Physical Action Primitives for Collaborative Decision Making in Human-Human Manipulation

Zhanibek Rysbek[1], Ki Hwan Oh[1], Bahareh Abbasi[2], Miloš Žefran[1], and Barbara Di Eugenio[3]

***Abstract*— Human-human collaboration is characterized by a back-and-forth, where an action of one agent elicits the response of the other. This interaction is inherently multimodal and includes both high-level modalities such as language and low-level ones such as force exchanges. In this work, we investigate human collaborative manipulation: we show distinct patterns that can be identified in low-level physical data and that can be interpreted as primitives used by humans to negotiate about various aspects of the motion and to execute the motion. These primitives provide a high-level interpretation of the interaction and can be used to connect low-level behavior to language. We describe the human study used to collect the data, the data analysis process, and discuss how the identified primitives could be used by a robot's interaction manager to mediate physical Human-Robot Interaction (pHRI).**

I. Introduction

When humans collaborate on a task, they use a multitude of modalities to communicate and coordinate their actions. Perhaps the most obvious communication modality is language; it operates at a highly abstract (symbolic) level and is mostly unique to humans. On the other hand, physical collaboration typically involves low-level signals such as forces (wrenches) and velocities (twists). Much of the literature on physical Human-Robot Interaction (pHRI) focuses on these low-level signals and how to control them.

The ultimate goal of this research is to bridge the gap between the low-level signal domain and the high-level symbolic domain to allow pHRI to approach the richness of human-human collaboration. In particular, we want pHRI to mimic the back-and-forth typical of human collaboration, akin to taking turns in a dialogue. Such back-and-forth can be observed when a spoken utterance is meant to modify physical motion (like *Slow down*), or when the partners negotiate what strategy to use for a task. Ultimately, our goal is to implement an interaction manager that would allow the robot to respond to a multimodal input from the human partner with its own multimodal action. We build on our work in [1] but focus specifically on how a robot can respond at a symbolic level to low-level signals from the human user by switching between discrete low-level behaviors (primitives).

[1]Z. Rysbek, K.H. Oh and Miloš Žefran are with the Robotics Lab, Department of Electrical and Computer Engineering, the University of Illinois at Chicago, Chicago, IL 60607, USA.

[2]B. Abbasi is with the Department of Computer Science, California State University Channel Islands, Camarillo, CA 93012, USA.

[3]B. Di Eugenio is with the Natural Language Processing Lab, Department of Computer Science, the University of Illinois at Chicago, Chicago, IL 60607, USA.

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

In this paper, we study a human-human collaborative manipulation task. In contrast to our previous work [2], humans engage in deliberative rather than ballistic motion. In particular, they need to negotiate how to avoid an obstacle. We describe the experimental setup and the human subject study that we conducted to collect the data. The main contribution of our work is a data analysis that shows that negotiation during deliberative collaborative manipulation tasks involves discrete behaviors (primitives). We also present the framework for extracting the primitives from the collected data and show how they enable the back-and-forth in physical interaction. The important implication of our work is that such primitives can be subsequently used to implement collaborative negotiation in pHRI. To the best of our knowledge, we are the first to view physical interaction as dialog and to identify the primitives that are used to build such dialog. When adopted for pHRI, viewing the interaction in this way lets the robot play an active role, allowing it both to respond to the human and to initiate an action when necessary at a fast time scale. This contrasts with existing approaches where the robot mainly executes a previously learned behavior in an open-loop fashion and is unable to dynamically react to human action or take initiative.

The rest of the paper is organized as follows. Section II reviews the existing related literature and Section III describes our human study. In Section IV we introduce our proposed framework for extracting the interaction primitives, while Section V describes the analysis of the collected data. Finally, Section VI concludes the paper.

II. Related Work

One of the leading paradigms for pHRI is programming by demonstration (PbD) [3], [4], often in conjunction with impedance or admittance control [5], [6]. When it comes to manipulation and handover, several works [7]–[9] build on the seminal work in [10] where a minimum-jerk model is proposed for ballistic reaching motions. In [11], the model is used to compute interaction forces during collaborative manipulation involving ballistic motions. Building on that work, [12] describes a set of haptic features that are used to accurately classify the interaction into conflicting or harmonious, both for rotational and translational motions. A probabilistic framework for learning task-level interaction primitives from human demonstrations is proposed in [4]. The approach does not provide a mechanism for making decisions during the execution of the primitive. A method for learning robot impedance parameters that can be used to define a virtual attractor for the end-effector is proposed in [5]. Similar to the above, the approach assumes an open-loop execution of the motion, in contrast to what is observed when the agents need to negotiate during the task. Several researchers studied how the roles of the leader and the follower are assigned during collaborative manipulation. In [13], the authors study how humans adapt to their partners. They show that while humans initially rely on both visual and haptic feedback to coordinate with the collaborator, as the subjects get used to each other, they start relying on perceived stiffness alone. In [14], a theoretical framework is proposed to model how the collaborating humans switch between the leader and the follower roles. And finally, [15] shows that dyads modulate the forces applied on the object they manipulate to generate a haptic information channel; this allows them to match their performance during the joint manipulation task to that of a single individual.

III. Human Study

A. Experimental Setup

In order to study human-human collaborative decision-making, an experiment was designed in which two people were asked to collaborate and relocate a tray from one table (table A) to another (table B) while avoiding the obstacle located between the two. The layout of the environment is shown in Fig. 1. The distance between the tables was 2.6 m, which typically resulted in subjects requiring approximately 9 seconds to carry the tray from the start to the goal location. We placed a water bottle on the tray and instructed the subjects to prevent it from toppling to force them to coordinate their actions. Subjects were explicitly instructed to communicate only through physical actions which means that talking or communicating with gestures was not allowed. As can be seen in Fig. 1, subjects had the freedom to either circumnavigate the obstacle to the left or the right. The experimental task thus requires the subjects to engage in physical negotiation at several points during the motion and allows us to observe how such negotiation/collaborative decision-making is achieved through physical actions.

Several sensors were used to capture the interaction between humans during the collaborative task. The force data were collected by two RFT60 force-torque sensors (Robotous Inc.), each installed between the handle and the tray (depicted in Fig. 2 and 3). The force sensors were sampled at 1KHz and interfaced with a desktop PC using serial communication. At the center of the tray, a Raspberry Pi with an embedded 9-DOF IMU sensor (LSM9DS1 from AdaFruit) was attached to track the configuration (position and orientation) of the tray. The IMU data was transmitted via a local wireless network to the desktop PC. An online position tracking algorithm was implemented using the ArUco library [16]. For this purpose, we affixed fiducial ArUco markers to the surface and the sides of the tray (see Fig. 2). Three USB cameras were placed in a triangular configuration to record the movement of the subjects and the tray, allowing us to deal with occlusions of the markers. The weight of the tray was $2.3kg$, and its dimensions were $61cm \times 31cm$. The data

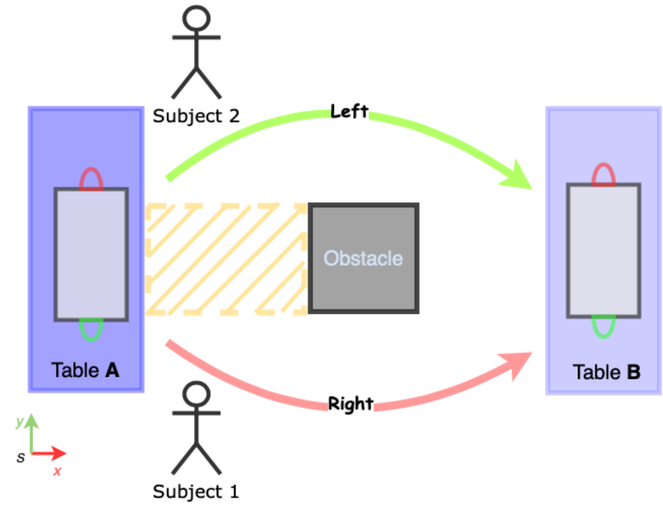


Fig. 1. Schematic of the environment. Subjects had to go around the obstacle either on the right or the left side.

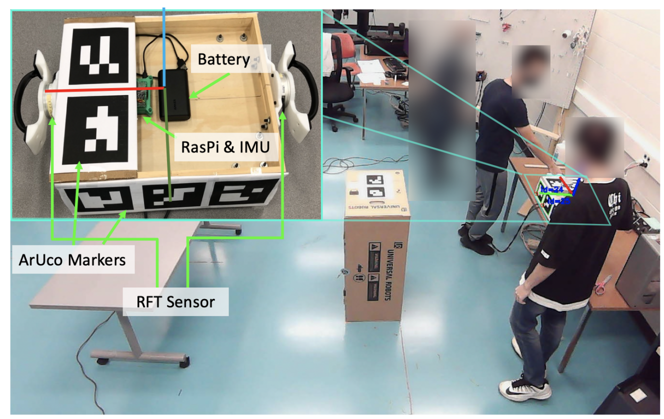


Fig. 2. Experimental setup and the tray used for this study.

collection was implemented in the Robot Operating System (ROS) environment.

The collected force-torque and IMU signals were pre-processed using a low-pass filter with the cutoff frequency of $12.5Hz$. The position data is obtained from three different cameras and processed with the robust-Lowess smoothing filter [17]. Then, all signals are re-sampled to 100Hz. As an alternative to ArUco pose estimation, the IMU data was used to track the orientation of the tray and processed using the Kalman filter [18]. The body and spatial twists were computed by differentiating the pose of the object over time. All the signal processing was performed in MATLAB.

B. Data Analysis

Four subjects were recruited for this experiment and were asked to form five dyads, resulting in 112 trials.[1] In each trial, the dyads chose between the *left* path and the *right* path (Fig. 1). The distribution of the right/left trajectories is 49/51%. In addition to deciding between the left/right motions, the

[1]The experiment design and data collection was reviewed and approved by the Institutional Review Board (IRB).

subjects were also negotiating whether the tray is placed at the goal configuration in the same orientation as at the beginning, or it is rotated by 180°. We call these movements parallel and serial motions, respectively.

The negotiation on whether to perform the parallel or the serial motion can take place at any time during the motion. Therefore, to simplify the problem, we decided to focus on the right/left decision-making process, as it consistently happens at the beginning of the trajectory. Moreover, to narrow down the scope of the analysis to the period when subjects actually negotiate whether to move left or right, we have annotated the time when subjects leave the yellow shaded area between table A and the obstacle (Fig. 1).

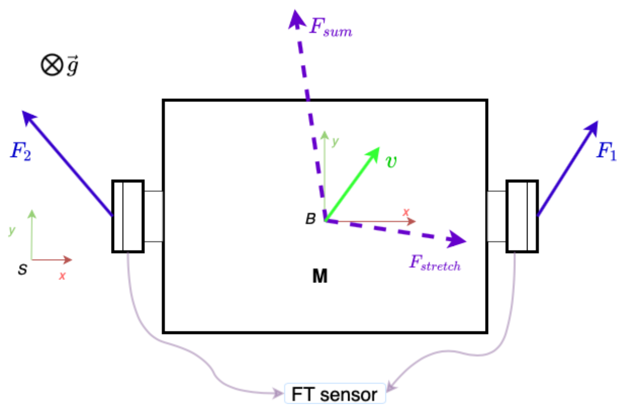


Fig. 3. Free body diagram from top view.

IV. Framework

A. Background

In a collaborative manipulation task, each subject applies a force to the manipulated object. In our experiment, F_1 and F_2 are the forces applied to the handles of the tray by subject 1 and subject 2 respectively (shown in Fig. 3). The resulting force, $F_{sum} = F_1 + F_2$, contributes to the linear movement of the tray and is directly proportional to its acceleration (Newton's Second law). The equation of motion for the center of mass, whose coordinates are described by a vector x, is:

$$m\ddot{x} = F_1 + F_2 \tag{1}$$

where m is the mass of the tray.

It is widely accepted that during collaborative manipulation humans communicate via the interaction force [11], [15]. This force does not impact the motion of the manipulated object, but it could stretch or compress it. Formally, we can write:

$$\begin{cases} F_1 = F_1^* + F_i \\ F_2 = F_2^* - F_i \end{cases} \tag{2}$$

where, F_1^* and F_2^* are the effective forces that affect the motion of the object, and F_i is the interaction force. The above equation is underdetermined so the interaction force can not be determined without additional assumptions. Several alternatives have been proposed in the literature [11], [19], [20].

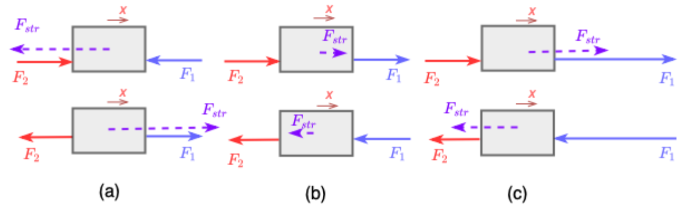


Fig. 4. Possible tension-compression scenarios in $F_{stretch}$. x-axis in red is body affixed frame.

Forces orthogonal to F_{sum} are one component of the interaction force and are one way to communicate between the dyads. Further, in this work, we exploit the *stretching force* $F_{stretch}$ which is defined as:

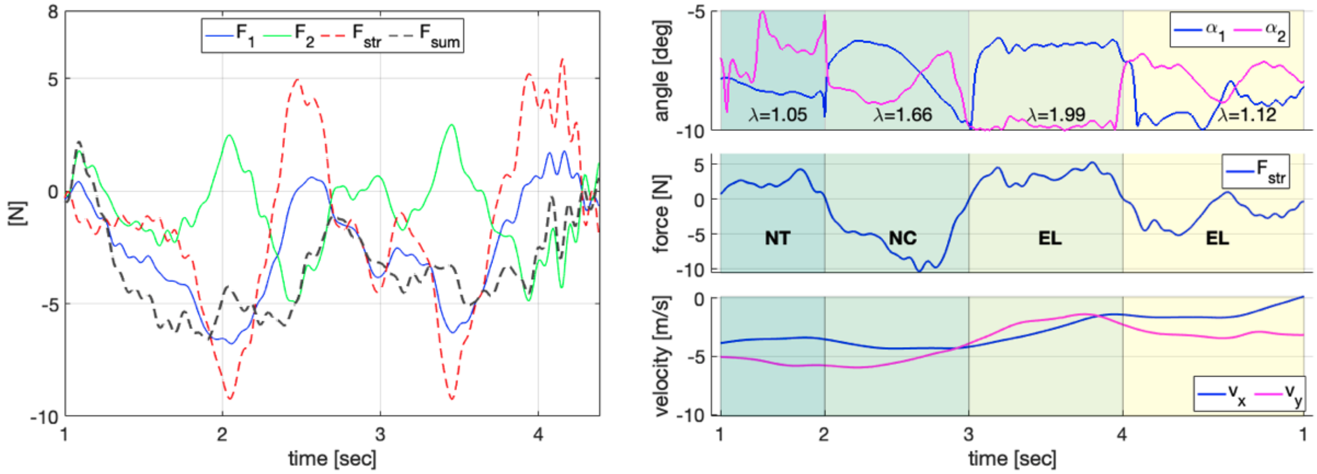
$$F_{stretch} = F_1 - F_2 \tag{3}$$

During the interaction, the total applied forces are usually significantly larger than F_{sum} (i.e. $|F_{sum}| << |F_1| + |F_2|$) [21]. This happens due to negotiation between the agents: excessive components of applied forces cancel each other and yield F_{sum} that is smaller in magnitude (refer to Fig. 5a). Meanwhile, $F_{stretch}$ magnifies the excessive components of the applied forces and cancels the components contributing to the motion. Therefore, $F_{stretch}$ acts as an indicator for determining the state of the interaction. This is in agreement with the finding in [15]. Note that, in this work, we focus on the $F_{stretch}$ along the x-direction in the body affixed frame that corresponds to the axis parallel to the handles. It is trivial to conclude that $F_{stretch}$ in the y-direction contributes to the rotational motion (refer to Fig. 3).

Fig. 4 depicts the possible outcomes of $F_{stretch}$ values. The applied forces pointing against each other result in high values of *tension* (aligned with x-axis) or *compression* (opposite to x-axis) (Fig. 4a). This could be an obvious sign of a conflicting situation. In contrast, smaller values of $F_{stretch}$ were observed when the dyad's applied forces are aligned causing it to bounce between tension/compression (Fig. 4b). This can be translated into agreement or pause actions by both subjects. A situation similar to the case in Fig. 4a may also occur when one of the subjects aggressively dominates the motion (Fig. 4c). However, in typical cooperative collaboration, this rarely happens. The presence of such alternation in the values of the stretching force is the indication of back-and-forth physical interaction.

These different outcomes of tension/compression scenarios are reflected as "hump"-like shapes in the $F_{stretch}$ signal (see Fig. 5a). In general, these humps (peaks) could happen due to several reasons:

- Conflicting scenario with clear disagreement between the subjects. In this case, the tray stays almost at rest.
- Significant difference between the velocity of each subject which makes the tray move.



(a) (b)

Fig. 5. *(a):* F_1, F_2 applied forces and $F_{stretch}$, F_{sum} along x direction in *body frame*; *(b):* Detected Primitives (four shaded areas) in force-velocity signals. Top plot: α_1 and α_2 correspond to an angle between $\angle(F_1, v)$ and $\angle(F_2, v)$ respectively; λ - leader-follower consistency measure. Middle plot: $F_{stretch}$. Bottom plot: velocity in x, y directions.

- Haptic-interactive negotiation. Subjects engage in probing actions trying to reach an agreement on the strategy.

Therefore, relying only on the stretching force leaves ambiguity in determining the correct interaction state. An additional modality, such as velocity, helps to infer the accurate interaction state during these peaks.

B. Event Detection Algorithm

We hypothesize that the humps in $F_{stretch}$ (events) correspond to the atomic parts of physical interaction. Depending on the context, each of them could represent the symbolic meaning behind the action of the agents. To interpret the data, we implemented an algorithm that identifies the beginning and end of each of these events. Furthermore, we argue that these detected events describe the back and forth physical interaction akin to dialog between humans.

The inputs to our event detection algorithm (inspired by [22]) are $F_{stretch}$, α_1, α_2 (Fig. 5b); the outputs are the start and the end time of the events. First, a possible set of start/end times is computed considering: zero crossings of $F_{stretch}$; local minima or maxima of $F_{stretch}$ (extrema); and zero crossings of $\alpha_1 - \alpha_2$. Then, a shortlist of event candidates is formed by disregarding negligible extrema and combining adjacent start/end times. Given that the fastest simple human reaction time is approximately 0.25 secs [23], we impose the additional constraint that no event can be shorter than 0.4 sec. Thus, shorter events are merged to satisfy the constraint. Moreover, outputs of our event detection algorithm are manually checked and start/end times are adjusted as needed.

V. Results and Discussion

From 112 trials, 283 instances of events were detected. We employed an unsupervised learning algorithm, K-medoids [24], to find the clusters corresponding to the interaction primitives. As the similarity measure, we used the dynamic time warping algorithm (DTW) [25] for both $F_{stretch}$ and v_y signals (in spatial frame). DTW is well-suited for our application since it can measure the similarity between two signals regardless of their duration. The relative distance between the primitives is computed according to:

$$E(P_i, P_j) = \frac{DTW(F^i_{stretch}, F^j_{stretch})}{F^{max}_{stretch}} + \alpha \frac{DTW(v^i_y, v^j_y)}{v^{max}_y} \quad (4)$$

where, α is a blending coefficient and $F^{max}_{stretch}$ and v^{max}_y are normalization factors (the maximum observed value for $F_{stretch}$ and v_y in all trials).

The clustering process consists of two stages. We first separate the *execution primitives* into one subset by applying a simple threshold on mean, min, and max value of v_y, and the sequence number of the primitive. The execution primitives always have to conclude the interaction as the subjects leave the area of interest (yellow area on Fig. 1). After that, we apply the K-medoids algorithm to each subset of the data. This is done to improve the clustering results, as many clusters overlap if K-medoids is applied to all the data. As a result, we obtain 6 different clusters. Fig. 6 represents the 2D-histogram (heatmap), where each cluster is represented in the $F_{stretch}$ and v_y plane. These 6 clusters represent different stages of interaction in the task that involves collaborative decision making: *Negotiation with Tension* (NT) – high tension in $F_{stretch}$, v_y distribution centered around the horizontal line; *Negotiation with Compression* (NC) – high compression in $F_{stretch}$, v_y distribution centered around a horizontal line; *Decision Making to the Left* (DL) – high tension, but v_y sharply increases; *Decision Making to the Right* (DR) – high tension, but v_y sharply decreases; *Execution to the Left* (EL) – tension and compression combined, v_y sustains higher positive velocity; *Execution to the Right* (ER) – tension and compression combined, v_y sustains higher negative velocity.

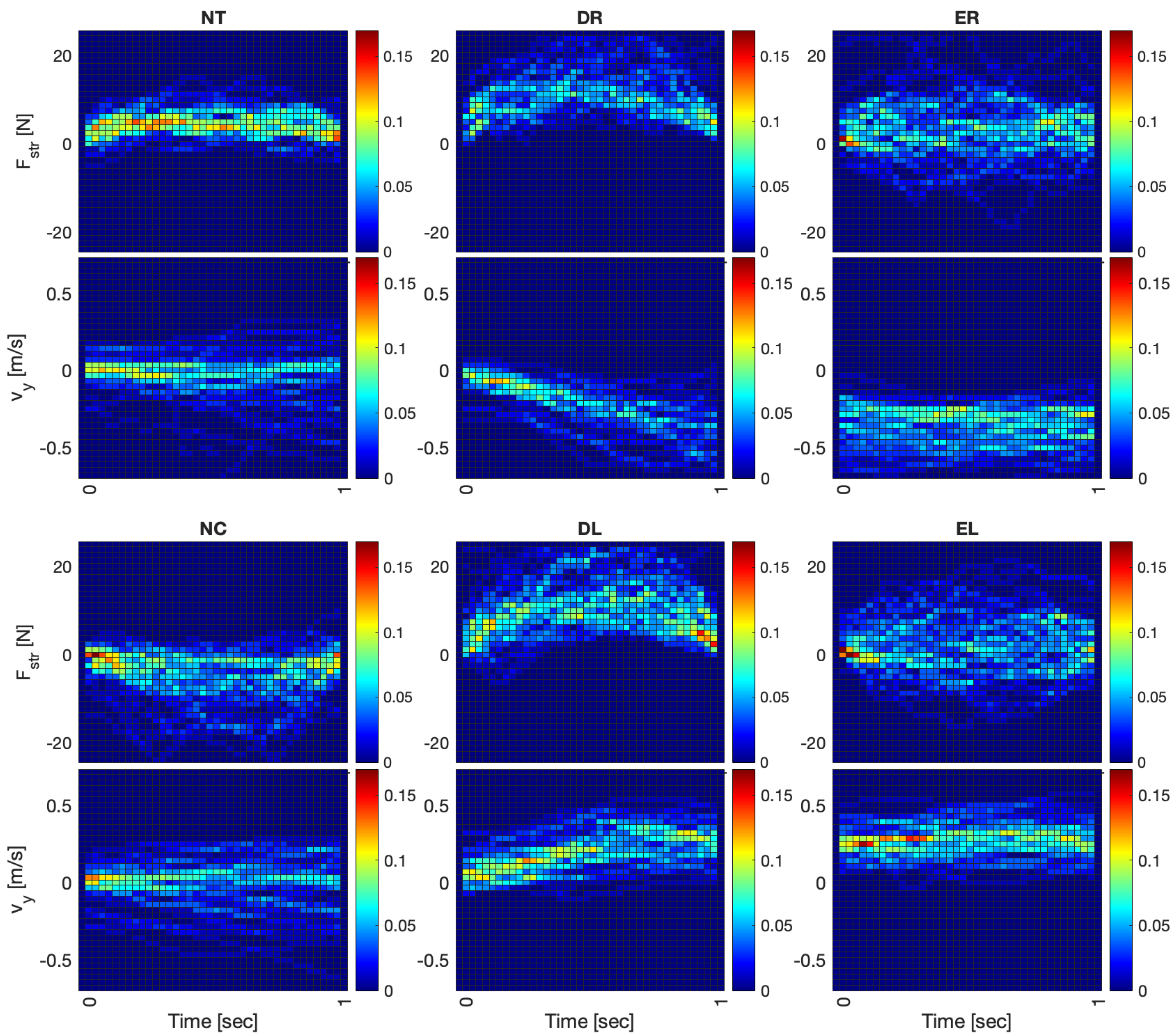


Fig. 6. Clusters grouped by $F_{stretch}$ and v_y shown in heatmaps. The color of each bin represents the percentage of the samples observed in that bin. *Note:* $F_{stretch}$ outliers are re-scaled to the [-20N, 20N] range for visualization purposes.

An example of a sequence of primitives in a trajectory is shown in Fig. 5b (middle plot).

Each collaborative interaction can be translated into a sequence of these primitives. We can learn these sequences using a state transition diagram. Fig. 7 shows the back-and-forth negotiation behaviors (i.e. primitives) that happened during the decision-making process in our collected data. As depicted in the figure, all the trajectories start from the "Start" state and end at the "End" state, where a final decision is made. Primitives that were the first in the interaction sequence are directly connected to "Start"; similarly, primitives that were the last in the sequence connect to the "End" state. Nearly half of the trajectory sequences start from NT or NC. This portion of interactions corresponds to the negotiation phase, where subjects exchange pull/push actions. In this phase, the decision is still unknown: the probability of *left/right* agreement is comparable (see transitions from NT $\rightarrow$ {DL, ER, EL} and NC $\rightarrow$ {DL, ER, EL}). The other half of the trajectories immediately start with decision-making primitives (Start $\rightarrow$ {DL, DR}), skipping physical negotiation sequences (NT, NC). Whenever they reach this state, the proposed *left/right* strategy does not change. This can be explained by the fact that the subjects guess each other's intent by other clues (such as eye contact, body posture) or one of the subjects takes on a follower role so that they do not take initiative. Moreover, these clusters could have negotiation patterns with pauses in acceleration but still moving in the proposed direction; however, they are too short to be detected. While our experiment design does not consider learning effects, we note that when dyads have multiple options to choose from, they start with a low success rate of guessing intent, but it increases to 83% after a short amount of practice [13].

Moreover, we argue that similar dynamic transitions cannot be observed in a ballistic type of motion, where the dyads are constrained to make the decision prior to the motion.

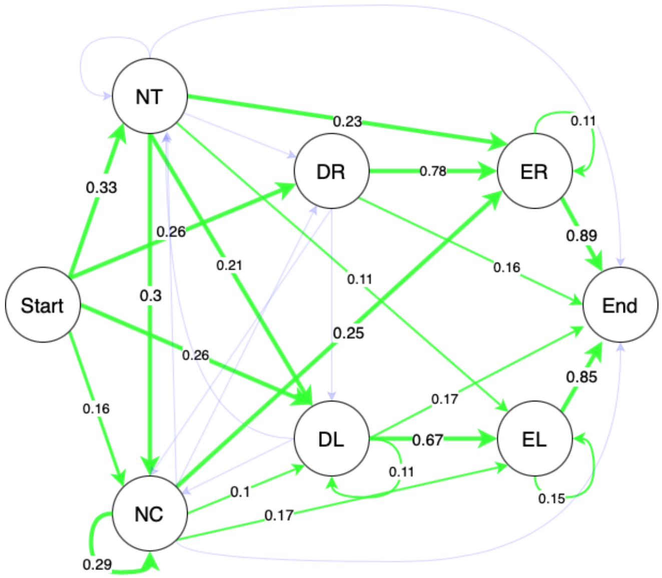


Fig. 7. Equivalent state machine from primitive sequences. Unlabelled lines have probability less than 0.09.

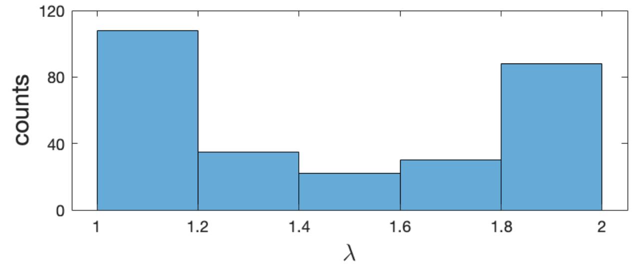


Fig. 8. Distribution of leader-follower consistency measure λ across all primitives.

In that case, PbD methods [3], [4] perform well, as the distribution of the trajectories remains relatively constant. In our approach, different kinds of initiative taken by the dyad members are observed in the primitive sequence which means that the interaction is much richer. Our approach thus allows the agents to exhibit similar behaviors.

During the proposed primitives, subjects often engage in a common maneuver, in which mostly one of them tends to be the consistent leader in the context, except for pause or conflicting states that are inherently leader-free. We introduce a leader-follower consistency measure, defined by:

$$\lambda = 1 + \frac{T(\alpha_1(t), \alpha_2(t))}{t^f_{prim} - t^0_{prim}} \tag{5}$$

where, $T(\alpha_1(t), \alpha_2(t)) \leq t^f_{prim} - t^0_{prim}$ is the duration when $\alpha_1(t) \geq \alpha_2(t)$ for $t \in [t^0_{prim}, t^f_{prim}]$; t^0_{prim}, t^f_{prim} are the start and end time of the primitive. If $\lambda = 1.5$ no distinctive leader is observed; the closer the value is to 1, subject 1 is leading and vice versa. From the distribution of λ (Fig. 8) one can see that either subject 1 or subject 2 is in a leader most of the time. The intermediate values of λ capture pause or conflicting situations where no particular role could be assigned. Analogous metrics exist in the natural language literature [26]; they measure how collaborative the dyads are when solving problems together (Knowledge Co-Construction dialogs).

To clarify, consider the top plot in Fig. 5b. One can notice that in the first primitive, subject 1 started to go to the right, hence he is the leader in that context ($\lambda \approx 1$). But subject 2 intervened and led the object to the left side in two stages: opposing the initially proposed right direction in the 2nd primitive ($\lambda \approx 1.5$ – conflicting case); increasing the velocity in the 3rd primitive to which subject 1 agreed ($\lambda \approx 2$). Finally, they finish the negotiation by executing the decision (4th primitive).

VI. Conclusions and Future Work

Given a relatively simple deliberative collaborative manipulation task, we observed that two people are capable to negotiate about many aspects of the task such as the direction of motion around an obstacle, the type of rotation of the object during the motion, and the height and the speed of the manipulated object. To simplify the analysis, we narrowed the part of the motion we analyzed to the initial phase of the object transfer when the two subjects make a decision on whether to avoid the obstacle on the left or the right. We propose a framework to identify the interaction primitives during this part of the task. Specifically, we show that the stretching force along the manipulated object provides reliable information about the interaction state. Through clustering, we identify a meaningful set of primitives that describe how the subjects make a decision. Using these interaction primitives, we argue that an interaction manager can be designed for a physical human-robot interaction task that can deal with the physical and linguistic communication of the agents and that extends our previous work in [1].

Unlike the prevailing programming by demonstration approaches that often consider the entire trajectory as a single primitive, we envision to create a more fine-grained controller for the collaborative robot that enables it to negotiate with the human partner and allows it to both respond to the human multimodal inputs as well as take the initiative and act. An interesting extension to this work would be to devise a real-time event detection algorithm and verify the proposed scheme on a more complex collaborative manipulation task. We also plan to implement our proposed framework on a real robot using impedance control [27] and show its effectiveness through a more extensive human study.

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] E. Noohi, M. Zefran, 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, Aug. 2016.

[3] 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.

[4] 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.

[5] L. D. Rozo, S. Calinon, D. Caldwell, P. Jimenez, and C. Torras, "Learning Collaborative Impedance-Based Robot Behaviors," in *Twenty-Seventh AAAI Conference on Artificial Intelligence*, Jun. 2013. [Online]. Available: http://www.aaai.org/ocs/index.php/AAAI/AAAI13/paper/view/6243

[6] 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. [Online]. Available: http://ijr.sagepub.com.proxy.cc.uic.edu/content/31/13/1656

[7] B. Corteville, E. Aertbelien, H. Bruyninckx, J. De Schutter, and H. Van Brussel, "Human-inspired robot assistant for fast point-to-point movements," in *2007 IEEE International Conference on Robotics and Automation*, Apr. 2007, pp. 3639–3644.

[8] M. Huber, M. Rickert, A. Knoll, T. Brandt, and S. Glasauer, "Human-robot interaction in handing-over tasks," in *Robot and Human Interactive Communication, 2008. RO-MAN 2008. The 17th IEEE International Symposium On*. IEEE, 2008, pp. 107–112. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4600651

[9] K. W. Strabala, M. K. Lee, A. D. Dragan, J. L. Forlizzi, S. Srinivasa, M. Cakmak, and V. Micelli, "Towards seamless human-robot handovers," *Journal of Human-Robot Interaction*, vol. 2, no. 1, pp. 112–132, 2013. [Online]. Available: http://hri-journal.org/index.php/HRI/article/view/114

[10] 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.

[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] Z. Al-saadi, D. Sirintuna, A. Kucukyilmaz, and C. Basdogan, "A Novel Haptic Feature Set for the Classification of Interactive Motor Behaviors in Collaborative Object Transfer," *IEEE Transactions on Haptics*, pp. 1–1, 2020.

[13] 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, Apr. 2017.

[14] P. Evrard and A. Kheddar, "Homotopy-based controller for physical human-robot interaction," in *RO-MAN 2009 - The 18th IEEE International Symposium on Robot and Human Interactive Communication*, Sep. 2009, pp. 1–6.

[15] R. Van der Wel, G. Knoblich, and N. Sebanz, "Let the force be with us: Dyads exploit haptic coupling for coordination," *Journal of experimental psychology. Human perception and performance*, vol. 37, pp. 1420–31, 03 2011.

[16] F. J. Romero-Ramirez, R. Muñoz-Salinas, and R. Medina-Carnicer, "Speeded up detection of squared fiducial markers," *Image and Vision Computing*, vol. 76, pp. 38–47, 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0262885618300799

[17] W. S. Cleveland, "Robust locally weighted regression and smoothing scatterplots," *Journal of the American Statistical Association*, vol. 74, no. 368, pp. 829–836, 1979. [Online]. Available: https://www.tandfonline.com/doi/abs/10.1080/01621459.1979.10481038

[18] D. Roetenberg, H. J. Luinge, C. T. M. Baten, and P. H. Veltink, "Compensation of magnetic disturbances improves inertial and magnetic sensing of human body segment orientation," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 13, no. 3, pp. 395–405, 2005.

[19] 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.

[20] 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.

[21] E. Noohi and M. Zefran, "Quantitative measures of cooperation for a dyadic physical interaction task," in *Humanoid Robots (Humanoids), 2014 14th IEEE-RAS International Conference on*. IEEE, 2014, pp. 469–474.

[22] B. Abbasi, E. Noohi, S. Parastegari, and M. Žefran, "Grasp taxonomy based on force distribution," in *Robot and Human Interactive Communication (RO-MAN)*. IEEE, 2016, pp. 1098–1103.

[23] D. L. Woods, J. M. Wyma, E. W. Yund, T. J. Herron, and B. Reed, "Factors influencing the latency of simple reaction time," *Frontiers in Human Neuroscience*, vol. 9, Mar. 2015.

[24] L. Kaufman and P. J. Rousseeuw, *Finding groups in data: an introduction to cluster analysis*. John Wiley & Sons, 2009, vol. 344.

[25] H. Sakoe and S. Chiba, "Dynamic programming algorithm optimization for spoken word recognition," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 26, no. 1, pp. 43–49, 1978.

[26] C. Howard, B. Di Eugenio, P. Jordan, and S. Katz, "Exploring Initiative as a Signal of Knowledge Co-Construction During Collaborative Problem Solving," *Cognitive Science*, vol. 41, no. 6, pp. 1422–1449, Aug. 2017.

[27] N. Hogan, "Impedance Control: An Approach to Manipulation: Part I—Theory," *Journal of Dynamic Systems, Measurement, and Control*, vol. 107, no. 1, pp. 1–7, 03 1985. [Online]. Available: https://doi.org/10.1115/1.3140702