MITSUBISHI ELECTRIC RESEARCH LABORATORIES
https://www.merl.com

Open Human-Robot Collaboration Systems (OHRCS): A Research Perspective

Suresh, Prasanth; Romeres, Diego; Dosh, Prashant; Jain, Siddarth

TR2024-150 November 02, 2024

Abstract

Human-robot collaboration (HRC) is the paradigm of humans and robots working synergistically in a shared workspace toward common goals. Prior research models such collaborative scenarios as multiagent systems composed of a fixed number of agents. Such models where the number and type of agents remain constant throughout, are termed closed systems. Conversely, a human-robot collaborative where the team size dynamically changes during the task is called an open HRC system (OHRCS). OHRCS allows for a realistic representation of human-robot collaboration by allowing agents to join or leave the task as needed. In this paper, we posit that many real-world HRC scenarios are better modeled as OHRCS. We present our vision of OHRC, present potential applications, examine the benefits of openness in HRC, and provide some avenues for future research.

IEEE International Conference on Cognitive Machine Intelligence (CogML 2024)

Mitsubishi Electric Research Laboratories, Inc.
201 Broadway, Cambridge, Massachusetts 02139

Open Human-Robot Collaboration Systems (OHRCS): A Research Perspective

Prasanth Sengadu Suresh
THINC Lab
School of Computing
University of Georgia
Athens, GA, USA
ps32611@uga.edu

Diego Romeres
Mitsubishi Electric Research Labs
Cambridge, MA, USA
romeres@merl.com

Prashant Doshi
THINC Lab
School of Computing
University of Georgia
Athens, GA, USA
pdoshi@uga.edu

Siddarth Jain
Mitsubishi Electric Research Labs
Cambridge, MA, USA
sjain@merl.com

***Abstract*—Human-robot collaboration (HRC) is the paradigm of humans and robots working synergistically in a shared workspace toward common goals. Prior research models such collaborative scenarios as multiagent systems composed of a fixed number of agents. Such models where the number and type of agents remain constant throughout, are termed *closed* systems. Conversely, a human-robot collaborative where the team size dynamically changes during the task is called an *open* HRC system (OHRCS). OHRCS allows for a realistic representation of human-robot collaboration by allowing agents to join or leave the task as needed. In this paper, we posit that many real-world HRC scenarios are better modeled as OHRCS. We present our vision of OHRC, present potential applications, examine the benefits of openness in HRC, and provide some avenues for future research.**

***Index Terms*—Human-robot collaboration, agent-openness, multiagent systems**

I. Introduction

Collaboration plays a crucial role in the success of human endeavors. As robots become an integral part of human society, they need to work hand-in-hand with human teams to contribute effectively. However, accomplished human teams employ a variety of techniques for enhanced coordination and collaboration [1]. In order for humans and robots to collaborate in a shared workspace, all agents must factor in the behavior of the other(s) during decision-making. This scenario naturally lends itself to a multiagent system [2] that can be used to model and learn behavioral policies.

This paper focuses on multiagent systems involving flexible collaboration between humans and robots. Previous work on human-robot collaboration (HRC) often assumes a predetermined and unchanging set of agents throughout the task. However, in certain scenarios, collaborative efficiency could be enhanced if the human joins in when needed and exits upon completing their role in the task. For example, envision a factory floor setup (see Fig. 1) where a collaborative robot begins assembling a part; upon reaching a stage that requires human assistance, signals for help through an alarm, or a flashing light. The human factory worker arrives, assists the robot with the collaborative subtask, and leaves to attend to other tasks. We term such a system, where human or robotic agents can enter and exit the task as needed, an *open* HRC (OHRC) system [3]. Observe that OHRC manifests a form of open agent systems [4] through its ability to allow agents to join or depart as needed. Open agent systems literature terms this artifact as agent openness (AO). A noteworthy feature of OHRC is that while the human is absent from the environment, the robot could switch to *industrial mode* of operation, thereby working at its maximum speed to optimize throughput, and switching back to *collaborative mode* when the human enters to safely team-up with them. This combines the speed of an industrial robot, the safety and agility of a collaborative robot, and the dexterity of a human, to enhance HRC competency.

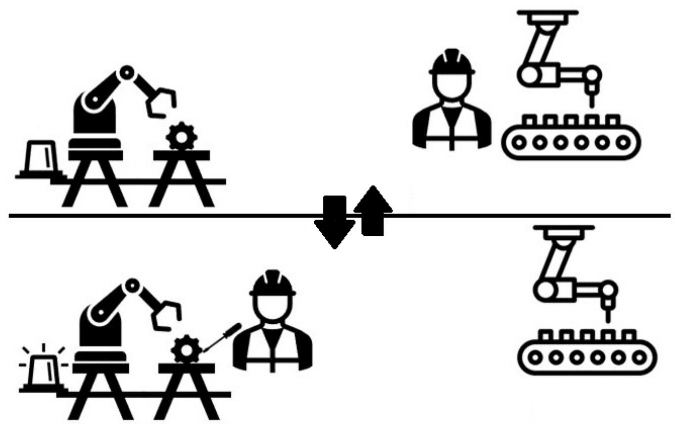

Fig. 1: Open HRC - Illustration of a factory floor scenario where a collaborative robot and a human operator are working on two different tasks independently. The robot reaches a point where it requires human assistance, signals, and calls a human operator. The human operator arrives, assists the robot in completing the collaborative subtask, and then leaves to attend to other tasks.

A related paradigm that discusses flexible synergy is ad hoc teamwork [5, 6]. However, an important distinction is that ad hoc teamwork deals with real-time adaptation to unknown teammates without prior coordination while OHRC focuses on collaboration within an open environment with dynamically

P. Doshi was supported in part by the NSF under grant number IIS-2312657.

changing teams. We begin by positioning OHRCS in the context of related settings. These include viewing HRC as a classic multiagent system and comparing and contrasting OHRC with the related paradigm of ad hoc teamwork [5]. Since OHRC manifests open agent systems [4], we review recent approaches for modeling and acting optimally in such systems under agent openness. Next, we introduce human-robot task domains where agent openness organically manifests in collaboration. These motivating domains illustrate several modeling-, behavioral-, and implementation-centric challenges OHRC presents. Toward this, we describe these challenges and propose potential solution approaches, aiming to inspire future research endeavors.

II. Related Problems

In this section, we provide a concise overview of various pertinent issues, previous endeavors aimed at addressing them, and their associated constraints. We begin by examining studies that have formulated HRC as a multiagent system, along with the models they have employed for this purpose. Subsequently, we delve into the concept of ad hoc teamwork, which operates without assuming prior coordination or knowledge of agent types. We then explore open agent systems, where team members dynamically join and depart tasks, while goals may evolve. Finally, we analyze why prevailing methodologies and their adopted models are inadequately suited for OHRCS. This paper aims to summarize the limitations of current approaches in addressing OHRCS challenges; we refer the reader to surveys on HRC [17], ad hoc teamwork [5], and open agent systems [4] for a more comprehensive review of the topics.

A. Human-Robot Collaboration as Multiagent Systems

The primary aim of HRC is to develop systems that efficiently avail the combined strengths of human and robotic agents. Robots are assigned tasks that are repetitive or physically demanding, enabling humans to focus on activities requiring cognitive and dexterous skills. Achieving effective collaboration in shared workspaces necessitates both human and robotic agents to respond appropriately to each other's actions and environmental changes.

The selection of a suitable multiagent decision-making model depends on the agents' natures and objectives. For instance, in scenarios where agents are competitive, the task can be framed as a Markov game, with an equilibrium condition ensuring the best strategy profile [18]. Conversely, if agents are collaborative and possess complete knowledge of one another, a fully centralized framework like multiagent MDP (MMDP) can be employed. Alternatively, one might consider the human agent(s)' actions as impacting only the transition dynamics, allowing them to be marginalized; in such cases, a single-agent MDP suffices to learn just the robot's behavior.

In their exploration of HRC tasks, Nikolaidis et al. 2012 [19] and Chen et al. 2020 [20] model their HRC tasks as a single-agent partially observable MDP (POMDP) [21]. While Nikolaidis et al. 2015 [22] and 2017 [23] consider slightly simpler cases where only portions of the world state are partially observable and adopt a mixed-observability MDP (MOMDP) [24] to model the task. Conversely, Nikolaidis et al. 2017 [25] model their task as a two-player Markov game, where the human adapts their behavior with evolving expectations of the robot's capabilities. Alternatively, Unhelkar et al. 2020 [26] and Seo et al. 2022 [27] split their HRI scenario into an agent model and a task model, where the task model - MMDP, captures the task attributes and the agent Markov model, the mental states of the other agent.

A recent approach by Wang et al. 2022 [28] and Van der Spa et al. 2024 [29] model a simulated HRC handover and combined manipulation task using an MMDP and use inverse learning methods [30, 31] to learn collaborative joint policies. Along the same lines, recent works by Yuan et al. 2022 [32] and Jiang et al. 2024 [33] use a variational inference to discover a latent strategy of the human teammate for enhanced collaboration, and represent their task as a decentralized scenario. Sengadu et al. 2023 [34] models their HRC scenarios as a Dec-MDP [35] where the human and robot are each aware of their own local state and some task attributes only. A vector of policies (one for each agent) is learned and the robot policy is tested on a simulated patient assistance task [36] and a realistic collaborative produce sorting task [37]. While all these previous works solve key problems in HRC, notice that they all use *closed* models to represent their HRC scenarios and hence are unsuitable for OHRCS.

B. Ad Hoc Teamwork

A line of work that aims to address ad hoc teamwork [6] makes related contributions. Mirsky et al. 2022 [5] defines ad hoc teamwork as *"To create an autonomous agent that is able to efficiently and robustly collaborate with previously unknown teammates on tasks to which they are all individually capable of contributing as team members"*. While research into ad hoc teamwork has made great strides in the past decade, one key difference between such teamwork and OHRC is that the latter focuses on agents learning to collaborate within an open, dynamically changing team that allows both human and robotic agents to enter and exit the task as needed. For example, the task may begin as a dyadic human-robot team but evolve to engage more humans or vice versa. Therefore, OHRC expands the scope of ad hoc teamwork to settings where agents can join or leave the system while the task is ongoing. As such, the decision-making complexity of OHRC is considerably higher than that of ad-hoc teamwork, as we will see in the upcoming sections.

C. Open Agent Systems

Multiagent systems with the additional relaxation of openness introduce several new challenges since each agent must now also account for: agents that have entered or exited presently - agent openness (AO), type of agents that departed or joined in - type openness (TO), and any changes in goals - task openness (TaO). While prior research in multiagent systems has achieved broad success within closed systems, most real-world domains tend to be *open* [4]. Hence, developing

TABLE I: Table summarizing prior works in Open Agent Systems along multiple aspects. Terminology used: Ad Hoc - working with previously unseen teammates; Planned-CoOp - collaborating with known teammates towards a common goal; Self-interest - optimizing agent-specific cumulative return; AO- Agent Openness, TaO - Task Openness, TO - Type Openness.

Method/Paper	AO	TaO	TO	Experiment type	Task nature	Key contribution
Simao et al. 2001 [7]	✓	✓	✗	N/A	Planned-CoOp	Social reasoning framework
Jumadinova et al. 2013 [8]	✓	✓	✗	Sim	Ad Hoc	Teacher-learner framework
Chen et al. 2015 [9]	✓	✓	✗	Sim	Ad Hoc	Flexible ad hoc framework
Open-Dec-POMDP [10]	✓	✗	✗	Sim	Planned-CoOp	Open agent collab model
I-PBVI [11]	✓	✗	✓	Sim	Self-Interest	Posthoc & predictive planning
I-POMCP [12]	✓	✓	✗	Sim	Self-Interest	Scalable planning framework
CI-POMCP-PF [13]	✓	✓	✗	Sim	Self-Interest	Planning with communication
LIA2C [14]	✓	✓	✗	Sim	Self-Interest	DTDE-MARL for open systems
GPL-SPI [15]	✓	✗	✗	Sim	Ad Hoc	Graph RL for open systems
Fastap [16]	✗	✗	✓	Sim	Planned-CoOp	MARL for fast policy adaptation

sophisticated methods to model openness becomes crucial, especially when considering realistic domains such as HRC.

Each agent may only have limited information at every timestep regarding goals, active agents, their types, abilities, and level of openness. Therefore, the first challenge becomes modeling the task appropriately to capture such uncertainties. One of the earliest works in the field, Simao et al. 2001 [7], addresses the concept of agent openness from a social reasoning perspective that enables an agent to reason about others when the agents' organization is not available a priori. They hypothesize that such a mechanism can be used both as a basis for coalition and decision-making to adapt to dynamic changes. On the other hand, Jumadinova et al. 2013 [8] consider distributed collaboration among multiple agents with both TaO and AO. They use a teacher-learner framework to decide what capabilities to learn from the other agents and validate their model using a simulated agent-based modeling tool [38]. They conclude that agents learning all capabilities, regardless of usefulness, outperform others.

The work by Chen et al. 2015 [9] considers two types of openness - AO and TaO, to ascertain how they contribute to learning in an ad hoc setting. They assume that all agents know the level of AO and TaO and perform simulated experiments where agents can bid on tasks based on their knowledge level. The authors conclude that while TaO makes it harder for the agents to solve the tasks, AO improves the performance, as newer agents could bring additional capabilities to help solve the tasks. Cohen et al. 2017 [10] provide a modified Dec-POMDP framework [35] that can be applied to open systems by factoring in the coalition model during decision-making. They use an offline best-response algorithm and Monte Carlo tree search (MCTS) [39] to plan appropriate actions under different coalitions. They validate their model using a simulated urban firefighting domain where new agents are called in to extinguish fires within a grid. However, in their coalition transition model, agents make the decision to transition to a new coalition solely based on the previous coalition, which is unrealistic. Additionally, the methods used for planning such as the best-response algorithm and MCTS do not scale well to real-world domains.

Using an IPOMDP-Lite framework, Hoang et al. 2013 [40] and Chandrasekaran et al. 2016 [11] model an open system, and use an agent interaction graph for post hoc reasoning. They validate their model on the simulated wildfire-suppression domain and compare their learned policies with heuristic-based, and random policies. Eck et al. 2020 [12] extend this method to be more scalable by selectively modeling the neighboring agents and providing theoretical bounds for the same using regret analysis. They consider more complex scenarios of the wildfire-suppression domain and compare their learned policy with heuristics-based policies.

A novel model called CI-POMDP, proposed by Kakarlapudi et al. 2022 [13], improves upon the IPOMDP-Lite framework by introducing communication between agents. They posit that by communicating, agents can better navigate the challenges of an open system. They use the wildfire-suppression domain to validate their claims, compare various levels of communication, and analyze corresponding costs. Recent work by He et al. 2023 [14] uses a multiagent reinforcement learning (MARL) paradigm called decentralized training, decentralized execution (DTDE) policy gradient method with an underlying I-POMDP model to solve a simulated open-organization problem where employees can join or quit their jobs at any point. They propose a novel method to factor partial-observability in open agent systems. Finally, in most recent works Rahman et al. 2023 [15] and Zhang et al. 2023 [16] use multiagent RL to address the complexities of openness. The former proposes a partially observable open stochastic Bayesian game to model AO and uses a graph-based policy learning approach to learn RL policies. The latter studies decision-making through multiagent RL when other agents' policies abruptly change during the task; which could be considered TO. Both works evaluate their method on simulated toy domains like Level-based Foraging, WorldPack, and PredatorPrey.

Some common denominators (as summarized in Table I) that can be considered limitations of the aforementioned prior works are that they evaluate their methods on small, simulated, toy domains like the urban firefighting or the open-organization domain, and largely do not provide theoretical or empirical analysis of convergence, regret, scalability, or sample complexity (when applicable). Methods using IPOMDPs or Game Theory approaches do not apply to OHRCS since they typically model competitive or self-interested scenarios.

Finally, OHRC aims to learn a human-centric solution that prioritizes the human agent's time and energy. Therefore, it is essential to demonstrate the effectiveness of the proposed methods on a physical system such as a robot in pragmatic settings, to ascertain their applicability to OHRC. Considering these factors, existing techniques may not be directly usable in OHRC, although they address similar concepts.

III. Open Human-Robot Collaborations

Open HRC (OHRC) refers to the paradigm of modeling HRC with agent openness (AO), where agents (human and robotic) can join or leave the task at any point. In this paper, we only analyze AO, however, certain HRC scenarios may manifest type openness (TO) and task openness (TaO) as well. In OHRC, new agents join the task either when called in by a currently active agent or of their own volition, and any agent can choose to leave the task at any point. To explicate this, we start by classifying OHRC broadly into two categories:

- **Collaboration-for-efficiency (CE):** Whereas tasks can be completed by the robotic agent alone, collaborating with human(s) could improve task efficiency. An example of CE would be tasks like line-sorting and packing [34], where both precision and recall need to be maximized.
- **Collaboration-due-to-requirement (CR):** This pertains to tasks that contain subtasks that a robotic agent cannot complete alone. An example of CR would be when the task involves highly dexterous manipulation or when the subtask requires multiple agents collaborating simultaneously to complete (see Fig. 2b).

These two categories will be illustrated in the context of each domain of interest in the following sections.

A. Use-Inspired Domains for Motivation

In order to tether the concepts of OHRCS to realistic examples, we provide two motivating domains: Collaborative robots in manufacturing, and assistive robotics. In both domains, we discuss cases where CE and CR may be applicable and how they can be addressed.

1) Collaborative Robots in Manufacturing: Consider the task of collaborative furniture assembly [3] as shown in Fig. 2a. The goal of this task is to assemble a table consisting of multiple parts: base, leg-support1, leg-support2, leg1, leg2, two screws for each leg-support to connect them to the base, and two screws for each leg to screw them into their respective supports, as shown in Fig. 2c. The task can be completed in multiple valid orders. For instance, one may position both leg-supports on the base and screw them in before positioning their corresponding legs and screwing the legs into their respective supports. Alternatively, one may position leg-support1, screw it into the base, place the leg1, screw it into the leg-support1, and analogously repeat the sequence for the other parts to complete the assembly. Notice that the simple positioning actions can be done independently by the robot, while the screwing action requires the assistance of a human.

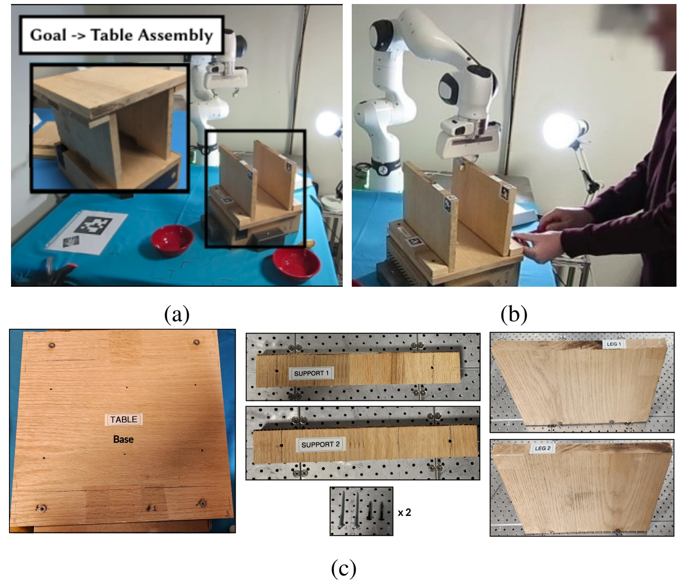

Fig. 2: A wooden table assembly task involving placing and screwing actions. (a) The assembled table. (b) The human screws one of the legs in while the robot holds it in place. (c) The base, supports and screws, and legs of the table.

- **CE:** In this furniture assembly example, the robot can independently position all the parts on the base at their respective locations. However, if a human is present in the environment, they can assemble parts in parallel, thus completing the positioning sooner. Nonetheless, availing a human's help for trivial positioning tasks is an ineffective use of their time and energy. Therefore, it is imperative to optimize the available resources to enhance task efficiency while minimizing human effort. Only if the task is time-sensitive or safety-critical can the robot engage the human's services from the beginning to speedily complete it.
- **CR:** Once the positioning of the supports is complete, the supports need to be screwed into the base before assembling the legs of the table. This screwing subtask requires dexterous manipulation and another agent to hold the part aligned with the screw holes. Even if the robot could screw in the part, since the part requires another hand to hold it in place, collaborating with a human is crucial. However, as mentioned before, since the assembly can be performed in multiple orders, the robot, in addition to learning the optimal method to assemble the parts, needs to learn when to call the human for help, and how to collaborate with them.
- **Other complexities:** Since human agents have limited time and energy, they may decide to leave after completing a subtask to focus on other tasks. Specifically, in the furniture assembly case, if the robot calls the human after positioning the first support, the human may arrive to assist with screwing the support in but may leave after that. In that case, the robot must call the human again

after placing the second support. Alternatively, the human may position the first leg alongside its support while the robot positions the second support. Since these subtasks may take different durations, if one agent finishes their subtask sooner and needs assistance, the other may choose to complete their current subtask before assisting them or disregard their current subtask and assist them immediately. Lastly, there may be a case where the human is forced to wait a few timesteps due to the unavailability of valid subtasks; leaving and rejoining at that point may be inefficient. For example, consider the case where the only subtasks left are positioning and screwing the last leg in place. Until the robot positions the leg, the human cannot screw it in, and exiting at this point only to rejoin shortly after may not be worthwhile.

2) Assistive Robotics: With the recent advances in robotics, assistive robots have become an increasingly necessary tool to provide care and assistance to the elderly and people with motor impairments. An assistive robot may aid them in daily activities such as bed-bathing, itch-scratching, feeding, etc., as shown in Fig. 3. As such, usually, one robot interacts with the subject to aid them; however, certain subtasks, such as bed-bathing, may require collaboration with additional agents to be accomplished. For example, to assist a patient in bathing, the robot may need to stand them up or turn them over and may require a hand on either side to lift them and turn them around. Evidently, this is an OHRC, and modeling it as a closed system may diminish the quality of patient care.

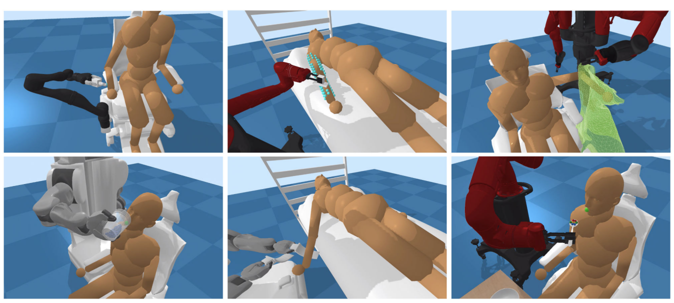

Fig. 3: Illustration of a few assistive tasks that could be performed by a collaborative robot such as itch-scratching, changing, feeding, bed-bathing, etc. Image credit: Assistive Gym [36].

- **CE:** In certain cases, multiple subtasks could be performed in parallel to enhance assistance, for example, the person could be feeling cold while the robot is engaged in feeding them. In this case, the robot could either stop feeding them to cover them with a blanket or finish feeding them before covering them with a blanket. Since these subtasks could be completed in parallel with another agent on the scene, a new agent (human or robotic) could arrive, cover the subject with a blanket, and then leave to attend to other tasks.
- **CR:** Alternatively, the robot could accomplish all other tasks on its own but when it reaches the bed-bathing scenario, as mentioned before, the subject may need to be held on both sides and flipped over. In this case, the robot may call another human or robot for assistance. Once the subject has bathed, the new agent could leave to attend to other tasks.

B. Novel Challenges of OHRC

In order to systematically characterize the challenges in OHRC, we classify them into three overarching categories:

1) **Modeling challenges:** The first important challenge to address in OHRCS is establishing an efficient way to model openness in HRC. Since HRC by definition is collaborative, typically the model has a single reward function designed toward the common goal. However, since the agents involved in HRC cannot perfectly observe all the attributes of the other agents (for example, human joint angles), the model has to support a decentralized execution approach. Models such as decentralized Markov decision processes (Dec-MDP) [35] can be extended to model agent openness. The global state and global action at a given timestep can be formed using the pooled states and pooled actions of the *agents active at that timestep*, thus incorporating openness. This would require the learning paradigm to accept inputs of states and actions that vary in size at every timestep based on the currently active agents. This global state, global action, and a common reward function can be used for training, to obtain decentralized policies (one for each agent that participated in the task during training). These decentralized policies would map the agent's local state to their local action and factor in the current team of that agent, to learn the appropriate action mapping.

   Certain modeling challenges of HRC may also manifest in OHRC, such as designing an appropriate reward function. However, this becomes more challenging in OHRC since multiple combinations of teams may accomplish the task similarly. Additionally, AO may be stochastic, where the agent may enter and exit the task with a certain probability, further increasing the decision-making complexity. Different confounding factors like sensor noise may render determining the presence or absence of a human uncertain.

   Finally, a critical challenge in navigating a shared workspace is recognizing and correctly modeling the *necessary interactions*. For instance, in the furniture assembly domain, the screwing subtask requires the robot to hold the part in place while the human screws it in. Such interactions can be termed *necessary* interactions. On the contrary, if both the robot and human are positioning parts in parallel and both agents try to position a part at the same location simultaneously, these can be termed as *adverse* interactions.

   OHRC further complicates such interactions due to agent openness. The robot needs to factor in the human's presence or absence first and then decide on a strategy to navigate potential interactions. Therefore, designing a model that can appropriately handle the challenges and complexities of OHRC is an avenue for future research.

2) **Behavioral challenges:** Learning behavioral policies for agents involved in OHRC is an ambitious task. These agent policies have to account for: the currently active agents in the environment, actively collaborating agents, when new agents are needed, and when the current agent(s) must exit the task. A popular approach used in multiagent reinforcement learning [41] is centralized training and decentralized execution, which could be extended to consider agent-openness, to learn OHRC policies. Additional factors such as partial observability and stochasticity may complicate behavioral policy learning by providing noisy and uncertain feedback about the currently active agents and the environment.
   Agent-based modeling techniques [26, 42, 43], Theory of Mind [44, 45] or Prospect Theory [46] could be incorporated to learn the entry and exit pattern of an agent, based on prior experience, their current energy level, level of rationality, etc. Such latent decision-making factors may also manifest in other forms of openness, such as task-openness (TaO) or type-openness (TO). For example, the human, upon joining the task may decide to focus on a different subtask than the one initially intended. The robot must be capable of recognizing and adapting to such TaO. Furthermore, humans may modify their behavior mid-task by, say, becoming less collaborative due to fatigue, and such TO must be handled by the robot's policy.
3) **Implementation challenges:** While all the modeling and learning challenges are quite complex, implementing learned policies in the real world and achieving seamless collaboration in OHRC may be the most challenging of all. Depending on the robot and the task itself, several unforeseen challenges may arise. For example, physical limitations such as dynamic environmental occlusions or abruptly malfunctioning sensors may make it difficult to assess the presence of a particular agent, further complicating the decision to call a new agent.
   Furthermore, due to limited signal strength or environmental distractions, agents may not receive the call to join the task (say a loud machine on the factory floor prevents the human from noticing the robot's call for help). Humans may also decide to abruptly exit the task due to an emergency, or their shift ending, which may throw off the learned policy since it may not have happened during training. Sensing such abnormalities and adapting accordingly is one of the major implementation challenges. A data-driven model that learns from human-human team experiences (e.g. inverse learning methods [47, 48]) combined with feedback-based improvement (such as active learning [49] or using clarification requests [50]) could mitigate such difficulties.

IV. Concluding remarks and future work

Navigating the multifaceted challenges of OHRC and translating its principles into practical applications underscores the necessity of addressing these challenges. While achieving seamless OHRC may prove challenging, integrating agent-openness shows promise in unlocking a plethora of applications for effective human-robot collaboration. By enabling agents to enter and exit the system as needed, OHRC not only enhances collaboration but also empowers humans to engage in multiple tasks simultaneously. By combining an industrial robot's expeditiousness, with a collaborative robot's safety and agility, and a human's dexterity, OHRC nurtures a symbiotic relationship that dynamically adapts to evolving demands.

A. Future work

As the first work that discusses agent-openness from a HRC perspective, we assert that most real-world HRC domains ought to be modeled as OHRC to capture the richness and flexibility of real-world collaborations. By systematically addressing the challenges presented, the potential of OHRC to revolutionize industries such as manufacturing, healthcare, and service robotics can be fully realized. Future research must focus on refining the protocols that govern OHRCS, ensuring robustness and efficiency. We expect that by empirically and theoretically analyzing the complexity and effectiveness of OHRC compared to HRC, future works will bolster our claim and move towards open collaboration. Ultimately, our work aims to inspire further exploration and innovation, setting a foundation for a new era of human-robot collaboration that is as flexible and dynamic as the challenges it seeks to overcome.

References

[1] K. B. Reed and M. A. Peshkin, "Physical collaboration of human-human and human-robot teams," *IEEE transactions on haptics*, vol. 1, no. 2, pp. 108–120, 2008.

[2] J. Ferber and G. Weiss, *Multi-agent systems: an introduction to distributed artificial intelligence*. Addison-wesley Reading, 1999, vol. 1.

[3] P. S. Suresh, S. Jain, D. Romeres, and P. Doshi, "Open human-robot collaboration using decentralized inverse reinforcement learning," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2024.

[4] A. Eck, L.-K. Soh, and P. Doshi, "Decision making in open agent systems," *AI Magazine*, 2023.

[5] R. Mirsky, I. Carlucho, A. Rahman, E. Fosong, W. Macke, M. Sridharan, P. Stone, and S. V. Albrecht, "A survey of ad hoc teamwork: Definitions, methods, and open problems," in *European Conference on Multiagent Systems*, 2022.

[6] P. Stone, G. Kaminka, S. Kraus, and J. Rosenschein, "Ad hoc autonomous agent teams: Collaboration without pre-coordination," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 24, no. 1, 2010, pp. 1504–1509.

[7] J. S. Simão and Y. Demazeau, "On social reasoning in multi-agent systems," *Inteligencia Artificial. Revista Iberoamericana de Inteligencia Artificial*, vol. 5, no. 13, p. 0, 2001.

[8] J. Jumadinova, P. Dasgupta, and L.-K. Soh, “Strategic capability-learning for improved multiagent collaboration in ad hoc environments,” *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 44, no. 8, pp. 1003–1014, 2013.

[9] B. Chen, X. Chen, A. Timsina, and L.-K. Soh, “Considering agent and task openness in ad hoc team formation.” in *AAMAS*, 2015, pp. 1861–1862.

[10] J. Cohen, J.-S. Dibangoye, and A.-I. Mouaddib, “Open decentralized pomdps,” in *2017 IEEE 29th International Conference on Tools with Artificial Intelligence (ICTAI)*. IEEE, 2017, pp. 977–984.

[11] M. Chandrasekaran, A. Eck, P. Doshi, and L. Soh, “Individual planning in open and typed agent systems,” in *Proceedings of the Thirty-Second Conference on Uncertainty in Artificial Intelligence*, 2016, pp. 82–91.

[12] A. Eck, M. Shah, P. Doshi, and L.-K. Soh, “Scalable decision-theoretic planning in open and typed multiagent systems,” in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, no. 05, 2020, pp. 7127–7134.

[13] A. Kakarlapudi, G. Anil, A. Eck, P. Doshi, and L.-K. Soh, “Decision-theoretic planning with communication in open multiagent systems,” in *Uncertainty in Artificial Intelligence*. PMLR, 2022, pp. 938–948.

[14] K. He, P. Doshi, and B. Banerjee, “Latent interactive a2c for improved rl in open many-agent systems,” *arXiv preprint arXiv:2305.05159*, 2023.

[15] A. Rahman, I. Carlucho, N. Höpner, and S. V. Albrecht, “A general learning framework for open ad hoc teamwork using graph-based policy learning,” *Journal of Machine Learning Research*, vol. 24, no. 298, pp. 1–74, 2023.

[16] Z. Zhang, L. Yuan, L. Li, K. Xue, C. Jia, C. Guan, C. Qian, and Y. Yu, “Fast teammate adaptation in the presence of sudden policy change,” *arXiv preprint arXiv:2305.05911*, 2023.

[17] A. Dahiya, A. M. Aroyo, K. Dautenhahn, and S. L. Smith, “A survey of multi-agent human–robot interaction systems,” *Robotics and Autonomous Systems*, vol. 161, p. 104335, 2023.

[18] M. L. Littman, “Markov games as a framework for multi-agent reinforcement learning,” in *Machine learning proceedings 1994*. Elsevier, 1994, pp. 157–163.

[19] S. Nikolaidis and J. Shah, “Human-robot teaming using shared mental models,” *ACM/IEEE HRI*, 2012.

[20] M. Chen, S. Nikolaidis, H. Soh, D. Hsu, and S. Srinivasa, “Trust-aware decision making for human-robot collaboration: Model learning and planning,” *ACM Transactions on Human-Robot Interaction (THRI)*, vol. 9, no. 2, pp. 1–23, 2020.

[21] G. E. Monahan, “State of the art—a survey of partially observable markov decision processes: theory, models, and algorithms,” *Management science*, vol. 28, no. 1, pp. 1–16, 1982.

[22] S. Nikolaidis, R. Ramakrishnan, K. Gu, and J. Shah, “Efficient model learning from joint-action demonstrations for human-robot collaborative tasks,” in *Proceedings of the tenth annual ACM/IEEE international conference on human-robot interaction*, 2015, pp. 189–196.

[23] S. Nikolaidis, Y. X. Zhu, D. Hsu, and S. Srinivasa, “Human-robot mutual adaptation in shared autonomy,” in *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*, 2017, pp. 294–302.

[24] S. C. Ong, S. W. Png, D. Hsu, and W. S. Lee, “Planning under uncertainty for robotic tasks with mixed observability,” *The International Journal of Robotics Research*, vol. 29, no. 8, pp. 1053–1068, 2010.

[25] S. Nikolaidis, S. Nath, A. D. Procaccia, and S. Srinivasa, “Game-theoretic modeling of human adaptation in human-robot collaboration,” in *Proceedings of the 2017 ACM/IEEE international conference on human-robot interaction*, 2017, pp. 323–331.

[26] V. V. Unhelkar, “Effective information sharing for human-robot collaboration,” Ph.D. dissertation, Massachusetts Institute of Technology, 2020.

[27] S. Seo and V. V. Unhelkar, “Semi-supervised imitation learning of team policies from suboptimal demonstrations,” *arXiv preprint arXiv:2205.02959*, 2022.

[28] C. Wang, C. Pérez-D’Arpino, D. Xu, L. Fei-Fei, K. Liu, and S. Savarese, “Co-gail: Learning diverse strategies for human-robot collaboration,” in *Conference on Robot Learning*. PMLR, 2022, pp. 1279–1290.

[29] L. van der Spaa, J. Kober, and M. Gienger, “Simultaneously learning intentions and preferences during physical human-robot cooperation,” *Autonomous Robots*, vol. 48, no. 4, p. 11, 2024.

[30] J. Ho and S. Ermon, “Generative adversarial imitation learning,” *Advances in neural information processing systems*, vol. 29, 2016.

[31] B. D. Ziebart, A. L. Maas, J. A. Bagnell, A. K. Dey *et al.*, “Maximum entropy inverse reinforcement learning.” in *Aaai*, vol. 8. Chicago, IL, USA, 2008, pp. 1433–1438.

[32] L. Yuan, J. Wang, F. Zhang, C. Wang, Z. Zhang, Y. Yu, and C. Zhang, “Multi-agent incentive communication via decentralized teammate modeling,” in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 36, 2022, pp. 9466–9474.

[33] R. Jiang, X. Zhang, Y. Liu, Y. Xu, X. Zhang, and Y. Zhuang, “Multi-agent cooperative strategy with explicit teammate modeling and targeted informative communication,” *Neurocomputing*, vol. 586, p. 127638, 2024.

[34] P. Sengadu Suresh, Y. Gui, and P. Doshi, “Dec-airl: Decentralized adversarial irl for human-robot teaming,” in *Proceedings of the 2023 International Conference on Autonomous Agents and Multiagent Systems*, 2023, pp. 1116–1124.

[35] C. V. Goldman and S. Zilberstein, “Decentralized control of cooperative systems: Categorization and complexity analysis,” *Journal of artificial intelligence research*, vol. 22, pp. 143–174, 2004.

[36] Z. Erickson, V. Gangaram, A. Kapusta, C. K. Liu,

and C. C. Kemp, “Assistive gym: A physics simulation framework for assistive robotics,” *IEEE International Conference on Robotics and Automation (ICRA)*, 2020.
[37] P. S. Suresh and P. Doshi, “Marginal map estimation for inverse rl under occlusion with observer noise,” in *Uncertainty in Artificial Intelligence*. PMLR, 2022, pp. 1907–1916.
[38] “Repast framework,” https://repast.github.io/index.html, 2023.
[39] C. B. Browne, E. Powley, D. Whitehouse, S. M. Lucas, P. I. Cowling, P. Rohlfshagen, S. Tavener, D. Perez, S. Samothrakis, and S. Colton, “A survey of monte carlo tree search methods,” *IEEE Transactions on Computational Intelligence and AI in games*, vol. 4, no. 1, pp. 1–43, 2012.
[40] T. N. Hoang and K. H. Low, “Interactive pomdp lite: Towards practical planning to predict and exploit intentions for interacting with self-interested agents,” *arXiv preprint arXiv:1304.5159*, 2013.
[41] S. V. Albrecht, F. Christianos, and L. Schäfer, “Multi-agent reinforcement learning: Foundations and modern approaches,” *Massachusetts Institute of Technology: Cambridge, MA, USA*, 2023.
[42] N. Gilbert, *Agent-based models*. Sage Publications, 2019.
[43] S. Nikolaidis, J. Forlizzi, D. Hsu, J. Shah, and S. Srinivasa, “Mathematical models of adaptation in human-robot collaboration,” *arXiv preprint arXiv:1707.02586*, 2017.
[44] A. M. Leslie, O. Friedman, and T. P. German, “Core mechanisms in ‘theory of mind’,” *Trends in cognitive sciences*, vol. 8, no. 12, pp. 528–533, 2004.
[45] P. J. Gmytrasiewicz and P. Doshi, “A framework for sequential planning in multi-agent settings,” *Journal of Artificial Intelligence Research*, vol. 24, pp. 49–79, 2005.
[46] M. Kwon, E. Biyik, A. Talati, K. Bhasin, D. P. Losey, and D. Sadigh, “When humans aren’t optimal: Robots that collaborate with risk-aware humans,” in *Proceedings of the 2020 ACM/IEEE international conference on human-robot interaction*, 2020, pp. 43–52.
[47] S. Arora and P. Doshi, “A survey of inverse reinforcement learning: Challenges, methods and progress,” *Artificial Intelligence*, vol. 297, p. 103500, 2021.
[48] A. Hussein, M. M. Gaber, E. Elyan, and C. Jayne, “Imitation learning: A survey of learning methods,” *ACM Computing Surveys (CSUR)*, vol. 50, no. 2, pp. 1–35, 2017.
[49] B. Settles, “Active learning literature survey,” 2009.
[50] T. Williams, F. Yazdani, P. Suresh, M. Scheutz, and M. Beetz, “Dempster-shafer theoretic resolution of referential ambiguity,” *Autonomous Robots*, vol. 43, pp. 389–414, 2019.