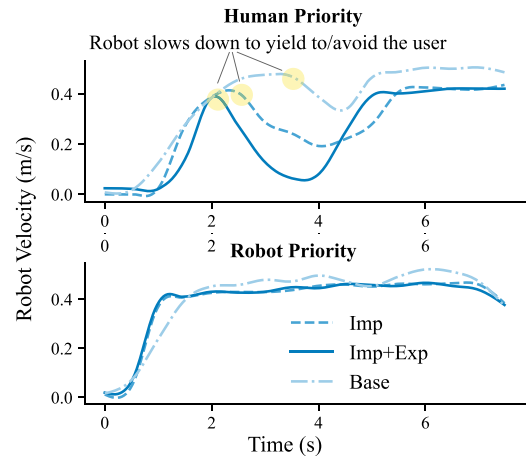


Fig. 11. Velocity profiles from sample trials. Top: Human priority trials. Bottom: Robot priority trials.

implicit communication, users were more confused, especially in robot priority trials (about 50% of the time users passed in front of the robot despite that the robot's intent was not to yield). When the robot performed collision avoidance naively (baseline condition), users were the most confused, and acted more conservatively.

Fig. 11 depicts the robot's velocity over time from sample trials of different communication modes and task priorities. In human priority trials (top row), when the robot planned for communication (implicit only and implicit + explicit modes), it started to slow down relatively early, compared to the baseline mode. By doing so, the robot implicitly expressed its intent. Indeed, significantly more users chose to pass in front of the robot with these communication modes [see Fig. 10(a)]. The reason that the planner generated this behavior is that, with the human model, the planner was able to predict the effect of robot motions on the user's navigation behavior. More specifically, our human model would predict that, given the robot slows down early, the human is more likely to believe that the robot is yielding priority and will choose to pass first, which would generate higher expected rewards. This encouraged the robot to

slow down earlier. On the other hand, the baseline planner did not perform any prediction and simply solved for a collision-free trajectory. As a result, the robot did not slow down until it got very close to the user, which caused the user to believe that it did not want to yield. The result suggests that legible motions emerge when reactions of the human are taken into consideration during planning.

Our results also suggest that users can navigate more efficiently when the robot planned for communication. Fig. 10(b) shows that for human priority trials, users' path length is significantly shorter in the implicit + explicit mode than the baseline mode. For robot priority trials, both the implicit + explicit mode and the implicit only mode result in significantly shorter path lengths than the baseline mode. Knowing the robot's intent, users can better coordinate their movements with the robot to navigate more efficiently. In the implicit + explicit mode, users in general got feedback early in the trial. In the implicit only mode, users had to infer the robot's intent, and tended to be less efficient. This result does not show statistical significance; users' behaviors have high variance in the implicit only mode. In Fig. 10(a), we observe the same high variance in the implicit only mode.

An important feature of the study is that we never explained to the user how the robot may move differently in the three communication modes. One of the main purposes of the training trials is to let the user figure out the robot's movement pattern. We observe that users, in general, were able to adapt to the robot's communication mode in a few trials. Most users started conservatively, trying to keep relatively large distances from the robot. In the implicit + explicit mode, most users adapted to trust the explicit communication quickly. In the baseline mode, users made their decisions largely based on the initial distance from the robot, as they found it difficult to tell the robot's intent from its motion. Interestingly, in the implicit only mode, we observed that users' behaviors varied based on their strategy of interacting with the robot. Aggressive users tend to realize earlier that the robot would slow down when it wanted to yield priority; they were more likely to cut in front of the robot. This helped them to interpret the implicit communication. Timid users, on the other hand, tend to keep larger distances or let the robot pass first. Sometimes this behavior resulted in the robot not slowing down in human priority trials. Unlike legibility planning as in [31], our algorithm does not explicitly optimize legibility. In addition, the algorithm responds to human actions in real time. As a result, the robot would not wait if the user is too far, even in human priority trials, which could confuse the user. This explains in part the high variance observed in the implicit only mode.

In the implicit + explicit mode, the robot may not always use explicit communication. Again, this is the result of optimizing the overall reward function. We assigned a negative reward for performing explicit communication to prevent overcommunicating. In trials where the robot started too close to or too far from the intersection point, it often chose to not use explicit communication because the chance of collision is extremely low.

Finally, we show that users gained more trust in the robot with the proposed planner in Fig. 10(c). We found statistical significance between each pair of communication modes. When the robot communicated both implicitly and explicitly, users trusted the robot the most. This is also the condition where users

understood the robot's intent the best [based on the previous analysis shown in Fig. 10(a)]. The correlation suggests that the easier users can understand the robot's intent, the more they trust the robot—transparency is beneficial in social navigation. When users do not trust the robot, they tend to act conservatively and be less efficient (fewer trials where the user passed in front of the robot, and longer path lengths in the baseline mode). The trust measurement in this work is exploratory, as we only asked users to rate their overall trust in the robot. In order to obtain more accurate results, a more specific and validated questionnaire should be used.

VIII. CONCLUSION

In this article, we presented a planning framework that leverages implicit and explicit communication to generate efficient and transparent social navigation behaviors. This approach for mobile robot proactive communication is inspired by humans' ability to use both explicit and implicit communication to avoid collisions during navigation. The planner relied on a new probabilistic model that predicts human movements. We evaluated the planner both in simulation and in a user study with a physical mobile robot. Results showed that the proposed planning algorithm can generate proactive communicative actions, which better expressed the robot's intent, reduced users' effort, and increased users' trust of the robot compared to collision avoidance with and without a model that predicts users' movements.

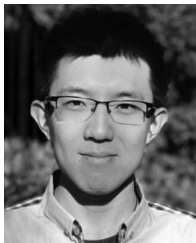
There are numerous ways to expand on this work. First, the model of human navigation behavior makes certain assumptions and approximations as described in Section IV. As a result, the planner cannot deal with certain scenarios (e.g., if the human suddenly stops and remains stationary). Expanding the data collection for IRL, including additional behaviors, and relaxing assumptions, would improve the model and thus, the planner. Second, our planner is computationally expensive. One approach to reducing the computational complexity of planning is to use sampling-based methods, instead of performing optimization. Third, we can generalize our approach to consider richer interaction scenarios and other communication modalities. We are also interested in identifying the advantages of combining implicit and explicit communication in other application scenarios. For example, in collaborative tasks, such as assembly, prompt explicit communication can be used to align each agent's goal, while implicit communication, such as forces, can improve physical interactions between partners. Finally, we would like to explore in future work how implicit and explicit communication affects users' trust toward the robot. The result in this article is an initial demonstration of the benefit of communication in terms of trust. However, more studies, such as looking into specific aspects of trust [56], are needed to gain deeper understanding.

While this work only explored a relatively simple scenario, the result has clearly demonstrated the benefit of combining implicit and explicit communication in robot social navigation. We believe that the computational framework for modeling and planning communicative actions can potentially be adapted to other applications and improve interactions between robots and humans.

REFERENCES

- [1] P. Trautman, J. Ma, R. M. Murray, and A. Krause, "Robot navigation in dense human crowds: The case for cooperation," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2013, pp. 2153–2160.
- [2] G. Ferrer and A. Sanfeliu, "Proactive kinodynamic planning using the extended social force model and human motion prediction in urban environments," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2014, pp. 1730–1735.
- [3] M. Kollmitz, K. Hsiao, J. Gaa, and W. Burgard, "Time dependent planning on a layered social cost map for human-aware robot navigation," in *Proc. IEEE Eur. Conf. Mobile Robots*, 2015, pp. 1–6.
- [4] H. Kretschmar, M. Spies, C. Sprunk, and W. Burgard, "Socially compliant mobile robot navigation via inverse reinforcement learning," *Int. J. Robot. Res.*, vol. 35, no. 11, pp. 1289–1307, 2016.
- [5] B. Kim and J. Pineau, "Socially adaptive path planning in human environments using inverse reinforcement learning," *Int. J. Social Robot.*, vol. 8, no. 1, pp. 51–66, 2016.
- [6] M. Pfeiffer, U. Schwesinger, H. Sommer, E. Galceran, and R. Siegwart, "Predicting actions to act predictably: Cooperative partial motion planning with maximum entropy models," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2016, pp. 2096–2101.
- [7] Y. F. Chen, M. Everett, M. Liu, and J. P. How, "Socially aware motion planning with deep reinforcement learning," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2017, pp. 1343–1350.
- [8] L. Peterson and M. J. Peterson, "Short-term retention of individual verbal items," *J. Exp. Psychol.*, vol. 58, no. 3, pp. 193–198, 1959.
- [9] C. Breazeal, C. D. Kidd, A. L. Thomaz, G. Hoffman, and M. Berlin, "Effects of nonverbal communication on efficiency and robustness in human-robot teamwork," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2005, pp. 708–713.
- [10] R. A. Knepper, C. I. Mavrogiannis, J. Proft, and C. Liang, "Implicit communication in a joint action," in *Proc. ACM/IEEE Int. Conf. Human-Robot Interact.*, 2017, pp. 283–292.
- [11] C. I. Mavrogiannis and R. A. Knepper, "Towards socially competent navigation of pedestrian environments," in *Proc. Robot., Sci. Syst. Workshop Social Trust Auton. Robots*, 2016.
- [12] A. D. Dragan and S. S. Srinivasa, "Generating legible motion," in *Proc. Robot., Sci. Syst.*, 2013. [Online]. Available: <http://www.roboticsproceedings.org/rss09/p24.html>
- [13] D. Fox, W. Burgard, and S. Thrun, "The dynamic window approach to collision avoidance," *IEEE Robot. Autom. Mag.*, vol. 4, no. 1, pp. 23–33, Mar. 1997.
- [14] T. Kruse, A. K. Pandey, R. Alami, and A. Kirsch, "Human-aware robot navigation: A survey," *Robot. Auton. Syst.*, vol. 61, no. 12, pp. 1726–1743, 2013.
- [15] D. Helbing and P. Molnar, "Social force model for pedestrian dynamics," *Phys. Rev. E*, vol. 51, no. 5, pp. 4282–4286, 1995.
- [16] F. Zanlungo, T. Ikeda, and T. Kanda, "Social force model with explicit collision prediction," *Europhys. Lett.*, vol. 93, no. 6, 2011, Art. no. 68005.
- [17] F. Farina, D. Fontanelli, A. Garulli, A. Giannitrapani, and D. Prattichizzo, "Walking ahead: The headed social force model," *PloS ONE*, vol. 12, no. 1, 2017, Art. no. e0169734.
- [18] P. Ratsamee, Y. Mae, K. Ohara, M. Kojima, and T. Arai, "Social navigation model based on human intention analysis using face orientation," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2013, pp. 1682–1687.
- [19] D. Mehta, G. Ferrer, and E. Olson, "Autonomous navigation in dynamic social environments using multi-policy decision making," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2016, pp. 1190–1197.
- [20] Y. Che, C. T. Sun, and A. M. Okamura, "Avoiding human-robot collisions using haptic communication," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2018, pp. 5828–5834.
- [21] J. Van Den Berg, S. J. Guy, M. Lin, and D. Manocha, "Reciprocal n-body collision avoidance," in *Proc. Robot. Res.*, 2011, pp. 3–19.
- [22] S. J. Guy, M. C. Lin, and D. Manocha, "Modeling collision avoidance behavior for virtual humans," in *Proc. Int. Conf. Auton. Agents Multiagent Syst.*, 2010, vol. 2, pp. 575–582.
- [23] P. Trautman and A. Krause, "Unfreezing the robot: Navigation in dense, interacting crowds," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2010, pp. 797–803.
- [24] A. Y. Ng and S. Russell, "Algorithms for inverse reinforcement learning," in *Proc. Int. Conf. Mach. Learn.*, 2000, pp. 663–670.
- [25] S. Levine and V. Koltun, "Continuous inverse optimal control with locally optimal examples," in *Proc. Int. Conf. Mach. Learn.*, 2012, pp. 475–482.
- [26] P. Abbeel and A. Y. Ng, "Exploration and apprenticeship learning in reinforcement learning," in *Proc. Int. Conf. Mach. Learn.*, 2005, pp. 1–8.
- [27] B. D. Ziebart, A. L. Maas, J. A. Bagnell, and A. K. Dey, "Maximum entropy inverse reinforcement learning," in *Proc. Assoc. Adv. Artif. Intell. Conf. Artif. Intell.*, 2008, pp. 1433–1438.
- [28] M. Kuderer, H. Kretschmar, C. Sprunk, and W. Burgard, "Feature-based prediction of trajectories for socially compliant navigation," in *Proc. Robot., Sci. Syst.*, 2012. [Online]. Available: <http://www.roboticsproceedings.org/rss08/p25.html>
- [29] L. Tail, J. Zhang, M. Liu, and W. Burgard, "Socially compliant navigation through raw depth inputs with generative adversarial imitation learning," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2018, pp. 1111–1117.
- [30] A. Thomaz *et al.*, "Computational human-robot interaction," *Found. Trends Robot.*, vol. 4, no. 2/3, pp. 105–223, 2016.
- [31] A. D. Dragan, K. C. Lee, and S. S. Srinivasa, "Legibility and predictability of robot motion," in *Proc. ACM/IEEE Int. Conf. Human-Robot Interact.*, 2013, pp. 301–308.
- [32] D. Sadigh, S. Sastry, S. A. Seshia, and A. D. Dragan, "Planning for autonomous cars that leverage effects on human actions," in *Proc. Robot., Sci. and Syst.*, 2016. [Online]. Available: <http://www.roboticsproceedings.org/rss12/p29.html>
- [33] D. Sadigh, S. S. Sastry, S. A. Seshia, and A. Dragan, "Information gathering actions over human internal state," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2016, pp. 66–73.
- [34] D. Sadigh, N. Landolfi, S. S. Sastry, S. A. Seshia, and A. D. Dragan, "Planning for cars that coordinate with people: Leveraging effects on human actions for planning and active information gathering over human internal state," *Auton. Robots*, vol. 42, no. 7, pp. 1405–1426, 2018.
- [35] A. Bestick, R. Bajcsy, and A. D. Dragan, "Implicitly assisting humans to choose good grasps in robot to human handovers," in *Proc. Int. Symp. Exp. Robot.*, 2016, pp. 341–354.
- [36] S. Nikolaidis, D. Hsu, and S. Srinivasa, "Human-robot mutual adaptation in collaborative tasks: Models and experiments," *Int. J. Robot. Res.*, vol. 36, no. 5/7, pp. 618–634, 2017.
- [37] C.-M. Huang and B. Mutlu, "Modeling and evaluating narrative gestures for humanlike robots," in *Proc. Robot., Sci. Syst.*, 2013, pp. 57–64.
- [38] M. Lohse, R. Rothuis, J. Gallego-Pérez, D. E. Karreman, and V. Evers, "Robot gestures make difficult tasks easier: The impact of gestures on perceived workload and task performance," in *Proc. SIGCHI Conf. Human Factors Comput. Syst.*, 2014, pp. 1459–1466.
- [39] T. Hashimoto, S. Hiramatsu, T. Tsuji, and H. Kobayashi, "Realization and evaluation of realistic nod with receptionist robot SAYA," in *Proc. IEEE Int. Symp. Robot Human Inter. Commun.*, 2007, pp. 326–331.
- [40] K. Baraka and M. M. Veloso, "Mobile service robot state revealing through expressive lights: Formalism, design, and evaluation," *Int. J. Social Robot.*, vol. 10, no. 1, pp. 65–92, 2018.
- [41] S. Scheggi, F. Chinello, and D. Prattichizzo, "Vibrotactile haptic feedback for human-robot interaction in leader-follower tasks," in *Int. Conf. Pervasive Technol. Related Assistive Environ.*, 2012, pp. 51:1–51:4.
- [42] S. Scheggi, F. Morbidi, and D. Prattichizzo, "Human-robot formation control via visual and vibrotactile haptic feedback," *IEEE Trans. Haptics*, vol. 7, no. 4, pp. 499–511, Oct./Dec. 2014.
- [43] D. Sieber, S. Musić, and S. Hirche, "Multi-robot manipulation controlled by a human with haptic feedback," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2015, pp. 2440–2446.
- [44] B. Mutlu, F. Yamaoka, T. Kanda, H. Ishiguro, and N. Hagita, "Nonverbal leakage in robots: Communication of intentions through seemingly unintentional behavior," in *Proc. 4th ACM/IEEE Int. Conf. Human-Robot Interact.*, 2009, pp. 69–76.
- [45] H. Admoni, A. Dragan, S. S. Srinivasa, and B. Scassellati, "Deliberate delays during robot-to-human handovers improve compliance with gaze communication," in *Proc. ACM/IEEE Int. Conf. Human-Robot Interact.*, 2014, pp. 49–56.
- [46] A. Moon *et al.*, "Meet me where I'm gazing: How shared attention gaze affects human-robot handover timing," in *Proc. ACM/IEEE Int. Conf. Human-Robot Interact.*, 2014, pp. 334–341.
- [47] M. Morari, C. Garcia, J. Lee, and D. Pretz, *Model Predictive Control*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1993.
- [48] C. I. Mavrogiannis, W. B. Thomason, and R. A. Knepper, "Social momentum: A framework for legible navigation in dynamic multi-agent environments," in *Proc. ACM/IEEE Int. Conf. Human-Robot Interact.*, 2018, pp. 361–369.
- [49] A. Astolfi, "Exponential stabilization of a wheeled mobile robot via discontinuous control," *ASME J. Dyn. Syst., Meas., Control*, vol. 121, no. 1, pp. 121–126, 1999.
- [50] S. Johnson, *The Nlopt Nonlinear-Optimization Package*. (2014).

- [51] S. Hoogendoorn and P. H. L. Bovy, "Simulation of pedestrian flows by optimal control and differential games," *Optimal Control Appl. Methods*, vol. 24, no. 3, pp. 153–172, 2003.
- [52] S. Thrun, D. Fox, W. Burgard, and F. Dellaert, "Robust Monte Carlo localization for mobile robots," *Artif. Intell.*, vol. 128, no. 1/2, pp. 99–141, 2001.
- [53] M. Quigley *et al.*, "Ros: An open-source robot operating system," in *Proc. ICRA Workshop Open Source Softw.*, 2009, vol. 3, p. 5.
- [54] S. Garrido-Jurado, R. Muñoz-Salinas, F. Madrid-Cuevas, and M. Marín-Jiménez, "Automatic generation and detection of highly reliable fiducial markers under occlusion," *Pattern Recognit.*, vol. 47, no. 6, pp. 2280–2292, 2014.
- [55] A. D. Dragan, S. Bauman, J. Forlizzi, and S. S. Srinivasa, "Effects of robot motion on human-robot collaboration," in *Proc. 10th Annu. ACM/IEEE Int. Conf. Human–Robot Interact.*, 2015, pp. 51–58.
- [56] P. A. Hancock, D. R. Billings, K. E. Schaefer, J. Y. Chen, E. J. de Visser, and R. Parasuraman, "A meta-analysis of factors affecting trust in human-robot interaction," *Human Factors*, vol. 53, no. 5, pp. 517–527, 2011.



Yuhang Che received the B.S. degree from Shanghai Jiaotong University, Shanghai, China, and Purdue University, West Lafayette, IN, USA, in 2013, and the M.S. and Ph.D. degrees from Stanford University, Stanford, CA, USA, in 2015 and 2019, respectively, all in mechanical engineering.

He is currently with Waymo LLC, Mountain View, CA, USA. His research interests include human–robot interaction, motion and behavior planning, autonomous systems, teleoperation, and haptics.



Allison M. Okamura (Fellow, IEEE) received the B.S. degree from the University of California, Berkeley, Berkeley, CA, USA, in 1994, and the M.S. and Ph.D. degrees from Stanford University, Stanford, CA, USA, in 1996 and 2000, respectively, all in mechanical engineering.

She is currently a Professor of Mechanical Engineering with Stanford University. Her research interests include haptics, teleoperation, medical robotics, virtual environments and simulation, neuromechanics and rehabilitation, prosthetics, and engineering

education.

Prof. Okamura is the Editor-in-Chief for the IEEE ROBOTICS AND AUTOMATION LETTERS.



Dorsa Sadigh received the bachelor's and doctoral degrees in electrical engineering and computer sciences (EECS) from the University of California, Berkeley, Berkeley, CA, USA, in 2012 and 2017, respectively.

She is an Assistant Professor of Computer Science and Electrical Engineering with Stanford University, Stanford, CA, USA. She works specifically on developing efficient algorithms for autonomous systems that safely and reliably interact with people. Her research interests include the intersection of robotics,

learning and control theory, and algorithmic human–robot interaction.

Dr. Sadigh was the recipient of the National Science Foundation (NSF) and National Defense Science & Engineering (NDSEG) Graduate Research Fellowships, the Leon O. Chua Departmental Award, and the Arthur M. Hopkin Departmental Award.