Planning with Partner Uncertainty Modeling for Efficient Information Revealing in Teamwork

Shih-Yun Lo
University of Texas at Austin
yunl@utexas.edu

Elaine Schaertl Short
Tufts University
elaine.short@tufts.edu

Andrea L. Thomaz
University of Texas at Austin
athomaz@ece.utexas.edu

ABSTRACT

Communication among team members is important for efficient teamwork, to coordinate behavior and ensure that all team members have the information they need to complete the task. To enable effective communication and thus efficient teamwork, we propose a multi-agent planning approach to revealing information based on its benefit to joint team performance. By explicitly modeling the partner's knowledge and behavior, our approach allows a robot in a team to reason about *when* information is useful, *how* the communication is effective, and to communicate through efficient actions. That is, the robot provides only the necessary information for task completion, provides the information at the time that it is needed, and through the action(s) that optimizes team performance. We validated this approach in a human study in which participants walk together with a robot to a destination that is known only to the robot. We compared to a legible motion generation approach, and showed that users perceived our approach as more natural, socially appropriate, and fluent to team with, while being both more predictable and intent-clear. The ratings of our approach are equal or higher than legible motion across all 18 survey items.

ACM Reference Format:
Shih-Yun Lo, Elaine Schaertl Short, and Andrea L. Thomaz. 2020. Planning with Partner Uncertainty Modeling for Efficient Information Revealing in Teamwork. In *Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction (HRI '20), March 23–26, 2020, Cambridge, United Kingdom.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3319502.3374827

1 INTRODUCTION

We address the problem of robot communication in human-robot teams. To efficiently collaborate, individuals share information when parts of the task specification (subgoals, constraints, etc.) are not known to all members of the team. However, this communication can be time- and energy-consuming, communicating unnecessary information can be inappropriate or distracting, and robots need to be able to communicate efficiently without sacrificing task performance.

This work was supported by the US NSF (IIS-1564080, IIS-1724157) and US ONR (N000141612835, N000141612785).

HRI '20, March 23–26, 2020, Cambridge, United Kingdom

ACM ISBN 978-1-4503-6746-2/20/03…$15.00
https://doi.org/10.1145/3319502.3374827

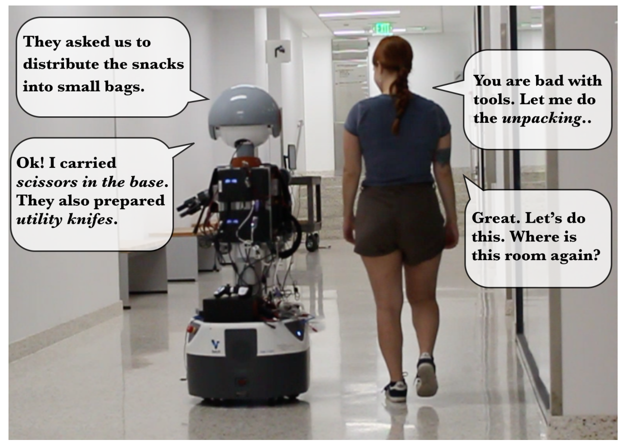

Figure 1: Simulated dialog in human-robot teaming to reach a designated location and provide service: the robot initiates communications to eliminate potential human uncertainty about task specification (here, snack distributing). Later, the robot responds with detailed specification (tool locations) given her subtask assignment.

We build on prior work that used information-revealing actions as part of task execution, and explicitly model partner uncertainty in order to reveal only relevant information during teamwork (an example is shown in Fig. 1). We use planning with nested inference in a multi-agent setting, interleaving action selection with partner inference; the optimization takes into account the impact of the action's embedded information to teamwork, which inherently balances the cost of information providing against the benefits to team performance. The planning process determines *when* to perform *which* actions that provide the most information with the least cost to robot performance. We refer to these behaviors as "corroborative acts": actions that embed intent-revealing in task execution and seamlessly deliver only the information needed to improve partner performance.

We compare this approach to prior work on information-revealing: specifically, work on generating intent-expressive, or legible, motions [7]. Legible motion has been shown to be effective in improving team performance, but has been evaluated under two assumptions: first, that it is always *necessary* to reveal information (about the goal), and second, that *sooner is better* when it comes to the timing to reveal information. However, depending on the task setting, these assumptions may not always hold: the information may not necessarily help partner performance (or may recklessly reduce robot performance), and demonstrating intent early may not be as effective compared to that at the critical timing. Information-providing gestures at the wrong time could appear inappropriate to the current task, or even confusing to the partner.

We conduct the evaluation under a guidance task, in which a robot and human walk together towards a goal known only to the robot. The two approaches are evaluated under the within-subjects design, with self-reported perceptions of teamming quality, including teamwork fluency [10], perceptions of the robot as a social agent and teammate, and intent clarity [6]. We show that our approach generates more efficient trajectory, and significantly outperforms legible motion on subjective measures, including being more fluent to team with, natural, socially appropriate, and both more predictable and intent-clear; our approach also performs equal or better across all 18 self-report questionnaire items. Furthermore, we note that our planning approach can be naturally extended to include other forms of communicative actions, and to include explicit as well as implicit communication.

2 RELATED WORK

Communication is an active research area to enable smooth human interaction with agents, including robots and machines. In dialog, the complex sentence structure and rich domain knowledge require the agent to reason about effective responses for command clarification [5] and long-duration interaction [13]. Using motions to communicate intents has been widely studied in human-robot interaction, through body language [8], gestures, and social cues for implicit communication [2, 3, 27], and has been shown to improve teamwork performance and perception of commitment [2, 8, 11].

Formulated as a robot motion planning problem, intent communication was specified as to reveal subgoal information during trajectory execution, and referred to as "legible motions", in manipulation and navigation domains [7, 16]. Similar formulation was applied to choose among abstracted action primitives for dynamic obstacle avoidance [18], and task assignment and dialog for information exchange [12]. With legibility gaining attention to equip motion planning with implicit communication and thus fluent teamwork [4, 6, 12], there remains the concern of information-overloading and exaggeration in such motions [3, 7]; moreover, the discussion on the trade-off between the cost of information revealing and the benefit to teamwork has remained in context-specific manners [7, 18].

For communication in teamwork in general, from the joint efficiency perspective, it is still unclear on: how to measure whether information-revealing is worth the cost, and how the benefits of information revealing (to smoother interaction) generalize to other domains and specifications. We address these aspects in this paper through the concept of "information value" embedded in our joint optimization formulation, and discuss the performance in a context-aware manner that can adopt different domain specifications.

We choose navigation as the target domain for evaluation, for its rich literature in human-robot interaction [29–31], essential role to enable robot service in human environments [17], and general applicability to integrate with other forms of interaction [19]. In navigation, the subgoals are of critical information for pedestrian interaction, to predict motions for collision avoidance and planning [1, 26]. When multiple agents walk together as a group, the collaborative setting requires group members to plan for one another to efficiently and smoothly adapt to environmental changes [14, 20]. The performance in subgoal identification can affect behavior patterns in interaction, e.g. ways to approach humans to initiate conversation [25, 28]. Behavior patterns highly affect human perceptions of the robot in close-distance interaction, e.g. better liking for active yielding behaviors in group navigation [20, 22]. For intent communication, in navigation domains, as goals are relatively far away from each other, efficient paths have been suggested as being more "legible" [16]. Intent communication in group navigation has remained in the form of robot following human and passive subgoal identification [14, 22], whereas leader behaviors have limited discussion on general group shape features [20].

3 PLANNING WITH TEAMMATE UNCERTAINTY MODELING

In teamwork, agents collaborate under certain task specification. A robot agent r at time t is described through its physical state $s_t^r \in S$ and task specification $\theta_t^r \in \Theta^r$, where S is the state space and $\theta_t^r \in \Theta^r$ is the set of task specifications. We assume that any agent r has full observability to the physical state in the environment and its own task specification, s_t^r and θ_t^r, and may only have partial information to that of the teammate(s), θ_t^{-r}. The uncertain information is kept in the belief state, $b_t^r \in B^r$, where B^r is the belief space. Same for agent $-r$. Here, for an agent $-r$ concerned with its uncertainty in teammate θ^r, $b_t^{-r} \in \triangle\Theta^r$, where $\triangle\Theta^r$ is the probability space over the teammate task specification [1].

In teamwork, subtask specifications are often dependent among teammates. Communication is then essential to prevent failure or inefficiency caused by false beliefs or insufficient information. Communication requires domain knowledge. We model a teammate agent $-r$ concerned with the team based on s_t^{-r}, s_t^r, θ_t^{-r} and belief over the partially-observable teammate state θ_t^r. $-r$ therefore makes decisions taking into account *self* uncertainty. On the other hand, an agent r maintaining belief over both teammate state θ_t^{-r} and teammate's belief over its own state $b_t^{-r}(\theta_t^r)$ has the ability to make decisions also based on *teammate* uncertainty, which requires *nested inference*. This way, the agent not only can reason about the effect of uncertainty to oneself, but also that to the teammates.

Interacting with nested inference, or *theory of mind*, is commonly seen among humans, to predict and infer about others' behaviors based on the explicit modeling of others and projection of oneself. For example, when people invite friends who for the first time come to their houses to cook, they introduce the kitchen, as they know the others do not know where things are yet. Without this act, it can lead to inefficiency, e.g. guests have to constantly ask for cooking utensil locations.

Decision-making with nested inference is an enhanced capability for robot teaming with humans. Yet, it should be noted that information revealing can be costly; unnecessary information exchange can also be bandwidth-exhausting and disruptive to ongoing tasks and social interaction (e.g. conversation). Consider teamwork with a robot preparing a dessert and the human making the main dish. If the robot knows the kitchen better and introduces every commonly used items, which the human may not need, the hospitality can make teamwork extremely inefficient.

[1] In case where agent $-r$ has partial observability of other variables, e.g. teammate state s_t^r, we keep them in the belief space B^{-r}

In this section, we model teamwork as a multi-agent planning problem, to evaluate action values not only to oneself but to the team. We incorporate layers of team member beliefs into the planning objectives, and formulate *information revealing* as choosing communicative actions along the teamwork optimization process given teammate uncertainty.

3.1 Teammate Modeling

When working in a shared workspace with other agents, the reward received by an agent $-r$ after taking an action ,$a_t^{-r} \in A^{-r}$, is dependent on not only its state, task specifications, but also those of the others: $R^{-r}(s_t, a_t^{-r}, a_t^r|\theta_t^{-r}, \theta_t^r)$. Here A^{-r} is $-r$'s action space, $s_t \in S^r \times S^{-r}$ denotes the joint state space. The transition function of each agent is as follows: $\mathcal{T}^r : S^r \times A^r \rightarrow S^r$ and $\mathcal{T}^{-r} : S^{-r} \times A^{-r} \rightarrow S^{-r}$, which can be truncated into: $\mathcal{T} : S \times A \rightarrow S$, where $A = A^r \times A^{-r}$. At each time t, agent $-r$ makes an (probabilistic) observation $o_t^{-r} \in O^{-r}$ out of the observation space O^{-r}, which can be used to update its belief b_t^{-r} over the hidden variables.

We define the belief space of a team member based on its layer of nested inference (over teammate states): an agent r with zero inference capability has no explicit modeling of uncertainty; we refer to such agent's policy as $\pi^{r,0} : S^r \times S^{-r} \times \Theta^r \rightarrow A^r$. In teamwork, this type of agent assumes what it knows is already the complete domain knowledge, and teammates also know such information. During planning, $\pi^{r,0}$ predicts about $-r$ through $\pi^{-r,0}$, as if the the knowledge it possesses is common knowledge. In collaborative tasks, where teammates share one objective function, the policy $\pi^{r,0}$ and its modeling of $\pi^{-r,0}$ are interchangeable:

$$a_t^{r*} = \underset{a_t^r}{\arg\max}\, \max_{a_t^{-r}} \mathbb{E}_{s_{t+1}\sim\mathcal{T}(s_t,a_t)}[R^r(s_t, a_t^r, a_t^{-r}|\theta_t^r, \hat{\theta}_t^{-r}) + V^r(s_{t+1}|\theta_t^r, \hat{\theta}_t^{-r})], \tag{1}$$

which can be solved by dynamic programming. This shared-knowledge behavior assumption is common in computational models for large-number-entity interaction simulation [9, 21]. The multi-agent planning formulation enables agents to reason about their action effects, in addition to their own task value, but also that of the partner's, and it serves as the basic teamwork model in this work.

Here we consider (human) teammates $-r$ as decision-makers under *first-order inference*, who maintain their belief spaces over Θ^r, the task specification space of agent r: $b_t^{-r} \in \triangle\Theta^r$. Since θ^r is directly observable to agent r and r's uncertainty over $-r$ task specification is not explicitly modeled in $-r$'s first-order belief, $-r$ has the mental capacity to model r through the zero-inference model, $\pi^{r,0}$, for prediction (which plans with $\pi^{-r,0}$ for prediction).

The observation function Ω^{-r} is then defined to acquire (possibly indirect) observations over Θ^r, $\Omega^{-r} : \Theta^r \times A^{-r} \rightarrow O^{-r}$. At each time step t, belief b_t^{-r} is updated based on newly received observation o_t^{-r}, the last action a_{t-1}^{-r}, and the last belief b_{t-1}^{-r} as follows :

$$b_t^{-r}(\theta_t^r) = \beta \sum_{\theta_t^r \in \Theta^r} b_{t-1}^{-r}(\theta_{t-1}^r) p(\theta_t^r|a_{t-1}^r, \theta_{t-1}^r, o_t^{-r}), \tag{2}$$

where β is a normalizing constant, and $p(\theta_t^r|\theta_{t-1}^r, a_{t-1}^r, o_t^{-r})$ is the belief propagation, calculated based on Ω^{-r} and $\mathcal{T}^{-r}$. In prediction time (for planning), o_t^{-r} is sampled under the distribution of belief b_t^{-r} using Ω^{-r} for belief propagation.

Assuming agent $-r$ is Bayesian rational, it uses $\pi^{r,0}$ for prediction and makes a decision that maximizes the expected accumulated future reward, or the value function $V(s_t^{-r}|\theta_t^{-r}, b_t^{-r})$:

$$V(s_t^{-r}|b_t^{-r}) = \max_{a_t^{-r}\in A^{-r}} \mathbb{E}_{b_t^{-r}, a_t^r\sim\pi^{r,0}}[R^{-r}(s_t^{-r}, a_t^{-r}, a_t^r|b_t^{-r})] + \mathbb{E}_{b_{t+1}^{-r}}[V(s_{t+1}^{-r}|\theta_t^{-r}, b_{t+1}^{-r}]). \tag{3}$$

The optimal action a_t^{-r*} is calculated to maximize the value estimate $V(s_t^{-r}|\theta_t^{-r}, b_t^{-r})$:

$$a_t^{-r*} = \arg\max_{a_t^{-r}\in A^{-r}} \mathbb{E}_{b_t^{-r}, a_t^r\sim\pi^{r,0}}[R^{-r}(s_t^{-r}, a_t^{-r}, a_t^r|\theta_t^{-r}, b_t^{-r})] + \mathbb{E}_{b_{t+1}^{-r}}[V(s_{t+1}|\theta_t^{-r}, b_{t+1}^{-r}]), \tag{4}$$

and we denote this policy as: $a_t^{-r*} \sim \pi^{-r,1}(s_t|\theta_t^{-r}, b_t^{-r})$, the policy of agent $-r$ with one layer of inference, under the Bayesian rationality assumption.

With one-layer inference, the form of information exchange is limited to **passive inquiry**: agents ask questions, wait, or explore/seek for feedback. The decision-making under partial observability yields agent behaviors of reasoning under uncertainty. Using this policy for prediction, it serves as a *quantitative measure* of long-term delay or performance degradation due to insufficient information. With $\pi^{-r,1}$ for prediction, the robot can reason about its action effects, not only based on its own task value, but also the partner's performance improvement *due to information gain.*

3.2 Planning under Teammate Uncertainty

We now consider agent r with belief over the agent $-r$'s first-layer inference, $b_t^{-r} \in \triangle\Theta^r$, which is the second-layer inference with belief over partner belief and its own uncertainty (possessed with first-layer inference): $b_t^r \in \triangle\Theta^{-r} \times \triangle\triangle\Theta^r$.

At each time step t, belief b_t^r over the truncated belief space is updated; it contains two subsets, one is over the first-layer belief, $b_t^r(\theta_t^{-r})$, which is updated based on newly received observation o_t^r, last action a_{t-1}^r, and last belief b_{t-1}^r. The other is over the second-layer belief, $b_t^r(b_t^{-r})$, which is a belief over the belief of teammate. To make this update, the belief propagation of teammate inference is made upon the simulated observation $\hat{o}_t^{-r}$, which can only be acquired by an approximated observation function $\hat{\Omega}^{-r}$. $\hat{\Omega}^{-r}$ samples $\hat{o}_t^{-r}$ given sampled $\hat{b}_t^{-r}$ and $\hat{\theta}_t^{-r}$ (from its belief b_t^r) and predicted $a_t^{-r} \sim \pi^{-r,1}(s_t|\hat{\theta}_t^{-r}, \hat{b}_t^{-r})$. The belief update of second-layer inference model is then:

$$b_t^r = \beta \sum_{is_t^r \in IS^r} b_{t-1}^r(is_{t-1}^{-r}) p(is_t^r|a_{t-1}^r, a_{t-1}^{-r}, o_t^r, \hat{o}_t^{-r}), \tag{5}$$

where $is^r \in IS^r$ is the interactive state in the space that agent r maintains its belief over: $IS^r = \Theta^{-r}\triangle\Theta^r$, with the second-layer inference. The value function and optimal action function are the same as in Eq. 3 and Eq. 4, except for to use the one-layer inference policy $a_t^{-r} \sim \pi^{-r,1}(s_t|\theta_t^{-r}, b_t^{-r})$ instead of the zero-layer one. The optimal policy of second-layer inference is denoted as: $a_t^{r*} \sim \pi^{r,2}(s_t|\theta_t^r, b_t^r)$, under the Bayesian rationality assumption.

Under this inference setting, information exchange can be done through **active revealing**. When agent r detects teammate uncertainty in its belief update process of $b_t^r(b^{-r_t})$, it can decide if to reveal information based on $\pi^{r,2}$. In scenarios where teammate uncertainty is certain, agents with $\pi^{r,2}$ can provide information that improves teammate performance.

3.3 Efficient and Appropriate Communication

Consider again the kitchen scenario with the robot assigned to make a dessert and the human assigned to the main dish, where the human is about to look for salt to season the dish: with teamwork planning over nested beliefs, the robot can make a minimal communication effort and show only the salt but not other possibly-relevant information; moreover, if the robot needs sugar in the same cabinet, it can grab the salt at the same time and place it where the teammate can see, which seamlessly helps teamwork with the least delay to its own task. We call this behavior a *corroborative act*; it helps its teammate with critical information, while subtly balancing its own efficiency. This behavior relies on nested inference over teammate uncertainty on domain specification (salt can location), and multi-agent planning allows the robot to reveal such information because of predicted teammate inefficiency (searching for the salt), in a way that maximizes teammate performance (by placing it somewhere visible) while minimizing the cost of its own (by grabbing the salt while getting sugar). We claim this behavior is crucial for efficient and therefore fluent teamwork.

To demonstrate the effectiveness of planning over nested beliefs for information revealing, or corroborative act, in addition to the assumption that the physical states of both human and robot agents s_t are fully observable to both agents, we focus on a setting in which: (i) the human's task depends on the robot's, which the human is uncertain of, (ii) the robot knows the human's task specification, and (iii) the human does not know the robot's task specification and the robot knows such uncertainty of the human. This setting focuses on the situation where the human teammate needs information from the robot, the robot knows how its information will affect teammate performance, and it knows that teammate does not have its information yet.

This setting applies to teamwork with production line where tasks are dependent on that of the upstream; it also applies to teaching and guidance tasks, where teammate's need and uncertainty is of public information to the instructor. By focusing on the belief over teammate belief $b_t^r(b_t^{-r})$ but omitting that on θ_t^{-r} and partial observability to s_t, we focus the discussion on the improved teamwork with communicative planning using nested beliefs, which legible motion also follows for later comparison, and omit the details on maintaining complex beliefs and making inferences over the (possibly weakly observable) belief space.

4 INFORMATION REVEALING THROUGH MOTION PLANNING

Following the setting detailed in previous section, we compare this nested inference approach for robot "theory of mind", with legible motion [7, 18], which demonstrated intent-expressive motion planning for fluent teamwork [4, 6].

Note that we choose motion planning and implicit communication as the target community in this work, given that there is literature to compare with, but the formulation can be used in other planning or integration work, such as symbolic planning and integrated task and motion planning, and planning with explicit communicative actions.

4.1 Legible Motion

In legible motion planning [7, 12], the implicit information revealing along with robot motions is achieved through optimizing the (discounted) accumulated information gain and efficiency of the trajectory ξ, which is a sequence of waypoints for the robot to track to, through the following objective function:

$$\xi^* = \arg\max_{\xi} H(\xi) f(t) - \lambda C(\xi), \tag{6}$$

where C is a domain-specific cost function, defined to generate efficient and "predictable" motions [7]; H is the "legibility" function, which incorporates the information gain for partner inference of the correct subgoal. As it was assumed in their work that information has discounted importance over time, early information revealing was encouraged in legible motions by applying *discounted weights on H over time* using $f(t)$. As legibility was introduced as an an counter factor for generating predictable motions, information revealing was treated as an "anti-efficiency" factor. For such reason, legible motion formulation used a regulator λ to balance trajectory efficiency and information gain, following the common formulation in exploratory planning [24].

4.2 Corroborative Act

Here, we instead formulate information revealing in teamwork from the perspective of **improving team performance**, and let the planner decide when it is most efficient and effective to provide information. Following the formulation in Sec.3, corroborative acts maximize team performance through:

$$a_t^{r*} = \arg\max_{a_t^r \in A^r} \mathbb{E}_{b_t^r}[R^r(is_t^r, \pi^{-r,1}(is_t^r), a_t^r)] + \mathbb{E}_{b_{t+1}^r}[V(is_{t+1}^r]), \tag{7}$$

where R^r is the robot's reward function, parametrized to encourage task completion of both teammate and itself. is_t^r includes belief over teammate belief over robot task specification, parametrized by $\theta_t^r \in \Theta^r$: $b_t^r \in \triangle(\triangle\Theta^r)$, where Θ^r is the task specification parameter space. The belief update follows that in Eq. 5, but only on the second-layer inference:

$$b_t^r = \beta \sum_{b_t^{-r} \in \triangle S^r} b_{t-1}^r(b_{t-1}^{-r}) p(b_t^{-r} | a_{t-1}^{-r}, \hat{o}_t^{-r}), \tag{8}$$

where $p(b_t^{-r} | a_{t-1}^{-r}, \hat{o}_t^{-r})$ is the simulated belief propagation of agent $-r$. With this formulation, corroborative acts balance between teamwork efficiency and the cost of information revealing; it reveals information when it is critical to improve teammate future performance and it can do so in an efficient manner, utilizing the modeling of teammate under uncertainty for teamwork performance prediction. We later refer to our approach as *NICA*: Nested Inference for Corroborative Acts.

4.3 Corroborative Act v.s. Legible Motion

During the planning process of corroborative acts, information gain is assessed by the accumulated improvement over time it brings to team performance; therefore, information is revealed only when its value is worth the cost. For example, when the human partner is assembling parts packed in boxes which are delivered by the robot, with the boxes containing identical contents, the robot revealing which box it will deliver has marginal improvement for teammate performance. The first difference between corroborative act and legible motion is that the robot may not always reveal information, such as when the teammate already knows the information (through explicit modeling of teammate beliefs), or when the benefit to teammate performance is not worth the cost of information-revealing. Second, in situations where information is not immediately useful, it can be revealed any time before it becomes critical for selecting a cost-efficient motion. Compared to legible motion, which encourages early intent revealing, we aim to make succinct communication that is more cost efficient yet equally effective through the corroborative act formulation.

Therefore, from the timing perspective, when information revealing is critical for teammate performance right away, e.g. human teammate is waiting for robot signal of which cup to empty, corroborative act shall perform similarly to legible motion. When information marginally affects the optimal value at current state, however, corroborative act shall generate efficient motion for team task completion without salient information revealing. When information is critical to teammate but not until in later states, corroborative act may delay revealing, in case early revealing is less efficient.

Finally, in legible motion, λ has significant effect on robot behavior, which requires careful tuning and domain expertise. In corroborative acts, such tuning is built-in; adequate motion is generated, because exaggeration is cost-inefficient to teamwork performance and non-salient signal is ineffective for teammate inference.

5 VALIDATION

Here we choose group navigation as the domain for validation and we detail the belief propagation, the reward paramatrization, and the planning techniques to implement smooth group coordination; but the approach can be generalized to other domain specifications.

Agents traveling together is widely studied in group navigation [14, 20, 22]; when walking together, agents need to adapt group shape/formulation to adjust to environment changes, e.g. to form a line when passing through narrow corridor, or to avoid to the side when partner's path is occluded. They therefore need the ability to plan for one another to efficiently reach the subgoal, which follows the setting of collaborative tasks under shared task specification. In group navigation, agents often do not coordinate a priori where they are going, but decide on-the-fly based on real-time observations. Communication either explicitly or implicitly about local decisions (on subtask specifications), is therefore essential to smooth teamwork. Here we choose the guidance task to demonstrate the effectiveness of our approach; this setting applies to service robots for helping pedestrians unfamiliar with the environment, e.g. visitors in museums or travellers in airports, in which the human is (known to be) uncertain of the environment specification and the robot knows such information.

5.1 Domain: Joint Navigation

We here detail the domain design and the implementation for agents traveling in groups, including the subgoal inference model for group followers $p(b_t^{-r}|a_{t-1}^{-r}, \hat{o}_t^{-r}, b_{t-1}^{-r})$, and the reward function C for smooth coordination among travellers in groups. We then introduce teammate behavior under uncertainty using the first-layer inference policy: $\pi^{-r,1}(is_t^{-r})$, and the intuition behind the interaction with the robot using second-layer inference policy: $\pi^{r,2}(is_t^{r})$.

5.1.1 Domain Design. We consider a corridor domain, where potential goals are aligned along with the walking direction. We choose this domain for the following two reasons, which help to distinguish our approach from the legible motion formulation: (i) here, goal information has no critical influence on follower optimal action in Eq. 4 until the timing is one-step/horizon away from the first entrance; no matter which entrance is the final goal, the optimal action, until destination is reached, is to walk straight along the corridor. As a result, goal information is not critical to teammate performance until a later timing in the planning horizon, which contradicts the assumption in legible motion. (ii) information revealing is less costly later along a trajectory, as small deviation from the first entrance serves as a salient signal to infer the second entrance as the right goal. As a result, early revealing in legible motion appears costly compared to the succinct motion using our approach.

5.1.2 Belief Propagation and State Transition. We assume agent $-r$ has direct observability to the position and velocity of agent r, if r is within the visible space, defined as $[\frac{-\pi}{2}, \frac{\pi}{2}]$ from its walking direction, to sample $\hat{o}_t^{-r}$ for belief propagation. The simulated belief propagation function of agent $-r$, $p(\theta_t^r|\theta_{t-1}^r, a_{t-1}^{-r}, \hat{o}_t^{-r})$, is then affected by a_{t-1}^r through $\hat{o}_t^{-r}$. During the planning process, at each time t, given past action a_{t-1}^r, the robot simulates belief propagation to acquire $b_t^r(b_t^{-r})$. The sampled b_t^{-r} is then used to acquire teammate action a_t^{-r} by Eq. 4, to forward transit to s_{t+1}, along with the choice of robot action a_t^r.

5.1.3 Reward/Cost Function: C. To plan as a competent traveller and group member, we evaluate action sequences based on: group travel efficiency (motion duration, weighted by 1), desired group configuration regulation (weighted by 5 in quadratic form), human walking pace regulation (weighted by 5 in quadratic form), and head orientation regulation. We apply side-by-side walking as the desired group configuration, and 0.7 m/s as the desired human walking pace.

5.1.4 Search and Action Space. We apply tree search to finite-horizon T to solve the optimization with discretized dynamics/state-transition. Based on the observation that people tended to have a response time of 3 *sec* before intersections or potential collision, we apply the lookahead to 4 *sec*, and use the admissible Euclidean-distance heuristic for future value estimate V_T. We sample smooth actions by applying constant speed change and angular velocity of the range [-1,1] m/s^2 and [-22.5,22.5] deg/s. We choose 1 *sec* as the duration of 1 planning horizon, in account for that humans tend to react 1 *sec* after environmental changes.

5.1.5 $\pi^{-r,1}$ Performance. Due to the good expected value in front of the first entrance, the simulated follower using $\pi^{-r,1}$ stops there under an incorrect prior, which delays the group travel. Group followers hesitating in front of decision points are commonly observed in the real world [23]. The time delay then serves as a motivation and quantitative measure in assessment to information-revealing motions; and our approach balances this cost by incorporating $\pi^{-r,1}$ into the prediction step of our planning approach.

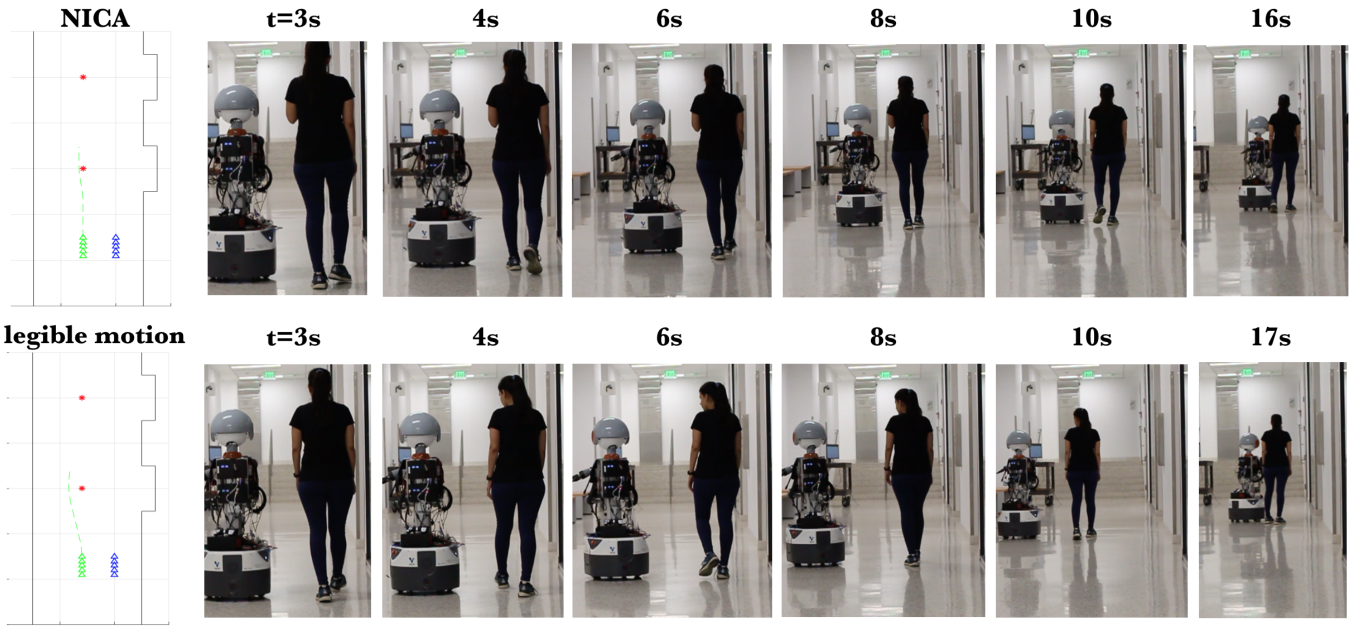

Figure 2: The trajectories of legible motion (bottom) and *NICA* (top) in simulated domains (left) and during human study (right). Legible motion generates a curved trajectory away from the first entrance starting $t = 4s$, whereas *NICA* produces a short-duration deviation when close to the first entrance at $t = 6s$.

To benchmark the performance comparison, we solve for legible motion using the same discrete setting, using accumulated reward instead of integral over time. For discounted weighting, we apply a discount factor $\gamma = 0.9$ over time horizon, in replacement of $f(t)$. We apply $\lambda = 0.05$, with which the robot curves its trajectory smoothly to bring the group to the center of the corridor but not all the way to the other side.

5.2 Experimental Design

We conducted a human study in a guidance task, where the human participants were instructed to reach an unknown target destination following a robot's guidance. For each participant, a training trial was run to familiarize the participant with the task and walking with the robot, to eliminate variance introduced by the novelty effect. Two conditions were then run within-subjects and we counterbalanced the conditions to prevent order effects: one implemented our approach and the other implemented legible motion [7] for trajectory generation. We collected participants' subjective evaluation of the robot's performance at the end of each experiment. The test took approximately 20-30 minutes, and a gift was delivered afterwards as compensation for their time.

We hypothesize the following, compared with legible motion:

(1) *NICA* generates motion that has high ratings on clarity of intent
(2) *NICA* generates motion is perceived as more natural and predictable
(3) *NICA* leads to a more comfortable experience during the collaboration
(4) *NICA* leads to higher rating of the robot being safe, intelligent, capable, thoughtful, and fluent to team with

5.3 Independent Variables

The trajectories run by the robot in the experiments can be seen in Fig. 2. The leftmost figures demonstrate the trajectories generated in the simulated corridor environment, by *NICA* (upper) and legible motion (lower). The agent on the left (marked in bright green) is the simulated robot, and that on the right (marked in blue) is the simulated pedestrian, who follows the robot.

With legible motions, the robot deviates its route to the left early, due to the encouragement in early information revealing, and makes a curve that moves away from the first entrance to indicate goal location. The robot then curves back and turns towards the final destination once the path deviation is sufficient to distinguish between the two velocity directions towards two goals. With *NICA*, the robot stays straight and does not deviate its route until it is *close* to the first entrance; this is efficient from the joint efficiency perspective, since early revealing has no additional benefit and small deviation when close to the goal serves sufficiently as a salient signal to not go to the first entrance. Due to the multi-agent planning formulation, the robot deviates to the other side of the human to prevent potential path occlusion and pressure against the wall.

The real-time performance while walking with a participant is shown in Fig. 2; we apply human tracking using a 2D laser scanner [15] to online detect potential collision. The overall experience took about 16 s to reach the destination. While legible motion started deviating to the left starting $t = 4s$, *NICA* did not do so until $t = 6s$. Due to the larger tracking error contributed by the

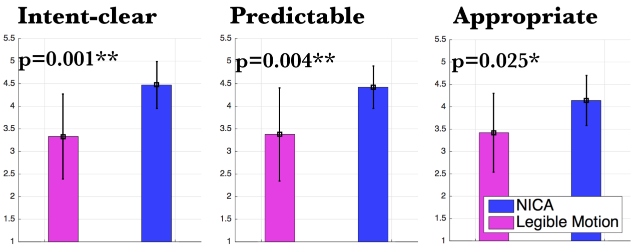

Figure 3: The statistical analysis on subjective measures, combining subscales. The bar values are the mean (M) and error bars are the standard deviation (SD). *NICA* significantly outperforms in perceptions of being socially appropriate, predictable and intent-clear.

	Legible Motion	NICA
Natural (p=0.028*)	2.92 (0.79)	**3.82 (0.83)**
Fluent to team with (p=0.05')	3.33 (0.78)	**4.00 (0.71)**
Efficient (p=0.09)	3.40 (1.17)	4.22 (0.67)
Capable of guidance (p=0.25)	4.08 (0.90)	4.54 (0.88)
Intelligent (p=0.63)	3.58 (0.63)	3.85 (0.80)
Safe (p=0.79)	4.17 (0.83)	4.33 (0.49)
Thoughtful (p=0.84)	3.70 (1.16)	3.90 (0.99)
Future interaction (p=0.96)	4.25 (0.87)	4.38 (0.77)
Trust (p=0.98)	3.83 (0.72)	3.92 (0.79)
Confident in its ability (p=1.00)	4.25 (0.62)	4.25 (0.62)

Table 1: Responses on individual questions, reported as "[M] ([SD])": *NICA* was perceived as more natural (significant), more fluent to team with (marginal significant), and had equal or better performance in the rest of the items. Listed results are ordered based on p value.

large orientation change, the turning-back of legible motion seemed more delayed, which resulted in a larger curved motion and slight delay in arrival time ($t = 17s$).

5.4 Measures

We conducted survey with 18 questions on robot (i) motion legibility and predictability based on Dragan's metrics [6] (1-5), (ii) trust and fluency of teamwork based on Hoffman's metrics for fluency in human-robot collaborations [10] (6-8), (iii) capability as a teammate for the guidance task, including: 9. the robot was thoughtful about how it guided me to the destination, 10. the robot was intelligent, 11. the robot was efficient, 12. I feel safe walking next to the robot, 13. the robot is capable of directing me, 14. I am confident in the robot's ability to help me and (vi) traits as a social agent, including: 15. the robot behaved naturally, 16. the robot was awkward to interact with, 17. the robot is socially appropriate, and 18. I feel uncomfortable with the robot. Responses were collected using 5-point Likert scale.

5.5 Experimental Setup and Protocol

The experiment took place in a public corridor, along which several classrooms are located. The experiment starts at a location close to the end of the corridor, where two classroom entrances were on the right as potential destination. The robot always took the participant to the first entrance during the training trial, and to the second entrance during official trials. This way, the goal location/distance is controlled among the two conditions yet participants would think that the goal is not fixed, to eliminate goal location as a confound.

The participant first read the instructions explaining a scenario in which the participant is being requested to reach a service location, with a robot sent to guide him/her to ensure timely arrival. The start location was shared among every trial, facing the end of the corridor. Neither the locations nor destination number were explicitly instructed.

In all training trials, the robot runs *NICA* and follows a straight line to the goal (first entrance) at a fixed speed, 0.7 m/s, and with a fixed distance to wall throughout the walking. The robot started slowing down 1 sec before reaching the destination. As a test of our assumptions about walking speed, participants were asked after the training trial whether the robot's speed was acceptable to them. All reported that the speed of the robot was fine to continue the experiment with. As we observed in pilot trials that the slowing-down motion also affects people's perceptions of how clear the goal is indicated, e.g. how quick the slow-down was, how far away the robot started slowing down, etc., participants were asked to focus their evaluation of the two conditions only on the walking portion of the interaction, not the final approach to the goal, which was common across both conditions, but sensitive to the quality of localization and pedestrian tracking. Because the study was within subjects and counter-balanced by order, we expect that this would not affect the validity of the study.

5.6 Data

We collected 16 participants in the human study, with 8 females and 8 males, who are visitors or students on campus. The experiment used a within-subjects design to enable participants to compare the two motions. The order of the conditions was counterbalanced to control for order effects. Among the experiments, there were occasional jittery motions, due to tracking delay from network communication issues. We filtered out 3 participants' entire data which had experienced the jittery behavior in more than one trial (including the training trial); we filtered out 1 additional response to legible motion due to observing jittery behavior during the test, in which legible motion went second.

6 RESULTS

The survey responses can be seen in Fig. 3, with responses to legible motions on the left side in bright magenta, and those to our planner on the right in dark blue. We conducted t-test to on combined subscales of legibility (question 1-3), predictability (4-5), and social appropriateness (question 16-18), with Cronbach's $\alpha > 0.7$. The results with significance ($p < 0.05$) are highlighted with $*$, and those with strong significance ($p < 0.01$) are highlighted with with $**$. For the rest of the individual questions, we conducted Wilcoxon signed-rank test and report p value for significance test, shown in Table. 1. Using t-test on combined subscales, *NICA* significantly outperforms the legible motion by being: more *socially appropriate* ($p = 0.027^*$), supporting our hypothesis (3) on interaction comfort, and more *intent-clear* ($p = 0.001^{**}$), which does not support our hypothesis (1) yet with better results. *NICA* is also perceived as more *predictable* ($p = 0.004^{**}$) using t-test on combined subscales and more *natural* ($p = 0.028^*$) using Wilcoxon signed-rank test, which supports our hypothesis (2). For hypothesis (4), most items are not supported yet *NICA* was rated more fluent to team with with marginal significance ($p = 0.05'$).

Among the scales and questions with significance using t-test and Wilcoxon signed-rank test: for traits of a social agent, including being socially appropriate (M=4.14, SD=0.56) and natural (M=3.82, SD=0.83) make the robot more suitable for long-duration interaction; being more fluent to team with (M=4.00, SD=0.71) makes the robot more suitable to engage teamwork with humans. While other traits in robot capabilities and interaction qualities had no significant trends, with results shown in Table. 1, ratings for robot being safe, capable, and future interaction were on average ranked above 4.0 (Somewhat Agree) for both planners; ratings for robot being intelligent and thoughtful, and trust-worthy were on average above 3.5 for both planners. Across all 18 self-report outcomes, *NICA* had equal or better performance compared to legible motions.

In navigation, humans often face towards their walking direction to maintain the ability to observe to the upcoming surroundings despite conversation and distraction. Efficient paths have also been suggested as be more intent-clear in the navigation domain previously [16]. During our legible motion experiments, we observed that participants constantly turned their heads and looked at the robot when it altered its direction and deviated away to the side . Many participants followed the robot all the way from the side to the middle of the corridor, as shown in Fig. 4. We assume these behaviors can be attributed to confusion or uncertainty.

7 DISCUSSION

7.0.1 Descriptive Responses to Robot Performance. After the survey questions for quantitative measure for each planner, we discussed robot performance with some participants, and collected their responses to: *fluency in teamwork*, *social appropriateness*, and *preferences for future interaction*. Overall, we found that participants often mentioned *distance* in their descriptions of social appropriateness. Some preferred constant distance to the wall during walking and referred to it as being more natural or like walking with a human. Some participants did not perceive the planners as either appropriate or inappropriate. One indicated that there was few social interaction, like verbal communication, and was not concerned if the robot was inappropriate. One did not find any differences among the planners or did only notice distance changes.

Goal indication was often described as a bonus for fluent teaming and future interaction. One described *NICA* as being "fluent and more predictable". For *future interaction*, some preferred legible motion because of its very clear intention. Some preferred *NICA* because it had some goal indication while maintaining constant distance to the wall.

7.0.2 Surrounding Conditions. While we controlled the workspace to be cleared during the experiment, e.g. we waited until pedestrians had passed, other pedestrians sharing the workspace have a strong influence on how participants perceived the robot's motion. One participant experienced interruption when the robot reached the goal and slowed down; he asked if the robot was supposed to do that when a person approached from behind. Two participants described legible motion as it can "weave through the crowd" and perceived it as more aware of the surroundings.

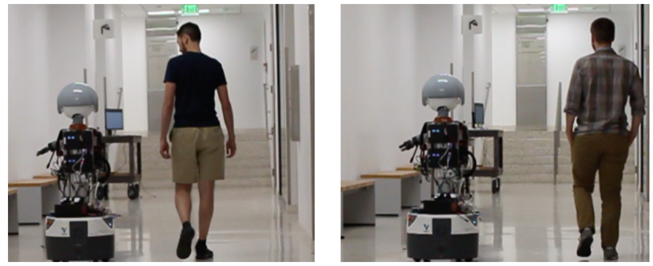

Figure 4: Shots of participants at the same time during the legible motion experiment: we often observed participants following the robot to the middle of the road (Left), while some stayed on their original routes or in between (Right).

7.0.3 Other Responses and Dominant Factors. While we reminded the participants to focus on the walking experience, many responded on the slowing-down behavior when approaching the destination. One suggested a red-glowing functionality when the robot is braking, given the human-following scenario. We recognize the implementation of goal-approaching behavior can have great influences on human's perception of robot performance.

While distance and motion smoothness appeared to be dominant factors of human perceptions of the experience, trajectories were sometimes off-route due to occasionally larger localization errors, contributed by the feature/landmark-poor long corridor with glass-wall reflections. We recognize this as the major source of noise for experiment condition, other than the occasional jittery motion due to network buffer issues.

8 CONCLUSION AND FUTURE WORK

In this work, we proposed that efficient information revealing is critical to human human-robot teamwork performance, and robot teammates need the ability to reason about human teammate uncertainty to enable efficient communication. To achieve this ability, we proposed nested inference modeling and incorporate it into a multi-agent planning formulation, to reveal information based on benifits to teamwork, which is concerned with both self efficiency and teammate performance improvement from the information-revealing process. We refer to the succinct actions in information revealing as corroborative acts, or *NICA* as the planner for motion generation, and validate the approach with human study on a robot guidance task. The results showed that *NICA* is perceived as significantly more natural, socially appropriate, and fluent to team with, while being both more predictable and intent-clear compared to the legible motion formulation. It was also suggested that *NICA* was perceived as more fluent to team with with marginal significance. *NICA* had equal or better performance across all 18 self-report outcomes.

While we implemented corroborative acts in motion planning in the navigation domain, the planning formulation can be generalized to other domains with more complex domain specifications and task structure. Other forms of communicative actions can also be considered. We also look forward to applying this approach to dynamic environments, where modeling of partial observability and communication become critical, to efficiently online coordinate teammates to adapt to unexpected changes.

REFERENCES

[1] Haoyu Bai, Shaojun Cai, Nan Ye, David Hsu, and Wee Sun Lee. 2015. Intention-aware online POMDP planning for autonomous driving in a crowd. In *2015 ieee international conference on robotics and automation (icra)*. IEEE, 454–460.

[2] Cynthia Breazeal, Cory D Kidd, Andrea Lockerd Thomaz, Guy Hoffman, and Matt Berlin. 2005. Effects of nonverbal communication on efficiency and robustness in human-robot teamwork. In *2005 IEEE/RSJ international conference on intelligent robots and systems*. IEEE, 708–713.

[3] Abhizna Butchibabu, Christopher Sparano-Huiban, Liz Sonenberg, and Julie Shah. 2016. Implicit coordination strategies for effective team communication. *Human factors* 58, 4 (2016), 595–610.

[4] Mai Lee Chang, Reymundo A Gutierrez, Priyanka Khante, Elaine Schaertl Short, and Andrea Lockerd Thomaz. 2018. Effects of integrated intent recognition and communication on human-robot collaboration. In *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 3381–3386.

[5] Robin Deits, Stefanie Tellex, Pratiksha Thaker, Dimitar Simeonov, Thomas Kollar, and Nicholas Roy. 2013. Clarifying commands with information-theoretic human-robot dialog. *Journal of Human-Robot Interaction* 2, 2 (2013), 58–79.

[6] Anca D Dragan, Shira Bauman, Jodi Forlizzi, and Siddhartha S Srinivasa. 2015. Effects of robot motion on human-robot collaboration. In *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*. ACM, 51–58.

[7] Anca D Dragan, Kenton CT Lee, and Siddhartha S Srinivasa. 2013. Legibility and predictability of robot motion. In *Human-Robot Interaction (HRI), 2013 8th ACM/IEEE International Conference on*. IEEE, 301–308.

[8] Michael J Gielniak and Andrea L Thomaz. 2011. Generating anticipation in robot motion. In *2011 RO-MAN*. IEEE, 449–454.

[9] Dirk Helbing and Peter Molnar. 1995. Social force model for pedestrian dynamics. *Physical review E* 51, 5 (1995), 4282.

[10] Guy Hoffman. 2019. Evaluating fluency in human–robot collaboration. *IEEE Transactions on Human-Machine Systems* 49, 3 (2019), 209–218.

[11] Guy Hoffman and Cynthia Breazeal. 2007. Effects of anticipatory action on human-robot teamwork efficiency, fluency, and perception of team. In *Proceedings of the ACM/IEEE international conference on Human-robot interaction*. ACM, 1–8.

[12] Ross A Knepper, Christoforos I Mavrogiannis, Julia Proft, and Claire Liang. 2017. Implicit communication in a joint action. In *Proceedings of the 2017 acm/ieee international conference on human-robot interaction*. ACM, 283–292.

[13] Emiel Krahmer, Marc Swerts, Mariet Theune, and Mieke Weegels. 2001. Error detection in spoken human-machine interaction. *International journal of speech technology* 4, 1 (2001), 19–30.

[14] Markus Kuderer and Wolfram Burgard. 2014. An approach to socially compliant leader following for mobile robots. In *International Conference on Social Robotics*. Springer, 239–248.

[15] Angus Leigh, Joelle Pineau, Nicolas Olmedo, and Hong Zhang. 2015. Person tracking and following with 2d laser scanners. In *Robotics and Automation (ICRA), 2015 IEEE International Conference on*. IEEE, 726–733.

[16] Christina Lichtenthäler, Tamara Lorenzy, and Alexandra Kirsch. 2012. Influence of legibility on perceived safety in a virtual human-robot path crossing task. In *RO-MAN, 2012 IEEE*. IEEE, 676–681.

[17] Felix Lindner. 2015. A conceptual model of personal space for human-aware robot activity placement. In *Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on*. IEEE, 5770–5775.

[18] Christoforos I Mavrogiannis, Valts Blukis, and Ross A Knepper. 2017. Socially competent navigation planning by deep learning of multi-agent path topologies. In *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 6817–6824.

[19] Marek P Michalowski, Selma Sabanovic, and Reid Simmons. 2006. A spatial model of engagement for a social robot. In *Advanced Motion Control, 2006. 9th IEEE International Workshop on*. IEEE, 762–767.

[20] Luis Yoichi Morales Saiki, Satoru Satake, Rajibul Huq, Dylan Glas, Takayuki Kanda, and Norihiro Hagita. 2012. How do people walk side-by-side?: using a computational model of human behavior for a social robot. In *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction*. ACM, 301–308.

[21] Mehdi Moussaïd, Niriaska Perozo, Simon Garnier, Dirk Helbing, and Guy Theraulaz. 2010. The walking behaviour of pedestrian social groups and its impact on crowd dynamics. *PloS one* 5, 4 (2010), e10047.

[22] Ryo Murakami, Luis Yoichi Morales Saiki, Satoru Satake, Takayuki Kanda, and Hiroshi Ishiguro. 2014. Destination unknown: walking side-by-side without knowing the goal. In *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction*. ACM, 471–478.

[23] Information omitted for blind review. [n.d.].

[24] Nicholas Roy, Wolfram Burgard, Dieter Fox, and Sebastian Thrun. 1999. Coastal navigation-mobile robot navigation with uncertainty in dynamic environments. In *Proceedings 1999 IEEE International Conference on Robotics and Automation (Cat. No. 99CH36288C)*, Vol. 1. IEEE, 35–40.

[25] Satoru Satake, Takayuki Kanda, Dylan F Glas, Michita Imai, Hiroshi Ishiguro, and Norihiro Hagita. 2009. How to approach humans?: strategies for social robots to initiate interaction. In *Proceedings of the 4th ACM/IEEE international conference on Human robot interaction*. ACM, 109–116.

[26] Volkan Sezer, Tirthankar Bandyopadhyay, Daniela Rus, Emilio Frazzoli, and David Hsu. 2015. Towards autonomous navigation of unsignalized intersections under uncertainty of human driver intent. In *Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on*. IEEE, 3578–3585.

[27] Julie Shah and Cynthia Breazeal. 2010. An empirical analysis of team coordination behaviors and action planning with application to human–robot teaming. *Human factors* 52, 2 (2010), 234–245.

[28] Chao Shi, Michihiro Shimada, Takayuki Kanda, Hiroshi Ishiguro, and Norihiro Hagita. 2011. Spatial formation model for initiating conversation. *Proceedings of robotics: Science and systems VII* (2011), 305–313.

[29] Elena Torta, Raymond H Cuijpers, James F Juola, and David Van Der Pol. 2012. Modeling and testing proxemic behavior for humanoid robots. *International Journal of Humanoid Robotics* 9, 04 (2012), 1250028.

[30] Araceli Vega, Luis J Manso, Douglas G Macharet, Pablo Bustos, and Pedro Núñez. 2018. Socially aware robot navigation system in human-populated and interactive environments based on an adaptive spatial density function and space affordances. *Pattern Recognition Letters* (2018).

[31] Michael L Walters, Kerstin Dautenhahn, René Te Boekhorst, Kheng Lee Koay, Christina Kaouri, Sarah Woods, Chrystopher Nehaniv, David Lee, and Iain Werry. 2005. The influence of subjects' personality traits on personal spatial zones in a human-robot interaction experiment. In *Robot and Human Interactive Communication, 2005. ROMAN 2005. IEEE International Workshop on*. IEEE, 347–352.