Implicit Communication in Human-Robot Collaborative Transport

Elvin Yang
University of Michigan
Ann Arbor, Michigan, USA
eyy@umich.edu

Christoforos Mavrogiannis
University of Michigan
Ann Arbor, Michigan, USA
cmavro@umich.edu

Abstract

We focus on human-robot collaborative transport, in which a robot and a user collaboratively move an object to a goal pose. In the absence of explicit communication, this problem is challenging because it demands tight *implicit* coordination between two heterogeneous agents, who have very different sensing, actuation, and reasoning capabilities. Our key insight is that the two agents can coordinate fluently by encoding subtle, communicative signals into actions that affect the state of the transported object. To this end, we design an inference mechanism that probabilistically maps observations of joint actions executed by the two agents to a set of joint strategies of workspace traversal. Based on this mechanism, we define a cost representing the human's uncertainty over the unfolding traversal strategy and introduce it into a model predictive controller that balances between uncertainty minimization and efficiency maximization. We deploy our framework on a mobile manipulator (Hello Robot Stretch) and evaluate it in a within-subjects lab study ($N = 24$). We show that our framework empowers the robot to be perceived as a significantly more fluent and competent partner compared to baselines lacking a communicative mechanism.

CCS Concepts

• **Computer systems organization → Robotic autonomy**; • **Human-centered computing → Human computer interaction (HCI)**; • **Computing methodologies → Cooperation and coordination**.

Keywords

Human-robot collaboration, Human-robot teams, Implicit communication

ACM Reference Format:
Elvin Yang and Christoforos Mavrogiannis. 2025. Implicit Communication in Human-Robot Collaborative Transport. In *Proceedings of HRI '25 (Third Workshop on Explainability in Human-Robot Collaboration)*. ACM, New York, NY, USA, 5 pages.

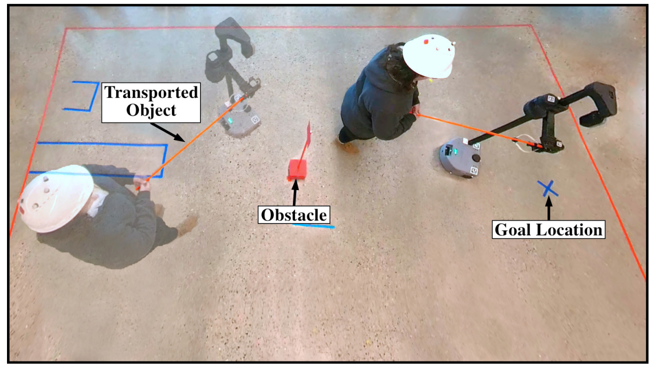

Figure 1: Footage from our study ($N = 24$) involving the collaborative transport of an object (orange stick) by a user and a mobile manipulator in a workspace with an obstacle (red color). The robot runs our controller (IC-MPC), designed to balance functional and communicative actions in collaborative tasks.

1 Introduction

Recently, there has been vivid interest in developing physically capable robot partners that could assist humans in *context*-rich, dynamic and unstructured domains [26] like homes [28, 31] and manufacturing sites [16]. An important task in this space involves the *collaborative transport* of objects that might be too large or too heavy to be transported by a single agent. This task is especially challenging as it not only requires efficient and fluent coordination between the two heterogeneous partners but also the simultaneous satisfaction of geometric, kinematic, and physics constraints.

Humans often tackle physically demanding collaborative tasks like transport by fluently coordinating their physical movements with their partners [25] even without a concrete plan, with minimal explicit coordination. This capability relies on sophisticated mechanisms connecting perception and action. A prevalent theory from action understanding, commonly referred to as the "teleological stance", highlights that agents' actions can often be explained by an underlying goal [1, 5, 7]. This idea has inspired researchers in human-robot interaction (HRI) to develop mechanisms that communicate a robot's intended goal to an observer through its actions [6, 11]. These mechanisms have produced intent-expressive robot behavior in manipulation [6], autonomous driving [24], and social robot navigation [17]. Likewise, we view communication—especially implicit communication [11], the ability to infer and convey information within physical actions—to be a critical skill of robots working in close physical collaboration with humans. Implicit communication can be low-latency, robust to environmental disturbances (e.g., noise, poor lighting conditions), and require less attention compared to explicit forms. While explicit communication remains highly relevant to team activities, implicit communication serves as an important complement that supports fluent teamwork.

To investigate the implications of implicit communication for physical human-robot teamwork, we instantiate a task of human-robot collaborative transport, where the goal of the human-robot team is to collaboratively move an object to a goal pose while avoiding collisions with static obstacles (see Fig. 1). In this task, the user is simultaneously an observer of the robot and a *dynamic actor*, persistently influencing and being influenced by the robot while it *physically collaborates* with them. While prior work in human-robot collaborative transport has emphasized fixed leadership roles for the two agents [2, 3, 15, 19, 22, 27], we consider a *dynamic negotiation* over a joint strategy of *workspace traversal*. We contribute a control framework that leverages *implicit communication* [11] through actions influencing the state of the transported object to enable the robot to negotiate an efficient traversal with its human partner. We move beyond past work on implicit communication, where the user is either *not an actor* [6] or *not physically* collaborating with their robot partner [14, 17]. We demonstrate our framework on a mobile manipulator and evaluate it in a lab study ($N = 24$) involving the collaborative transport of an object in a workspace with an obstacle obstruction. In our full paper [30], we perform a full analysis to show that our framework outperforms baselines lacking a communicative mechanism in terms of task completion and human impressions. Additionally, we include videos from the study (https://youtu.be/0NTDrobSifg) and code and data that could help the community iterate on our work (https://github.com/fluentrobotics/icmpc_collab_transport).

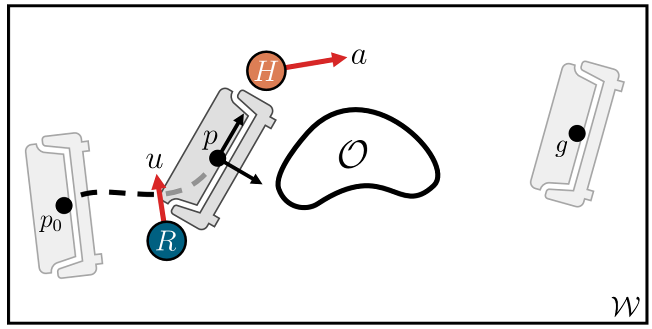

Figure 2: A human (H) and a robot (R) collaboratively move an object from an initial pose p_0 to a final pose g in a workspace $\mathcal{W}$. An obstacle O stands in their way. To avoid collisions with O and reach g, they have to coordinate on a strategy of workspace traversal. In this work, we engineer implicit coordination through the velocities a and u that the human and the robot exert on the object.

2 Problem Statement

We consider a human H and a robot R collaboratively transporting an object. The robot and the human grasp the object at a fixed height; this allows us to instantiate the problem on a planar workspace $\mathcal{W} \subseteq SE(2)$. Assuming a quasistatic setting, the object's state $p \in \mathcal{W}$ evolves according to $p_{k+1} = f(p_k, a_k, u_k)$, where $a \in \mathcal{A}$, $u \in \mathcal{U}$ represent human and robot velocities, respectively, and k is a time index. The workspace includes a set of obstacle-occupied regions $O \subset \mathcal{W}$. The goal of the human-robot team is to transport the object from an initial pose p_0 to a desired pose g in $\mathcal{W}$ (see Fig. 2) while avoiding collisions with O. We assume that the two agents do not communicate explicitly (e.g., via language), but they observe the actions of one another. Our goal is to design a control policy to enable the robot to efficiently and fluently collaborate with its human partner.

3 Balancing Functional and Communicative Actions in Human-Robot Collaborative Transport

3.1 Formalizing Joint Strategies of Workspace Traversal

Collaborative tasks involving multiple agents working together require consensus on a *joint strategy* ψ, i.e., a qualitatively distinct way of completing the task, out of the set of all possible joint strategies, Ψ. Often, this consensus is not established *a priori*; rather, it is dynamically negotiated during execution. The abstraction of a joint strategy effectively captures critical domain knowledge at a representation level. While prior work on collaborative transport has emphasized *role* assignment across the team (i.e., whether the robot or the human are leading or following each other) [10, 20, 23], realistic, obstacle-cluttered environments present additional important challenges, such as the decision over *how to pass through* an obstacle-cluttered workspace.

In this work, we formalize the space of workspace traversal strategies using tools from homotopy theory [12]. The human-robot team is tasked with transporting an object from its initial pose p_0 to a final pose g, resulting in an object trajectory $\boldsymbol{p} : [0, 1] \to \mathcal{W}$, where $\boldsymbol{p}(0) = p_0$ and $\boldsymbol{p}(1) = g$, belonging to an appropriate space of trajectories $\mathcal{P}$. Obstacles, defined as the connected components of O, naturally partition $\mathcal{P}$ into equivalence classes Ψ, where each $\psi \in \Psi$ represents a distinct workspace traversal strategy under which the transported object can travel from p_0 to g, i.e.,

$$\begin{aligned} &\mathcal{P} = \bigcup_{\psi \in \Psi} \psi \\ &\forall \psi^i, \psi^j \in \Psi : (\psi^i \cap \psi^j \neq \emptyset) \implies (\psi^i = \psi^j) . \\ &\forall \boldsymbol{p}^i, \boldsymbol{p}^j \in \psi : \boldsymbol{p}^i \sim \boldsymbol{p}^j \end{aligned} \tag{1}$$

These classes can be identified using a notion of topological invariance. The works of Kretzschmar et al. [13], Mavrogiannis et al. [18], Vernaza et al. [29] use winding numbers to describe topological relationships between the robot and obstacles or humans navigating around it. Here, we adapt this idea to collaborative transport by enumerating the set of homotopy classes between the object trajectory and obstacles in the workspace. Specifically, for any object trajectory $\boldsymbol{p}$ embedded in a space with m obstacles $o_1, \ldots, o_m$, we can define winding numbers

$$w_i = \frac{1}{2\pi} \sum_t \Delta\theta_t^i, \quad i = 1, \ldots, m, \tag{2}$$

where $\Delta\theta_t^i = \angle(p_t - o_i, p_{t-1} - o_i)$ denotes an angular displacement corresponding to the transfer of the object from p_{t-1} to p_t. The sign of w_i represents the passing side between the object and the i-th obstacle, and its absolute value represents the number of times the object encircled the i-th obstacle. For a trajectory $\boldsymbol{p}$, the tuple

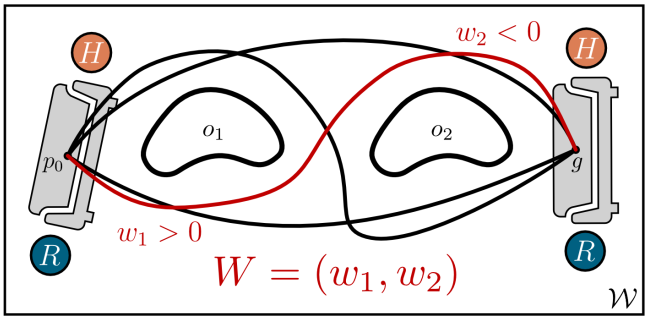

Figure 3: Illustration of our topological abstraction for representing strategies of workspace traversal. Representing workspace traversal strategies as tuples of winding number signs, W. In this scene with two obstacles, there are four possible strategies represented as continuous curves. The red curve highlights a strategy corresponding to passing on the right of o_1 ($w_1 > 0$), and the left of o_2 ($w_2 < 0$). This representation is applicable to any number of obstacles.

of winding number signs

$$W = (\operatorname{sign} w_1, \ldots, \operatorname{sign} w_m) \tag{3}$$

represents an equivalence class describing how the human-robot team transported the object past all obstacles in the environment. In this work, we model the space of joint strategies Ψ as set of distinct W, i.e., $|\Psi| = 2^m$.

3.2 Inferring Strategies of Workspace Traversal

We describe an inference mechanism that maps observations of team actions to a belief over a workspace traversal strategy. This mechanism is agnostic to the specific definition of the strategy. At time t, we assume that the robot observes the joint action $\alpha = (a, u)$, object state p, and task context $c = (g, O)$. Given α, p, and c, our goal is to infer the unfolding workspace traversal strategy, ψ, i.e.,

$$\mathbb{P}(\psi \mid \alpha, p, c). \tag{4}$$

Using Bayes' rule, we can expand (4) as

$$\mathbb{P}(\psi \mid \alpha, p, c) = \frac{1}{\eta}\mathbb{P}(\alpha \mid \psi, p, c)\,\mathbb{P}(\psi \mid p, c), \tag{5}$$

where the left-hand side expression is the *posterior distribution* of the joint strategy ψ, and on the right-hand side, η is a normalizer across α, $\mathbb{P}(\alpha \mid \psi, p, c)$ is the *joint action likelihood distribution* and $\mathbb{P}(\psi \mid p, c)$ is a *prior distribution* of the joint strategy before observing the joint action. We can rewrite the joint action likelihood distribution as

$$\mathbb{P}(\alpha \mid \psi, p, c) = \mathbb{P}(a \mid \psi, p, c)\,\mathbb{P}(u \mid \psi, p, c), \tag{6}$$

since the two agents choose their actions independently.

The distribution of (4) allows the robot to represent the belief of its partner over the unfolding traversal strategy. A natural measure of uncertainty over the observer's belief regarding that strategy can be acquired by computing the information entropy of ψ, conditioned on known α, p, c:

$$H(\psi \mid \alpha, p, c) = -\sum_{\psi \in \Psi} \mathbb{P}(\psi \mid \alpha, p, c) \log \mathbb{P}(\psi \mid \alpha, p, c). \tag{7}$$

Intuitively, the higher H is, the higher the uncertainty of the user over the unfolding ψ is assumed to be.

3.3 Integrating Human Inferences into Robot Control

We integrate the inference mechanism of (4) into a model predictive control (MPC) algorithm by using its entropy (7) as a cost. Given the context $c = (g, O)$ and the object state p at time t, the goal of the MPC is to find the sequence of future robot actions $\boldsymbol{u}^*$ that minimizes a cost function J over a horizon T. At every control cycle, the MPC solves the following planning problem:

$$\begin{aligned} (u_{t:t+T})^* = \arg\min_{u_{t:t+T}} &\; J(p_{t:t+T}, u_{t:t+T}) \\ s.t.\; p_{k+1} &= f(p_k, a_k, u_k), \\ a_k &\in \mathcal{A} \\ u_k &\in \mathcal{U} \end{aligned} \tag{8}$$

We split J into a running cost J_k and a terminal cost J_T

$$\begin{aligned} J(p_{t:t+T}, u_{t:t+T}) = &\sum_{k=0}^{T} \gamma^k J_k(p_{t+k}, u_{t+k}) \\ &+ J_T(p_{t+T}, u_{t+T}) \end{aligned}, \tag{9}$$

where γ is a discount factor, and the terminal cost penalizes distance from the object's goal pose g:

$$J_T(p_{t+k}, u_{t+k}) = ||p_{t+k} - g||^2. \tag{10}$$

The running cost J_k is a weighted sum of two terms, i.e.,

$$\begin{aligned} J_k(p_{t+k}, u_{t+k}) = &\, w_{obs} J_{obs}(p_{t+k}, u_{t+k}) \\ &+ w_{ent} J_{ent}(p_{t+k}, u_{t+k}), \end{aligned} \tag{11}$$

where

$$\begin{aligned} &J_{obs}(p_{t+k}, u_{t+k}) = \\ &\quad \max\left(0, -\log\left(\min_{o \in O} \frac{||p_{t+k} - o||}{\delta}\right)\right), \end{aligned} \tag{12}$$

is a collision avoidance cost penalizing proximity to obstacles, δ is a clearance threshold, J_{ent} is a cost proportional to the entropy defined in (7), and w_{obs}, w_{ent} are weights.

We refer to this control framework as *Implicit Communication MPC*, or **IC-MPC**. At every control cycle, IC-MPC plans a future robot trajectory that balances between functional objectives (collision avoidance, progress to goal) and communicative objectives (minimizing the partner's uncertainty over the upcoming joint strategy). The robot executes the first action u_t from the planned trajectory and then replans. This process is repeated in fixed control cycles until the task is completed.

4 User Study

We conducted an IRB-approved, within-subjects user study (U-M HUM00254044) in which a user collaborates with a robot to jointly transport an object to a designated pose. Each user experienced the same set of conditions, each corresponding to a collaborative algorithm running on the robot: ours, and two baselines. Through mailing lists, we recruited 24 participants (4 female, 18 male, 2 other), aged 18-29 from a university population. On average, participants

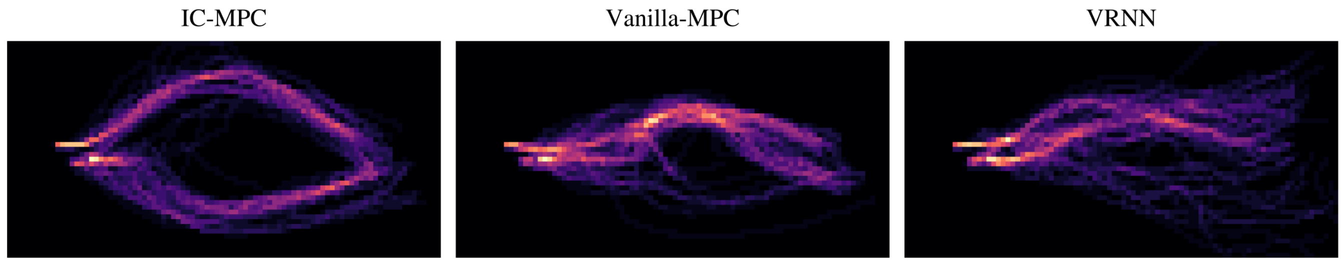

Figure 4: Spatial distribution of object trajectories within the workspace during the user study, including failure cases, itemized per algorithm. IC-MPC exhibits an almost uniform split between right and left, whereas baselines show mixed performance, including undesirable zig-zagging effects, an artifact of increased uncertainty over the unfolding traversal strategy.

Table 1: Friedman test results on subjective metrics.

Metric	$\chi^2(2)$	p	W_k
Warmth [4]	3.85	0.146	0.08
Competence [4]	18.86	$< .001$	0.39
Discomfort [4]	15.46	$< .001$	0.32
Fluency [8]	28.15	$< .001$	0.59

Table 2: Mean and standard deviation of subjective metrics. $^*p < .05$, $^{**}p < .01$, $^{***}p < .001$

Metric	IC-MPC	Vanilla-MPC	VRNN
Warmth [4] ↑	3.44 (1.89)	3.12 (1.98)	2.99 (1.82)
Competence [4] ↑	**6.06 (1.86)**	5.15 (1.82)*	4.02 (1.88)***
Discomfort [4] ↓	**2.15 (1.29)**	2.86 (1.84)*	3.22 (1.49)***
Fluency [8] ↑	**5.73 (1.02)**	4.64 (1.41)**	3.67 (1.49)***

rated of their familiarity with robotics technology as 4.1 (SD = 0.74) on a scale from 1 (not at all familiar) to 5 (very familiar). The study lasted 45 minutes, and each participant received $20. Full experimental details and analysis can be found in [30].

4.1 Experiment Design

Task Description. The user and the robot hold opposite ends of an object (a wooden stick) and transport it together from an initial pose to a goal pose. The team operates in a workspace with area 2.8×5.6 m^2. To study the coordination of the human-robot team over a discrete decision, a single static obstacle of area 0.15×0.15 m^2 is placed in the center of the workspace (see Fig. 1). Users collaborate with each algorithm three times to ensure they experience a diverse range of interactions.

Algorithms. We compare the performance of our framework (**IC-MPC**) against two baselines:

- *Vanilla-MPC*: A purely functional ablation of IC-MPC with no uncertainty-minimizing objective ($w_{ent} = 0$).
- *VRNN* [21]: A learning-based path planner based on a Variational Recurrent Neural Network that predicts the most likely future path of the object based on human demonstrations. The robot takes actions to track path predictions as closely as possible.

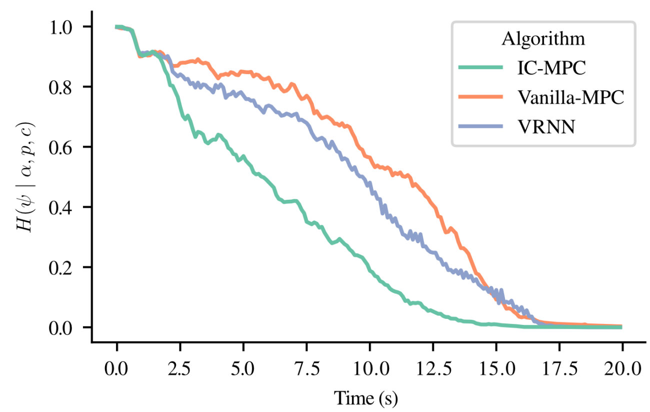

Figure 5: Entropy over the workspace traversal strategy as a proxy for strategy uncertainty, averaged across all trials for each algorithm. By directly minimizing the entropy, IC-MPC accelerates consensus on a traversal strategy. This reduces undesirable zig-zagging artifacts, present in the execution of baselines (see Fig. 4).

4.2 Discussion

We found that collected metrics did not uniformly pass the Shapiro-Wilk test of normality. Thus, for consistency, we use the non-parametric Friedman test to detect effects of the robot algorithms on dependent variables and the non-parametric paired Wilcoxon signed-rank test with Holm-Bonferroni corrections [9] for post-hoc pairwise comparison tests. Effect sizes are reported using Kendall's coefficient of concordance (denoted as W_k to disambiguate it from the test statistic of the Wilcoxon signed-rank test) and Cohen's d.

We report test statistics for subjective metrics in Table 1 and a summary of subjective metrics for each algorithm in Table 2. We found a significant effect of the robot algorithm on users' perception of *competence*, *discomfort*, and *fluency*. Post-hoc tests found that IC-MPC was judged by users as: significantly more *competent* compared to Vanilla-MPC ($W = 55, p = .021, d = 0.49$) and VRNN ($W = 26, p < .001, d = 1.09$); significantly less *discomforting* compared to Vanilla-MPC ($W = 30, p = .018, d = -0.44$) and VRNN ($W = 13.5, p < .001, d = -0.76$); a significantly more *fluent* collaborator compared to Vanilla-MPC ($W = 28, p = .002, d = 0.89$) and VRNN ($W = 11, p < .001, d = 1.61$). No statistically significant effect was found on users' perception of *warmth*.

The spatial distribution of object trajectories (Fig. 4) reveals that teams took *wider* paths around the obstacle when the robot was running IC-MPC compared to Vanilla-MPC and VRNN. Without implicit communication mechanisms to resolve ambiguity or achieve consensus on traversal strategy, both baselines would often drive straight towards the goal and attempt to pass the obstacle from directions opposite the user, leading to collisions. By acting *earlier* to take *wider* paths compared to Vanilla-MPC and VRNN, IC-MPC reduced uncertainty about the joint strategy faster than baselines (Fig. 5), thereby reducing the chance of similar collisions. As wider paths are longer than more direct paths, teams took more time on average to complete the task when the robot was running IC-MPC compared to VRNN.

Users noticed qualitative differences among algorithms. In open-ended responses, they described VRNN as "unpredictable" and "indecisive". One user described Vanilla-MPC as "a bad teammate that only does what they think is right". Two users who interacted with IC-MPC after one or both baselines whose comments were comparative in nature wrote that IC-MPC "felt more natural" and that "the collaboration on the task was a lot more seamless in this series of attempts".

At the start of each study session, we intentionally provided the vague instruction to "collaborate in whichever way feels natural". However, we received informal feedback from several participants that aspects of the interaction, including communication with the robot, *did not feel natural* or *intuitive*. Participants expressed confusion about *how* they could communicate with the robot and whether the robot was acknowledging, understanding, or ignoring what they were trying to communicate. Designing the robot to be expressive may facilitate interactions that are perceived as more natural.

References

[1] Chris L. Baker, Rebecca Saxe, and Joshua B. Tenenbaum. 2009. Action understanding as inverse planning. *Cognition* 113, 3 (2009), 329–349.

[2] Antoine Bussy, Pierre Gergondet, Abderrahmane Kheddar, François Keith, and André Crosnier. 2012. Proactive behavior of a humanoid robot in a haptic transportation task with a human partner. In *Proceedings of the IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. 962–967.

[3] Antoine Bussy, Abderrahmane Kheddar, André Crosnier, and François Keith. 2012. Human-humanoid haptic joint object transportation case study. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Sytems (IROS)*. 3633–3638.

[4] Colleen M Carpinella, Alisa B Wyman, Michael A Perez, and Steven J Stroessner. 2017. The robotic social attributes scale (RoSAS) development and validation. In *Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. 254–262.

[5] Gergely Csibra and György Gergely. 2007. 'Obsessed with goals': Functions and mechanisms of teleological interpretation of actions in humans. *Acta Psychologica* 124, 1 (2007), 60–78.

[6] Anca D. Dragan and Siddhartha Srinivasa. 2014. Integrating human observer inferences into robot motion planning. *Autonomous Robots* 37, 4 (2014), 351–368.

[7] György Gergely, Zoltán Nádasdy, Gergely Csibra, and Szilvia Bíró. 1995. Taking the intentional stance at 12 months of age. *Cognition* 56, 2 (1995), 165–193.

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

[9] Sture Holm. 1979. A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics* (1979), 65–70.

[10] Nathanael Jarrasse, Vittorio Sanguineti, and Etienne Burdet. 2014. Slaves no longer: review on role assignment for human–robot joint motor action. *Adaptive Behavior* 22, 1 (2014), 70–82.

[11] Ross A. Knepper, Christoforos Mavrogiannis, Julia Proft, and Claire Liang. 2017. Implicit Communication in a Joint Action. In *Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. 283–292.

[12] Ross A Knepper, Siddhartha S Srinivasa, and Matthew T Mason. 2012. Toward a deeper understanding of motion alternatives via an equivalence relation on local paths. *The International Journal of Robotics Research* 31, 2 (2012), 167–186.

[13] Henrik Kretzschmar, Markus Spies, Christoph Sprunk, and Wolfram Burgard. 2016. Socially compliant mobile robot navigation via inverse reinforcement learning. *The International Journal of Robotics Research* 35, 11 (2016), 1289–1307.

[14] Claire Liang, Julia Proft, Erik Andersen, and Ross A. Knepper. 2019. Implicit Communication of Actionable Information in Human-AI teams. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI)*. 1–13.

[15] Bruno G. C. Lima, Enrico Ferrentino, Pasquale Chiacchio, and Mario Vento. 2023. Assistive force control in collaborative human-robot transportation. In *Proceedings of the IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. 2646–2651.

[16] Eloise Matheson, Riccardo Minto, Emanuele GG Zampieri, Maurizio Faccio, and Giulio Rosati. 2019. Human–robot collaboration in manufacturing applications: A review. *Robotics* 8, 4 (2019), 100.

[17] Christoforos Mavrogiannis, Patrícia Alves-Oliveira, Wil Thomason, and Ross A. Knepper. 2022. Social Momentum: Design and Evaluation of a Framework for Socially Competent Robot Navigation. *Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction (HRI)* 11, 2 (2022), 1–37.

[18] Christoforos Mavrogiannis, Krishna Balasubramanian, Sriyash Poddar, Anush Gandra, and Siddhartha S. Srinivasa. 2023. Winding Through: Crowd Navigation via Topological Invariance. *IEEE Robotics and Automation Letters* 8, 1 (2023), 121–128.

[19] Erich Mielke, Eric Townsend, David Wingate, John L. Salmon, and Marc D. Killpack. 2024. Human-robot planar co-manipulation of extended objects: data-driven models and control from human-human dyads. *Frontiers in Neurorobotics* 18 (2024).

[20] Alexander Mörtl, Martin Lawitzky, Ayse Kucukyilmaz, Metin Sezgin, Cagatay Basdogan, and Sandra Hirche. 2012. The role of roles: Physical cooperation between humans and robots. *The International Journal of Robotics Research* 31, 13 (2012), 1656–1674.

[21] Eley Ng, Ziang Liu, and Monroe Kennedy. 2023. It takes two: Learning to plan for human-robot cooperative carrying. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*. 7526–7532.

[22] Stefanos Nikolaidis, David Hsu, and Siddhartha Srinivasa. 2017. Human-robot mutual adaptation in collaborative tasks: Models and experiments. *The International Journal of Robotics Research* 36, 5-7 (2017), 618–634.

[23] Stefanos Nikolaidis and Julie Shah. 2013. Human-robot cross-training: Computational formulation, modeling and evaluation of a human team training strategy. In *Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. 33–40.

[24] Dorsa Sadigh, Shankar Sastry, Sanjit A. Seshia, and Anca D. Dragan. 2016. Planning for Autonomous Cars that Leverage Effects on Human Actions. In *Proceedings of Robotics: Science and Systems (RSS)*.

[25] Natalie Sebanz, Harold Bekkering, and Günther Knoblich. 2006. Joint action: bodies and minds moving together. *Trends in Cognitive Sciences* 10, 2 (2006), 70–76.

[26] Mario Selvaggio, Marco Cognetti, Stefanos Nikolaidis, Serena Ivaldi, and Bruno Siciliano. 2021. Autonomy in Physical Human-Robot Interaction: A Brief Survey. *IEEE Robotics and Automation Letters* 6, 4 (2021), 7989–7996.

[27] J Ernesto Solanes, Luis Gracia, Pau Munoz-Benavent, Jaime Valls Miro, Marc G Carmichael, and Josep Tornero. 2018. Human–robot collaboration for safe object transportation using force feedback. *Robotics and Autonomous Systems* 107 (2018), 196–208.

[28] Tijn van der Zant and Luca Iocchi. 2011. Robocup@ home: Adaptive benchmarking of robot bodies and minds. In *International Conference on Social Robotics*. 214–225.

[29] Paul Vernaza, Venkatraman Narayanan, and Maxim Likhachev. 2013. Efficiently Finding Optimal Winding-Constrained Loops in the Plane. In *Proceedings of Robotics: Science and Systems (RSS)*.

[30] Elvin Yang and Christoforos Mavrogiannis. 2025. Implicit Communication in Human-Robot Collaborative Transport. arXiv:2502.03346 [cs.RO] https://arxiv.org/abs/2502.03346

[31] Sriram Yenamandra, Arun Ramachandran, Mukul Khanna, Karmesh Yadav, Devendra Singh Chaplot, Gunjan Chhablani, Alexander Clegg, Theophile Gervet, Vidhi Jain, Ruslan Partsey, Ram Ramrakhya, Andrew Szot, Tsung-Yen Yang, Aaron Edsinger, Charlie Kemp, Binit Shah, Zsolt Kira, Dhruv Batra, Roozbeh Mottaghi, Yonatan Bisk, and Chris Paxton. 2023. The HomeRobot Open Vocab Mobile Manipulation Challenge. In *Advances in Neural Information Processing Systems (NeurIPS): Competition Track*. https://aihabitat.org/challenge/2023_homerobot_ovmm/

Received 12 February 2025; revised 26 February 2025