2024 IEEE International Conference on Robotics and Automation (ICRA 2024)
May 13-17, 2024. Yokohama, Japan

Proactive Robot Control for Collaborative Manipulation Using Human Intent

Zhanibek Rysbek, Siyu Li, Afagh Mehri Shervedani, Miloš Žefran

Abstract— **Collaborative manipulation task often requires negotiation using explicit or implicit communication. An important example is determining where to move when the goal destination is not uniquely specified, and who should lead the motion. This work is motivated by the ability of humans to communicate the desired destination of motion through back-and-forth force exchanges. Inherent to these exchanges is also the ability to dynamically assign a role to each participant, either taking the initiative or deferring to the partner's lead. In this paper, we propose a hierarchical robot control framework that emulates human behavior in communicating a motion destination to a human collaborator and in responding to their actions. At the top level, the controller consists of a set of finite-state machines corresponding to different levels of commitment of the robot to its desired goal configuration. The control architecture is loosely based on the human strategy observed in the human-human experiments, and the key component is a real-time intent recognizer that helps the robot respond to human actions. We describe the details of the control framework, feature engineering and training process of the intent recognition. The proposed controller was implemented on a UR10e robot (Universal Robots) and evaluated through human studies. The experiments show that the robot correctly recognizes and responds to human input, communicates its intent clearly, and resolves conflict. We report success rates and draw comparisons with human-human experiments to demonstrate the effectiveness of the approach.**

I. Introduction

Robots that can physically interact with humans are becoming increasingly available. However, to be useful in practice, such robots must be able to interact with humans in a transparent and predictable way that is similar to how humans interact with other humans. Physical interaction is often related to multi-modality, an important field of study in Human-Robot Interaction (HRI) [1], [2], [3], [4], [5]. A challenging task within physical HRI (pHRI) is collaborative manipulation, where a dyad remains in physical contact throughout and communicates haptically [6], [7], [8], [5].

The specific aspect of the collaborative manipulation task studied in this work is the process through which the participants negotiate where to move, and how to move. The real-world scenario where such a negotiation takes place would be a sudden occurrence of an obstacle that requires the dyad to change direction of motion, with several options available; the dyad needs to agree on one. Another example would be moving a heavy pot to the table, where it can be placed at several possible locations; again, the dyad needs to choose one. Although such negotiations often involve several modalities, and in particular verbal communication, none of them except force interaction is essential. Therefore, in this study, we specifically exclude other communication modalities and investigate how participants use force exchanges alone to reach a consensus.

Collaborative manipulation has been studied extensively. A substantial body of work uses models of human reaching movements [9], [10] to control the interaction, often using rule-based intent recognition [11], [12]. In these studies, the main objective is to assist the human by accelerating or decelerating the manipulated object to minimize the effort [13], [12]. A similar approach is to use machine learning techniques to predict *a priori* defined conflicting states from haptic signals and use them to switch between different controllers [14].

Another popular approach is to use programming by demonstration. Initially targeting point-to-point reaching movements [15], the approach was later successfully adapted to pHRI applications [16], [17], [18]. While exhibiting good performance in learning low-level force and velocity characteristics of the task, the interaction with the human is not explicitly controlled and the robot has limited ability to respond to a large variety of human behaviors.

An important aspect of implementing a controller for collaborative manipulation is the low-level control of the point of contact. The predominant approach in the literature for contact-rich tasks is based on impedance/admittance control [19]. Specifically for pHRI, [20], [21] designed the impedance parameters to match the impedance of the human arm. To adapt to different phases of interaction, impedance gain scheduling is a common practice [13], [22], [23], [14]. For human-humanoid collaborative manipulation, there are several efforts [24], [25] that focus on low-level humanoid control and are inspired by human-human experiments for balancing tasks, where equilibrium trajectory position is modulated using admittance control.

A crucial part of pHRI control is accurate intent recognition. Most of the work is focused on inferring the intended goal based on the kinematic data for tasks that require positional coordination such as grasping, reaching motions, and ball catching [26], [27], [28], [29], [30]. Obtaining an accurate human intent from force data remains challenging since the force components are unpredictable due to task specifics. For example, in [3], stiff contact with the environment distorts the underlying intent; therefore, the authors had to use additional sensors. In our case, the interaction forces are influenced by the grasping forces and the walking

This work has been supported by the National Science Foundation grants IIS-1705058, CMMI-1762924, and CCF-2240532.

All the authors are with the Robotics Laboratory, Department of Electrical and Computer Engineering, the University of Illinois at Chicago, Chicago, IL 60607, USA.

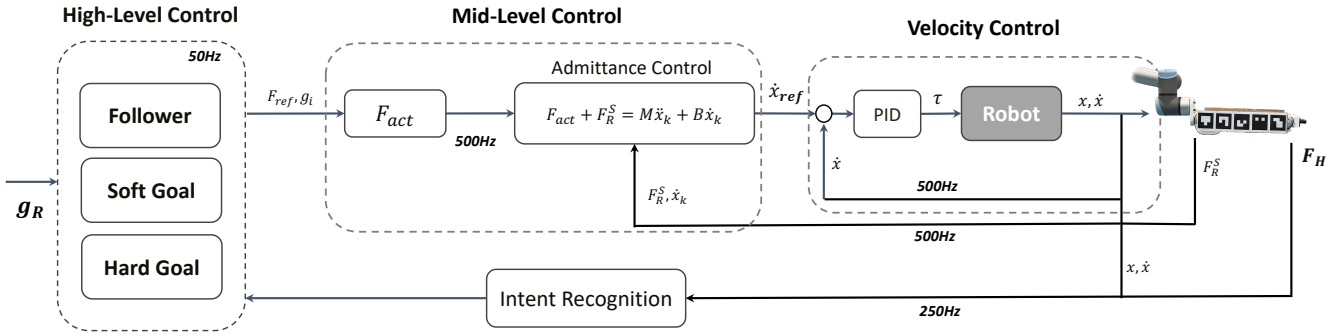


Fig. 1: Robot Control Architecture.

patterns [31].

In most of the existing work on collaborative manipulation, the robot follows the human lead. In contrast, our work focuses on enabling the robot to negotiate where to move and who will lead the motion. The main contributions of the work are three-fold: (a) A high-level controller that allows the robot to communicate its intent and respond to human actions. The control architecture is motivated by our studies of human-human collaborative manipulation [31]. (b) A real-time intent recognition module for robot control that uses novel force/kinematics features as described in [31]. In contrast, in the related work [14], [6], the haptic channel was used exclusively to infer the interaction state rather than the real-time human intent. (c) A human study that validates the framework. The study shows that the framework achieves the stated goals where humans feel that the robot communicates intent, and appropriately responds to their action.

II. Approach

In collaborative manipulation, there are three important variables to consider: *goal location*, *commitment*, and *what is known* by each agent. The *goal location* is often dependent on the manipulated object. Without loss of generality, we consider generic goal locations $g_i \in SE(2), i = 1, \ldots, n$ for a planar task. The *commitment* variable describes how strongly each agent wants to reach the goal. More committed agents tend to escalate the interaction forces while less committed agents may give up their goal and assume a follower role [8].

In this work, we consider a scenario where each participant is independently assigned a goal (in our case $n = 3$) and the level of commitment (*hard*, *soft*, or *none*, meaning that the participant has no goal and needs to follow). A hard goal requires the participant to go to the assigned goal, convincing the partner to comply if necessary. A soft goal should be reached if possible, but if necessary it can be given up and the partner followed instead. Given this task, the robot needs to have three principal controllers: 1) *Follower* controller, 2) *Soft Goal* controller, and 3) *Hard Goal* controller. For a detailed description of the task and the human-human experiments, see [31].

III. Control Architecture

A. Architecture

The proposed robot control architecture for the interaction task comprises three distinct layers, as illustrated in Fig. 1. The lower layer employs a Cartesian twist (velocity) controller operating at a frequency of 500Hz. The middle layer implements an admittance controller which provides compliance during the motion. This layer accepts the reference signal from the High-Level Controller (HLC). Finally, the HLC is a set of state machines responsible for imitating human behavior according to the roles outlined in Sect. II. Each state machine is triggered according to a predefined robot goal g_R and responds to human intent feedback (Fig. 1).

B. Admittance Control

The middle-layer controller consists of an admittance control law with inertia and damping terms [19], [32]. We use a force-torque sensor mounted between the end-effector and the object to measure force feedback F_R^S. The control equation in the discrete-time domain is:

$$M\ddot{x}[k] + B\dot{x}[k] = F_{act}[k] + F_R^S[k] \tag{1}$$

where M and B represent inertia and damping terms, $\dot{x}[k]$, and $\ddot{x}[k]$ are the twist and acceleration (in $se(3)$) and k is the time step. Assuming that $\dot{x}[k] = \dot{x}[k-1] + \Delta T\ddot{x}[k]$ where ΔT is the control rate, the admittance control law becomes:

$$\ddot{x}[k] = (M + B\Delta T)^{-1}\left(F_{act}[k] + F_R^S[k] - B\dot{x}[k-1]\right)$$

When implemented on the robot, a saturation block is added to keep the robot's speed within safe limits.

C. Robot Action Force

The robot action force block F_{act} in Fig. 1 accepts inputs specifying the desired force magnitude F_{ref} and the goal direction g_i. In turn, it will generate robot action force F_{act} directed toward g_i and interpolate it at 500 Hz. To avoid sudden force changes that might cause aggressive movements, the force is designed to reach the reference point within a time period $T_{transient}$. The value of $T_{transient}$ is chosen to be 0.2 seconds considering the human reaction time [33]. Moreover, F_{act} is set to zero when the object

reaches any of the goal sites. The robot action force is defined as follows:

$$F_{act}[k] = F_{act}[k-1] + \frac{F_{ref} - F_{act}[k-1]}{T_{transient}} \Delta T \quad (2)$$

Note that, F_{ref} is confined to the safe limits $[F_{min}, F_{max}]$ (see the *clip* function in Figs. 3 and 4), hence the magnitude of F_{act} has the same bounds. The dynamic response of F_{act} to F_{ref} can be seen in Fig. 2.

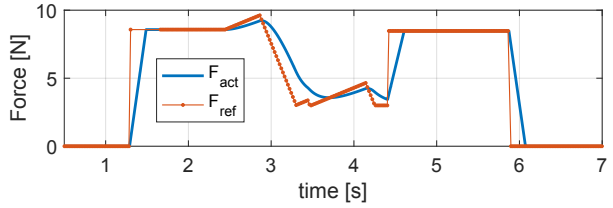


Fig. 2: Robot Force Reference Tracking Example.

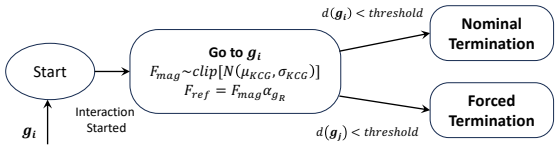
D. High-Level Reasoning

The high-level control module consists of three state machines corresponding to different types of robot goals as outlined in Sect. II. Each state machine runs at 50Hz.

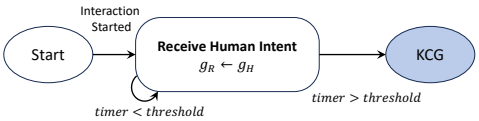
1) Known Common Goal (KCG): Known Common Goal (KCG) is a basic behavior that is a sub-block of the Follower, Soft, and Hard Goal controllers. KCG is invoked when the two agents agree on a goal. The controller simply generates a static F_{ref} towards g_i and stops when the target is reached (Fig. 3(a)). In addition, it stops at any other goal location in case a user decides to overpower the robot. Thus, it may terminate in two possible states: *Nominal Termination* and *Forced Termination*. It should be noted that several more elaborate versions of this behavior can be found in the literature [13], [11] so it is outside the scope of our work.

2) Follower: This module builds on the KCG with an added human intent perception block, as shown in Fig. 3(b). The human intent perception block accumulates the output of the intent recognizer, and when the timer threshold is reached, it switches to KCG by selecting the most likely goal. The timer threshold is designed to emulate the open-loop behavior of a human and provide enough time for them to perceive the partner's intent [10]. Note that the Follower controller does not require g_R as an input. If the human intent is misinterpreted or if it changes after the first round of interaction, the KCG module still has the flexibility to finish with *Forced Termination* state. We note that [14] implemented a similar version of this controller but inferred the intended continuous goal location from the measured force using a rule-based method.

3) Hard Goal: An important contribution of this paper is the ability of the robot to take initiative and lead the interaction. In the Hard Goal mode, the robot prioritizes its own goal, even though this creates a conflict. Inspired by observations from human-human experiments [31], the following logic is implemented as described in Fig. 4(a). First, the robot initializes the action force in the direction of the robot's goal. If a conflict is perceived via intent recognition during the current control step ($g_R \neq g_H$), then



(a) KCG



(b) Follower

Fig. 3: State machines for KCG and Follower controllers.

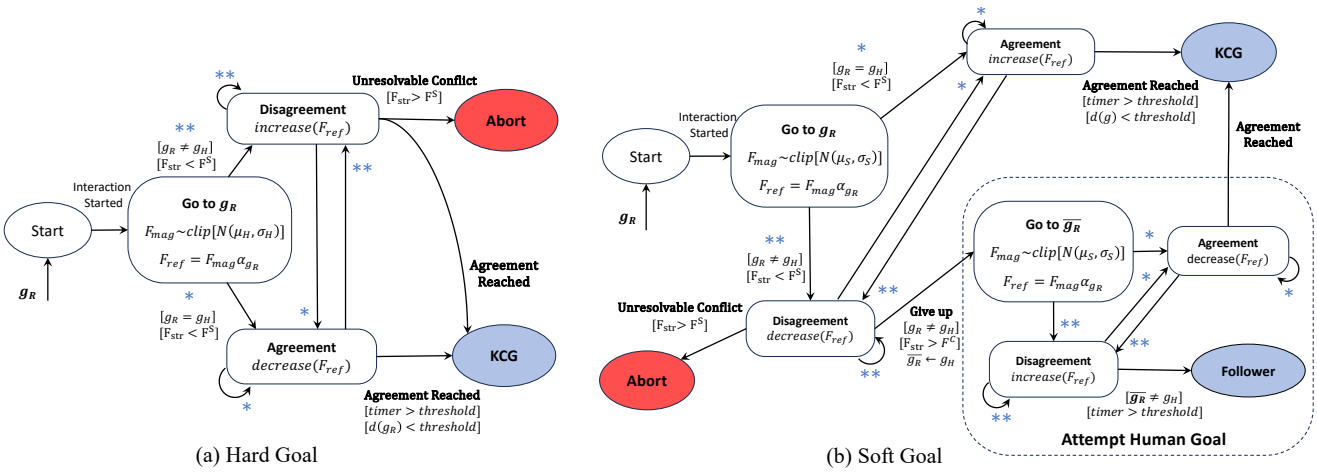
the magnitude of F_{ref} increases, otherwise ($g_R = g_H$) decreases. The rate of change depends on the current speed towards the robot goal g_R.

The stretch force $F_{str} = F_H - F_R$ is monitored for safety. If a human applies an excessive amount of force and overpowers the robot, the state machine transitions to the abort state, where the robot will gradually slow down and stop. We note that such situations occurred in human-human experiments. To resolve the conflict in this situation, humans either initiate a dialogue or use other nonverbal cues such as gestures, and restart the interaction.

4) Soft Goal: Another contribution of this work is a Soft Goal controller. In this mode, the robot can take the initiative or defer to the human. As can be seen in Fig. 4(b), the Soft Goal has an additional subtask **Attempt Human Goal** (AHG) compared to the Hard Goal. AHG is triggered if the stretch force F_{str} becomes large (but not large enough for abort). This threshold parameter is denoted as F^C. When triggering the AHG, the robot sets its goal to the perceived human intent and invokes the familiar agreement-disagreement cycle. However, if the robot misinterprets the human goal and spends too much time in the disagreement state, the robot switches to the Follower mode, trying to re-interpret human intent. If the robot interprets the human intent correctly, then the subtask terminates by transitioning to the KCG mode.

E. Robot Force Sampling

Since F_{ref} is the key control variable, it is important to carefully initialize the magnitude of the robot's force. Based on the predefined admittance gain, one can fine-tune the range of the robot force magnitude $F_{mag} = \|F_{ref}\|$ that will act in the spectrum from light to heavy. In this work, we chose to sample F_{mag} from three different regions: weak, medium, and strong. Each level is described by a normal distribution with means (μ) uniformly spread within $[F_{min}, F_{max}]$ and a fixed standard deviation of $\sigma_S = 0.6$N. At the HLC level, F_{ref} is sampled based on the predefined levels above. In addition to that, each controller randomly samples the strength level before initializing F_{ref}. For example, the Hard Goal controller is more likely to sample (μ_H) at a strong level than the Soft Goal controller (μ_S).



(a) Hard Goal (b) Soft Goal

Fig. 4: State machines for Hard and Soft Goal controllers.

The KCG mode mostly samples from the weak level. This variability makes the robot's behavior more human-like.

F. Intent Recognition

In this work, intent recognition serves both to detect conflict and to determine the direction of the human intended goal, which is the key variable that governs state transitions in HLC. The intent recognition for the human-robot setting was informed by our previous work [31] that studied the human-human interaction. In this work, instead of window intervals, instantaneous values of power signals are used. To train the model, several human-robot interaction trials were conducted in which the participant moved the object in different goal directions while the robot was in a passive admittance control mode. Details of the training procedure are described in Sect. IV-B.

IV. Implementation

A. Hardware Setup

We use a UR10e robot (Universal Robots), and a wooden tray with dimensions 61cm×31cm×10cm and weight $m = 2.1$kg. Three goal locations ($n = 3$) were chosen to resemble human-human experiments in [31], with a 40° separation angle and 0.5m away from the starting point to accommodate the reachable workspace of the manipulator. The tray is equipped with two RFT60 force-torque sensors (Robotous) connected to an onboard Raspberry Pi. Power is provided by a battery on the back of the tray. The Raspberry Pi transmits the force-torque data wirelessly to the main computer at 200Hz. The force sensor mounted between the robot end-effector and the tray was used as feedback for admittance control. To simplify the implementation, a user-side force sensor was used for intent recognition although the user force could be reliably estimated using the robot-side force sensor. To mitigate the noise in the feedback loop that causes a vibration in admittance control, the real-time Butterworth low-pass filter was implemented with a cutoff frequency of 5Hz. In order to reach the control rate of the robot, force-torque data was up-sampled to 600Hz for smooth admittance control. The Cartesian configuration of the manipulator and its velocity were also transmitted at 500Hz. The proposed system was implemented using the Robot Operating System (ROS).

To synchronize the motion between the robot and the human, three beeps are used. The first beep signals the human to grasp the object. The second beep is automatically triggered when the sensors detect that the human has grasped the tray handle. The third beep is played when the object reaches one of the goal locations.

B. Intent Recognition Training

Training data for the intent recognition consisted of 18 trials (12 for training, 6 for testing) from two participants. In each trial, the participant attempted to move the tray in a single goal direction and the robot passively followed. The distribution of goals was uniform. The intent recognition module uses only velocity, position, and force (human side) data. In addition, the projected force, power, velocity, and raw stretch force (F_{str} in the object frame) are used as features. This is a subset of Feature Set 2 in [31]. The input to the classifier was a 13-dimensional feature vector sampled at 250Hz. In turn, the classifier output was one of the 3 goal locations (g_1, g_2, g_3). The idle state (no human action) was determined by a simple threshold on the magnitude of F_H. Before training, each trial was annotated for an action phase in which a participant actively applied a force to move the object to a desired location.

We chose the Linear Discriminant Analysis (LDA) classification algorithm as our predictor [34]. The classification

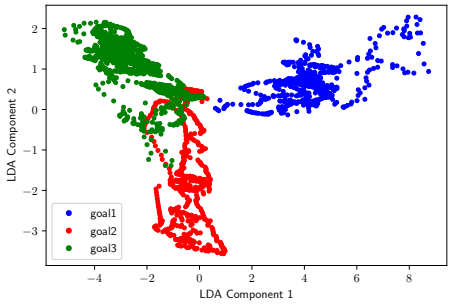


Fig. 5: Visualization of LDA Classification.

accuracy was 93. 47% for the test set. The separation of classes is shown in Fig. 5 for two principal components. Although the intent recognizer was trained for the passive mode controller, it demonstrated good generalization performance when the robot was actively applying force. Therefore, from a practical point of view, there was no need to generate new data with different robot behaviors. Studying how generalizable the trained model is across different humans is the subject of future work.

V. Human-Robot Study and Results

To validate the proposed controller, a human-robot experiment was designed, and approved by the IRB. The task is to collaboratively carry the tray to one of the 3 goal locations. Communication is restricted only to the haptic modality. Before each trial, the human participant was given a goal configuration (location and commitment) as defined in Sect. II; similarly, a goal configuration was selected for the robot. Neither the human nor the robot knew each other's goals. The experimental procedure replicates the human-human study in [31].

The experiment involved 10 healthy participants who were recruited from the University of Illinois Chicago campus. Each participant completed 24 trials, totaling 240 trials. The goal assignment for the human-robot dyad was drawn randomly but skewed, so more *soft vs. soft* conflicting interactions occurred. Before engaging in the task, the participants were given time to practice, adapt to the robot, and learn the goal types. This alleviates the learning effect. Following the experiment, the participants were asked to complete a qualitative feedback survey.

A. Controler Performance

1) Follower Controller: The number of instances where the robot assumed a follower role in the experiment was 20. In 16 instances, the controller terminated at Normal Termination (Fig. 3). This indicates that the intent recognition was correctly determining human input and transitioned to the KCG mode with a correctly recognized common goal. Out of 4 failed cases, 3 were due to humans mistakenly swapping the goal location. In one failed instance, the force applied by the human was small; the intent recognizer tends to fail when a human is not decisive. During all of the Follower controller instances, the human participant was assigned to different levels of commitment (9 Hard, 11 Soft) and actively applied force on the object. This demonstrates the generalizability and feasibility of the proposed intent-based feedback control scheme.

2) Hard Goal Controller: In 90 trials, the robot was assigned a hard goal. In 78 instances (87%), the controller behaved as expected, and the tray was delivered to the robot's intended goal. In 8 instances (9%), the controller terminated due to the safety abort. Only in 4 instances (4.5%) did the robot fail to move the object to the intended goal. There can be two reasons for such an outcome. An obvious possibility is that the participant failed to perceive the conflict. In fact, in the human-human experiment [31], the human participants

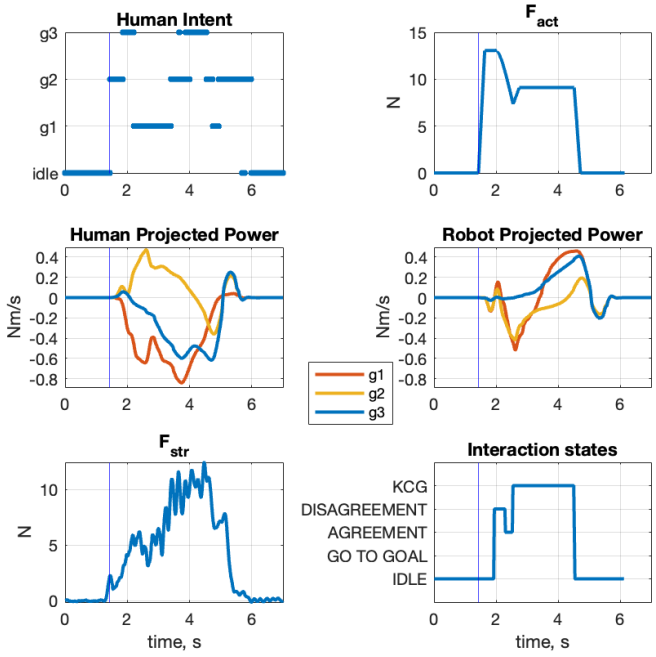


Fig. 6: An example of hard-soft case where $g_R = g_1$ and $g_H = g_2$. The blue vertical curve shows the instant of the first beep signal.

failed to perceive the hard goal intent of the partner in several instances. Another reason for a failure can be a small magnitude of F_{ref}. In this case, human participants may perceive the robot as having a soft goal, hence the participant decides to apply more effort to bring it to their assigned goal. As stated in Sect. III-E, F_{ref} can vary from trial to trial.

An example run of a hard goal controller is given in Fig. 6. The robot was assigned to go to g_1 and the human to g_2. From the projected power of the human, one can see that the human attempts to go to g_2. The intent recognition correctly detects the conflict and the controller increases F_{act}. At around 2 seconds, the human partner decides to give up and follow the robot lead; the controller appropriately transitions to the agreement phase. Shortly after, the negotiated goal settles at g_1, the controller transitions to the KCG mode, and the interaction terminates at g_1. Though not used by the controller, one can observe how the robot's intent is reflected in the robot's projected power.

3) Soft Goal Controller: Out of 240 trials, the robot had a soft goal in 130 trials and 117 of those were successful (90%), meaning that: 1) the tray ended up at either the robot's or human's goal if the human had a soft goal too; 2) the tray ended up at the robot's goal if the human had a follower role; 3) the tray ended up at the human's goal if the human had a hard goal. In the remaining 5 instances, the controller stopped due to a safety abort ($\approx$ 4%). In 2 instances ($\approx$ 2%), participants failed to reach the hard goal because they mistakenly swapped the goal location. In the remaining 6 instances ($\approx$ 5%), the dyads misunderstood the intent of the partner in *soft-soft* trials. Again, such misunderstandings also occur in human-human interactions.

We also report who prevailed in *soft-soft* trials (60 trials).

In 34 trials (57%) the tray ended up at the robot's goal and in 26 trials it ended up at the human's goal. This indicates that the robot controller was able to take the initiative and defer to the human partner.

An example run of the Soft Goal controller is given in Fig. 7. In this particular case, the intended goal of the human was g_3. In the human projected power curves, one can see a clear intent towards g_3, which is correctly captured by the intent recognizer. Also, note the magnitude of F_{str} that shows the level of the interaction conflict. In the conflicting phase, the controller correctly increases F_{act} in a disagreement mode and switches to the agreement phase as soon as the human partner decides to give up, indicated by the power curve dropping sharply in magnitude (blue curve, 2nd row, left). Since F_{str} did not cross the F^C threshold, the robot did not give up on its intent. Finally, after a short agreement phase, the dyad follows the KCG mode and completes the interaction. A video of the robot performance for all controllers is included in the supplemental material.

4) Switching Frequency: A valid concern in machine learning-based feedback controllers is fast switching due to misclassification. Therefore, we report the frequency of switches in the agreement-disagreement cycle and their average duration. In the hard goal cases (90 trials), the average durations of the agreement and disagreement states were 1.37s and 1.44s, respectively. The average number of switches was 0.8. In other words, the Hard Goal controller made one transition in 8 out of 10 cases and no transitions in the remaining 2 cases. Similarly, in the soft goal cases (130 trials), the average durations of the agreement and disagreement states were 1.6s and 0.9s, respectively. The average switching rate was 0.7, which is even less frequent than the Hard Goal controller. These results demonstrate the robustness of the proposed controller scheme against false transitions detected by the intent recognizer. Contributing factors to this are accurate tuning of timer thresholds (see Sec. III-D) and the accuracy of the intent recognizer.

B. Survey Results

In the follow-up survey, each participant was asked the following 5 questions: 1) How comfortable was the interaction? 2) How responsive was the robot? 3) Was the robot's behavior predictable? 4) If you had to perform this task of moving an object with a partner over a longer period of time, would it be acceptable to have this robot as the partner? 5) How similar was the robot's behavior to human behavior? For the responses, we used a 7-point Likert scale, with 7 being the most positive response [35]. The average values of the responses for the 10 participants are summarized in Tab. I. The results show that the participants very highly rated the interaction.

Question #	**Q1**	**Q2**	**Q3**	**Q4**	**Q5**
Average	5.8	6.2	5.4	5.3	5.2

TABLE I: Average responses of 10 participants on the Likert scale of 7.

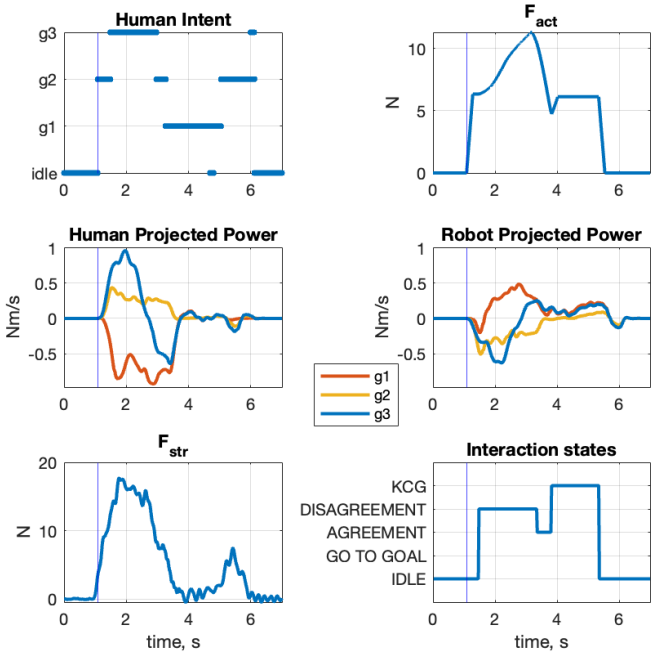


Fig. 7: An example of soft-soft case where $g_R = g_1$ and $g_H = g_3$. The blue vertical curve shows the instant of the first beep signal.

VI. Conclusion

The paper describes a hierarchical robot control framework that emulates human behavior in collaborative manipulation tasks. The framework is inspired by our study of human-human collaboration, which showed how humans use force exchanges to reach a consensus on where to move and who will lead. At the top level, the proposed controller, inspired by these human strategies, consists of finite-state machines that represent different levels of commitment to a desired goal configuration. Key to our architecture is a real-time intent recognizer that enables the robot to respond to human actions effectively. We provide insights into controller design, discuss the feature engineering process for the intent recognizer, and describe the recognizer training.

The framework was implemented on a UR10e robot and evaluated through a human study. The experimental results demonstrate the robot's ability to correctly recognize and respond to human input, effectively communicate its intent, and resolve conflicts. Success rates are reported, and comparisons with human-human experiments are made to illustrate the approach's effectiveness in emulating human collaborative behavior in robot-human interactions. A qualitative evaluation of the controller by human subjects shows that it was favorably perceived.

A natural extension of this work is to adapt the controller to a realistic setting where a human and a robot can co-manipulate various objects with different destination distributions. This will require an environment-invariant version of the presented intent recognizer. Another interesting path to explore is to expand the interaction modalities in the high-level controller with language and gestures.

REFERENCES

[1] B. Abbasi, N. Monaikul, Z. Rysbek, B. D. Eugenio, and M. Žefran, "A Multimodal Human-Robot Interaction Manager for Assistive Robots," in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Nov. 2019, pp. 6756–6762.

[2] N. Monaikul, B. Abbasi, Z. Rysbek, B. Di Eugenio, and M. Žefran, "Role Switching in Task-Oriented Multimodal Human-Robot Collaboration," in *2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, Aug. 2020, pp. 1150–1156.

[3] L. Peternel, N. Tsagarakis, and A. Ajoudani, "Towards multi-modal intention interfaces for human-robot co-manipulation," in *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2016, pp. 2663–2669.

[4] A. M. Shervedani, S. Li, N. Monaikul, B. Abbasi, B. Di Eugenio, and M. Žefran, "An end-to-end human simulator for task-oriented multimodal human-robot collaboration," in *2023 IEEE International Symposium on Robot and Human Interactive Communication*, 2023. [Online]. Available: https://arxiv.org/abs/2304.00584

[5] N. Gildert, A. G. Millard, A. Pomfret, and J. Timmis, "The Need for Combining Implicit and Explicit Communication in Cooperative Robotic Systems," *Frontiers in Robotics and AI*, vol. 5, 2018.

[6] R. Groten, D. Feth, H. Goshy, A. Peer, D. A. Kenny, and M. Buss, "Experimental analysis of dominance in haptic collaboration," in *RO-MAN 2009 - The 18th IEEE International Symposium on Robot and Human Interactive Communication*, 2009, pp. 723–729.

[7] A. Mörtl, M. Lawitzky, A. Kucukyilmaz, M. Sezgin, C. Basdogan, and S. Hirche, "The role of roles: Physical cooperation between humans and robots," *The International Journal of Robotics Research*, vol. 31, no. 13, pp. 1656–1674, Nov. 2012.

[8] Z. Rysbek, K. H. Oh, B. Abbasi, M. Žefran, and B. Di Eugenio, "Physical action primitives for collaborative decision making in human-human manipulation," in *2021 30th IEEE International Conference on Robot & Human Interactive Communication (RO-MAN)*, 2021, pp. 1319–1325.

[9] T. Flash and N. Hogan, "The coordination of arm movements: an experimentally confirmed mathematical model," *Journal of neuroscience*, vol. 5, no. 7, pp. 1688–1703, 1985.

[10] M. Kawato, "Internal models for motor control and trajectory planning," *Current Opinion in Neurobiology*, vol. 9, no. 6, pp. 718–727, Dec. 1999.

[11] E. Noohi, M. Žefran, and J. L. Patton, "A model for human–human collaborative object manipulation and its application to human–robot interaction," *IEEE transactions on robotics*, vol. 32, no. 4, pp. 880–896, 2016.

[12] V. Duchaine and C. Gosselin, "Safe, stable and intuitive control for physical human-robot interaction," in *2009 IEEE International Conference on Robotics and Automation*. IEEE, 2009, pp. 3383–3388.

[13] Y. M. Hamad, Y. Aydin, and C. Basdogan, "Adaptive human force scaling via admittance control for physical human-robot interaction," *IEEE Transactions on Haptics*, vol. 14, no. 4, pp. 750–761, 2021.

[14] Z. Al-Saadi, Y. M. Hamad, Y. Aydin, A. Kucukyilmaz, and C. Basdogan, "Resolving Conflicts During Human-Robot Co-Manipulation," in *Proceedings of the 2023 ACM/IEEE International Conference on Human-Robot Interaction*, ser. HRI '23. Stockholm Sweden: Association for Computing Machinery, Mar. 2023, pp. 243–251.

[15] S. Schaal, *Dynamic Movement Primitives -A Framework for Motor Control in Humans and Humanoid Robotics*. Tokyo: Springer Tokyo, 2006, pp. 261–280.

[16] H. Ben Amor, G. Neumann, S. Kamthe, O. Kroemer, and J. Peters, "Interaction primitives for human-robot cooperation tasks," in *2014 IEEE International Conference on Robotics and Automation (ICRA)*, May 2014, pp. 2831–2837.

[17] S. Calinon, P. Evrard, E. Gribovskaya, A. Billard, and A. Kheddar, "Learning collaborative manipulation tasks by demonstration using a haptic interface," in *International Conference on Advanced Robotics, 2009. ICAR 2009*, Jun. 2009, pp. 1–6.

[18] L. Rozo, D. Bruno, S. Calinon, and D. G. Caldwell, "Learning optimal controllers in human-robot cooperative transportation tasks with position and force constraints," in *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2015, pp. 1024–1030.

[19] N. Hogan, "Impedance control: An approach to manipulation," in *1984 American control conference*. IEEE, 1984, pp. 304–313.

[20] M. M. Rahman, R. Ikeura, and K. Mizutani, "Control characteristics of two humans in cooperative task," in *Smc 2000 conference proceedings. 2000 ieee international conference on systems, man and cybernetics. 'cybernetics evolving to systems, humans, organizations, and their complex interactions' (cat. no.0*, vol. 2, 2000, pp. 1301–1306 vol.2.

[21] K. Kosuge and N. Kazamura, "Control of a robot handling an object in cooperation with a human," in *Proceedings 6th IEEE International Workshop on Robot and Human Communication. RO-MAN'97 SENDAI*, 1997, pp. 142–147.

[22] H. A. Varol, F. Sup, and M. Goldfarb, "Multiclass real-time intent recognition of a powered lower limb prosthesis," *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 3, pp. 542–551, 2010.

[23] M. Jacobson, P. Kantharaju, H. Jeong, J.-K. Ryu, J.-J. Park, H.-J. Chung, and M. Kim, "Foot contact forces can be used to personalize a wearable robot during human walking," *Scientific Reports*, vol. 12, no. 1, p. 10947, Jun. 2022.

[24] D. J. Agravante, A. Cherubini, A. Bussy, P. Gergondet, and A. Kheddar, "Collaborative human-humanoid carrying using vision and haptic sensing," in *2014 IEEE International Conference on Robotics and Automation (ICRA)*, 2014, pp. 607–612.

[25] A. Bussy, A. Kheddar, A. Crosnier, and F. Keith, "Human-humanoid haptic joint object transportation case study," in *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2012, pp. 3633–3638.

[26] S. Jain and B. Argall, "Probabilistic Human Intent Recognition for Shared Autonomy in Assistive Robotics," *ACM Transactions on Human-Robot Interaction*, vol. 9, no. 1, pp. 2:1–2:23, Dec. 2019.

[27] Z. Wang, K. Mülling, M. P. Deisenroth, H. Ben Amor, D. Vogt, B. Schölkopf, and J. Peters, "Probabilistic movement modeling for intention inference in human–robot interaction," *The International Journal of Robotics Research*, vol. 32, no. 7, pp. 841–858, Jun. 2013.

[28] K. Mojtahedi, B. Whitsell, P. Artemiadis, and M. Santello, "Communication and Inference of Intended Movement Direction during Human–Human Physical Interaction," *Frontiers in Neurorobotics*, vol. 11, 2017.

[29] C. Schultz, S. Gaurav, M. Monfort, L. Zhang, and B. D. Ziebart, "Goal-predictive robotic teleoperation from noisy sensors," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*, 2017, pp. 5377–5383.

[30] G. Kang, H. S. Oh, J. K. Seo, U. Kim, and H. R. Choi, "Variable admittance control of robot manipulators based on human intention," *IEEE/ASME Transactions on Mechatronics*, vol. 24, no. 3, pp. 1023–1032, 2019.

[31] Z. Rysbek, K.-H. Oh, and M. Zefran, "Recognizing intent in collaborative manipulation," in *Proceedings of the 25th International Conference on Multimodal Interaction*, ser. ICMI '23. New York, NY, USA: Association for Computing Machinery, 2023, p. 498–506. [Online]. Available: https://doi.org/10.1145/3577190.3614174

[32] A. Q. Keemink, H. van der Kooij, and A. H. Stienen, "Admittance control for physical human–robot interaction," *The International Journal of Robotics Research*, vol. 37, no. 11, pp. 1421–1444, 2018.

[33] C. Murchison and H. Banister, *A Handbook of General Experimental Psychology*. Clark University Press, 1934. [Online]. Available: https://books.google.com/books?id=aEUczQEACAAJ

[34] T. Hastie, R. Tibshirani, J. H. Friedman, and J. H. Friedman, *The elements of statistical learning: data mining, inference, and prediction*. Springer, 2009, vol. 2.

[35] M. Schrum, M. Ghuy, E. Hedlund-botti, M. Natarajan, M. Johnson, and M. Gombolay, "Concerning Trends in Likert Scale Usage in Human-robot Interaction: Towards Improving Best Practices," *ACM Transactions on Human-Robot Interaction*, vol. 12, no. 3, pp. 33:1–33:32, Apr. 2023.