

The HRC Model Set for Human-Robot Collaboration Research

Sofya Zeylikman¹, Sarah Widder², Alessandro Roncone², Olivier Mangin² and Brian Scassellati²

Abstract—In this paper, we present a model set for designing human-robot collaboration (HRC) experiments. It targets a common scenario in HRC, which is the collaborative assembly of furniture, and it consists of a combination of standard components and custom designs. With this work, we aim at reducing the amount of work required to set up and reproduce HRC experiments, and we provide a unified framework to facilitate the comparison and integration of contributions to the field. The model set is designed to be modular, extendable, and easy to distribute. Importantly, it covers the majority of relevant research in HRC, and it allows tuning of a number of experimental variables that are particularly valuable to the field. Additionally, we provide a set of software libraries for perception, control and interaction, with the goal of encouraging other researchers to proactively contribute to our work.

I. INTRODUCTION

Robotics is a relatively recent experimental science. As with most scientific disciplines, it needs to provide guarantees for reproducibility and replicability. However, as young as robotics is, it is co-evolving alongside the same enabling technologies that allow research in the field. This can hamper the stability of the experimental tools and hence impedes the reproducibility of robotics experiments. The problem is particularly relevant to human-robot collaboration (HRC), an important and growing area within robotics research. HRC lacks standards for replicability, a factor that would accelerate progress and facilitate integration of advancements from independent contributions in the field. This is primarily due to the fact that HRC targets a broad range of topics, from rescue robots to collaborative manufacturing and service robotics. Defining what the term ‘*collaboration*’ entails is actually still an open question and an integral part of the research in this domain. Importantly, as of now, there exists no reference collaboration setup or standardized collaborative task in the field. This paper explicitly targets this problem, with the goal of promoting ease of reproduction, evaluation, and integration of HRC research.

We identify a number of critical points related to the problem of replicability and standardization of research in HRC—most of which apply to the broader robotics field as well. Firstly, research teams use a variety of robotic platforms that, for the most part, lack standardization. Secondly, experiment design is affected by limitations in the enabling technologies, e.g. inadequate perception/manipulation skills.

This work has been supported by a subcontract with Rensselaer Polytechnic Institute (RPI) by the Office of Naval Research, under Science and Technology: Communication for enhancing Human-Robot Collaboration.

¹ S. Z. is with the Center for Engineering, Innovation and Design, Yale University, New Haven, CT 06511, USA.

² These authors are with the Social Robotics Lab, Computer Science Department, Yale University, New Haven, CT 06511, USA.



Fig. 1. Human participant engaged in the joint construction of a chair with the Baxter Research Robot. More precisely, the robot is providing support by holding the chair in place while the human partner is screwing a leg.

Lastly, specific limitations of the robotic platform itself constrain the task domain. For example, as highlighted in Section II, of the few works that actively use a robot to physically engage in a collaborative task, a common choice is to rely on flat-pack furniture [1], [2], [3], [4]. Whilst this solution is advantageous for a number of reasons (i.e. big, flat components that are easy to grasp combined with fine manipulation tasks to be performed by the human partner), in many cases the final designs exceed the payload of a standard collaborative robot, which often is lower than 5 kg—e.g. the Baxter robot has a payload of 2.2 kg. As a result, the robot may not be physically capable of handling a completed (or close to completion) assembly, thus limiting its efficacy as a partner. Other elements that are relevant to the design of reusable experiments are: scalability to real-life problems, (low) cost and (high) availability of the constituent parts, ability to be intuitively understood by naive participants, and flexibility in terms of cognitive skills required (e.g. assembling a complex electronic circuit vs stacking blocks).

Not only do these elements hinder the design and replication of HRC experiments, but they also obfuscate evaluation and comparison of research. This is especially true when dissemination of the technology is deemed secondary to the scientific contribution. To address the issue, separation between task domain and algorithmic contribution needs to be encouraged, in order for comparative studies to aim attention at the algorithm itself while ‘freezing’ all the components related to robot platform and experimental evaluation. Further, lack of standardization and reproducibility impedes efforts to integrate advancements from within the field itself and outside areas. Hence, we claim that the

absence of a unified *language* for HRC experiments strongly limits the comparative evaluation of methods from the field.

In this article, we introduce a common ground for designing HRC tasks that is intended to be reused across a wide range of experiments. To this end, we aim at comprehensively covering the broad spectrum of works in HRC, with the goal of devising a reference design for which it is possible to tune the many aspects of collaboration as parameters. Specifically, we target the following properties: i) the robot is an effective contributor to the collaborative task—which rules out many domains due to current limitations of modern robotics systems; ii) the target domain enables by design a variety of capabilities (in terms of perception, control, and reasoning skills of both partners), to account for aspects of role assignment between peers—should it be predefined, negotiated, or take the form of one peer supporting the other; iii) the task is at least partially perceivable/understandable by both partners for them to share a common goal, and tuning information asymmetry in favor of either peer is possible; iv) communication—used to impart orders, negotiate, disclose internal states, or gather information—is treated as a first-class citizen in the design of the task; v) the target domain is scalable in terms of complexity of the task and complexity of the human-robot interaction.

In order to improve permeability between independent research in the field, we propose a reference set of tasks for reuse in HRC experiments that aims at easing implementation and reproducibility, and allows adaptation to a variety of robot environments and capabilities. We present a set of basic components that can be built into a variety of final designs, thus providing multiple collaborative tasks tailored to HRC research (Section III). Importantly, we do not focus on presenting a model set for benchmarking; rather, we aim at accelerating convergence toward a standardized task domain. We have designed the model set to fulfill the following requirements (Sections IV and V): i) accessibility of the basic components and low cost of the final designs; ii) modularity, scalability and ease of re-use of the components; iii) applicability of the model set to a wide range of experiments in the broader HRC field; iv) scalability of task complexity; v) lightweight design to suit the payload of the vast majority of collaborative platforms. With this set of assembly tasks we also provide open-source software for robot perception and manipulation, in order to lower the barrier to entry for new users (Section VI). By providing tools for perception, manipulation, and interaction we aim at promoting replicability; further, by encouraging consistency across experiments, we will add legitimacy and significance to contributions in the field.

II. BACKGROUND AND RELATED WORK

Historically, while benchmarking proved crucial to achieve scientific, replicable research, its applicability to robotics has been limited due to the complexity of the field. For this reason, previous work focused on establishing benchmarks in specific sub-fields of robotics research, e.g. manipulation [5], motion planning [6], [7], navigation [8], [9], service

TABLE I
SELECTION OF PRIOR WORK IN HRC (IN CHRONOLOGICAL ORDER).
FIRST COLUMN: GENERAL DOMAIN; SECOND COLUMN: APPROXIMATE
NUMBER OF PARTS AND TOOLS INVOLVED IN ASSEMBLY TASKS (IN
PARENTHESIS, NUMBER OF DISTINCT PARTS WHEN APPLICABLE); THIRD
COLUMN: NATURE OF THE ACTIONS PERFORMED BY THE ROBOT.

	Domain	Parts	Robot actions
[19]	manufacturing	–	pick, place
[20]	load sharing	–	collaborative hold
[21], [22]	buttons	3	press button
[23]	warehouse	–	pick, hold
[24]	sweeping	–	move
[25]	set table	–	pick, place
[26]	kitchen, set table	–	–
[27]	assembly	10	pick, pass
[28]	kitchen, set table	–	pick, place
[29]	assembly	5 (2)	hold, pass
[30]	virtual	–	pick, place
[31]	assembly	40 (20)	–
[32]	assembly	–	pick, place
[33]	warehouse	–	pick, place
[34]	manufacturing	–	pick, place
[35]	manufacturing	–	collaborative hold
[4]	assembly	5 (2)	hold
[2], [3]	assembly	5 (4)	pick, assemble
[36]	assembly	14 (6)	–
[37]	assembly	8	–
[38]	assembly	14 (3)	pick, place
[39]	assembly	12 (6)	–
[40]	diaper change	2	pick, pass
[41]	manufacturing	–	hold
[42]	kitchen	– (1)	pick, pass
[43]	assembly	3	–
[44]	assembly	5	pick, pass, hold
[1]	assembly	4 (4)	pick, place, pass, hold
[45]	assembly	29 (6)	pick, place, pass, hold
[46]	assembly	41 (10)	pick, place, pass, hold

robotics [10], [11] or human-robot teamwork [12]. Further, a number of robotic platforms have been designed with the specific purpose of mitigating the issue by fostering robotics research on shared hardware (e.g. PR2 [13], iCub [14], and Poppy [15]). Closer to this work are instead robot competitions [10], [16], [17], [18], that propose a standard set of challenges to homogenize evaluation of robot performance.

A great variety of tasks has been used in human-robot collaborative studies, each featuring specific properties that address one or several aspects of the collaboration. It is therefore particularly challenging to provide tools to standardize research in the field. A selection of relevant work is shown in Table I. We included papers that either explicitly or implicitly target HRC applications, with particular focus on contributions that feature practical deployments or concrete task domains. This selection criteria resulted in a wide variety of final applications, with works that e.g. focused on evaluating the communication between peers [21], [22], [29], [33], [1], leveraging the collaborative domain to target physical human-robot interactions [20], [35], or proposing a teaching scenario where the robot learns a task model [30], [37], [44].

This broad landscape of applications is then summarized in Table I in terms of the type of actions implemented on the robot platform. Remarkably, there is consistent overlap in robot capabilities; they are typically reduced to pick and place movements, button presses, hold actions, object passing, or basic assemblies. When applicable, the table also shows a basic measure of task complexity in terms of number of parts that need to be assembled. This directly relates to the intricacy of the task that needs to be represented (and eventually learned) by the robot, and the complexity of the physical interaction with the human partner. Of particular importance is the fact that these tasks are often tailored to the experiment, robot, and research space involved, and do not allow for a high degree of replicability. Further, there is frequently a hazy distinction between what constitutes the experimental scenario and the theoretical contribution itself. For example, papers addressing kitchen assistance domains such as setting the table feature a set of robot skills very similar to contributions that target advanced manufacturing. To alleviate these issues, we propose a framework aimed at converging toward a *common language* for human-robot collaboration. It is meant to provide a set of assembly tasks from which researchers can build a variety of HRC applications. Importantly, we demonstrate how the proposed model set covers the majority of the works detailed in Table I in terms of both task complexity and robot capabilities. In addition to that, we provide a set of software tools and libraries to facilitate the development of such experiments.

III. THE HRC MODEL SET

We propose a model set for human-robot collaboration experiments, aimed at standardizing repeatability and accelerating research in the field. We address the fact that, despite HRC being a growing area of study within the broader robotics umbrella, the difficulty to reproduce contributions in the field hinders its progress. To this end, we aim at providing a cohesive and comprehensive set of building blocks with the goal of: i) covering the needs of typical HRC works in terms of complexity and robot skills; ii) facilitating the setup of collaborative experiments to accelerate research in the field. In the following, we detail the hardware components that constitute the model set (Section III-A), and we later present a number of prototypical designs built on top of said components (Section III-B).

A. Hardware Components

The components of the model set can be divided in four main groups: *screws*, *dowels*, *plywood*, and *brackets*. These basic constituents can be assembled in a variety of different configurations, as detailed in Section III-B (please refer to Table II for a comprehensive list of all the parts that belong to the model set). As mentioned in Section I, we aim at providing a set of accessible tools *by design*. For this reason, the screws, dowels, and plywood can be easily purchased at a hardware store, craft supply, or online. For the purposes of this work, we use dowels with a diameter of $1/2$ in, standard #4 screws in $1/2$ in and $1/4$ in lengths, and birch plywood with

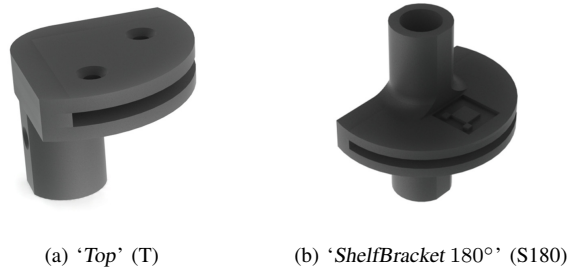


Fig. 2. Example 3D renderings for some of the brackets available in the model set. The design of the bracket was carried out with modularity in mind. From the basic ‘Top’ element in Fig. 2a, eight derivative brackets have been developed in order to extend the type of final designs the model set allows for. Optionally, a fiducial marker can be embedded in the design, to facilitate perception (Fig. 2b).

a thickness of $1/8$ in that has been laser cut into rectangles of different sizes (6×8 in, 6.75×2.5 in, and 6×16 in respectively). The only components that have been custom designed to suit the purposes of this work are the brackets, which are meant to connect dowels and plywood panels in a variety of ways, and are held in place by the screws. Their CAD models have been released under a *CC-by-sa* open-source license, and are freely available on GitHub¹.

In order to be able to flexibly extend the capabilities of the model set, we started from the basic element depicted in Fig. 2a, which simply connects a dowel with a plywood panel. This bracket, referred to as ‘Top’ (T) in Table II, is particularly suited for connecting e.g. a leg with the top of a table. Importantly, minimal changes to its design lead to new connections that can scale according to the specific needs of the user. In this work, we have designed a set of seven additional modules, hereinafter referred to as ‘ChairBack’ (CB), ‘ChairBracketL’ (CL), ‘ChairBracketR’ (CR), ‘Foot’ (F), ‘ShelfBracket 90°’ (S90), ‘ShelfBracket 180°’ (S180, Fig. 2b), and ‘Top 180°’ (T180). Each of these modules is used as a component of the prototypical assemblies described in Section III-B (whose constituent parts are listed in Table II).

B. Example Designs

The model set has been designed with modularity and scalability in mind. The main idea is to provide a number of basic parts that can be used to create a variety of different objects, tailoring to the specific needs of the user and the platform used. For the purposes of this work, we present four prototypical objects built from the parts described in Section III-A: a table, a chair, a shelf and an entertainment console. Fig. 3 illustrates a set of 3D renderings for the four basic configurations, whereas Fig. 4 shows the final objects for table and chair.

These four designs have been devised to showcase the flexibility of the model set. As highlighted in Table II, they

¹ github.com/scazlab/HRC-model-set hosts both STL and SolidWorks® files for the brackets, along with specifications for 3D printing, tutorials for assembling the example designs in Section III-B, and reference links for purchasing dowels, plywood panels and screws.

TABLE II

LIST OF COMPONENTS FOR THE EXAMPLE DESIGNS DETAILED IN SECTION III-B, WITH NUMBER OF PARTS, ESTIMATED COST AND WEIGHT.

	Brackets							Wooden	Plywood [in]			Screws [in]		Number of Parts	Cost [\$]	Weight [g]	
	CB	CL	CR	F	S90	S180	T180	T	Dowels	6x8	6.75x2.5	6x16	1/2				1/4
<i>Table</i>	0	0	0	4	0	0	0	4	4	1	0	0	8	8	29	58	208
<i>Chair</i>	2	1	1	4	0	0	0	2	7	1	1	0	14	8	41	84	312
<i>Shelf</i>	0	0	0	4	4	0	0	4	8	2	0	0	16	16	54	103	398
<i>Console</i>	0	0	0	8	12	4	4	4	24	6	0	2	48	24	136	302	1299

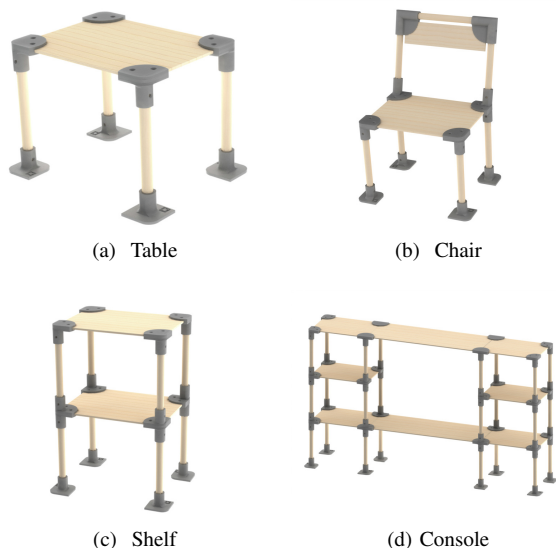


Fig. 3. 3D renderings of the example designs detailed in Section III-B. Please refer to Table II for a list of the components needed for each design.

differ in terms of number of parts, cost, and weight. The simplest construct—the table—is composed of only 29 parts, whereas the console represents the other extreme, with 136 parts—of which 72 are screws which need to be screwed in to complete the assembly. Further, some configurations feature a higher level of modularity than others: for example, combinations of 10 unique parts can be made into either the chair (41 total parts) or the more complex console (136 parts). Either may prove useful for specific experimental requirements. This ability to support a number of final designs with different properties enables granular control of the interaction between the robot and the human partner, and is considered an important asset of the model set. Please refer to Section IV for a detailed discussion on the features of the example designs in particular, and the model set at large.

IV. FEATURES AND USAGE

In this section, we detail the peculiar features of the proposed model set, by focusing on how it is easy to distribute, compatible with many robotic platforms, and an efficient solution for a variety of collaborative experiments.

A. Design Considerations for Robot Interactions

It is common practice in robotics research to shape the environment to suit the physical embodiment of the robot and to conform with the constraints dictated by the experiment. Similarly, an important requirement for the design of the



(a) Table

(b) Chair

Fig. 4. Final constructed objects (see Section III-B). Similarly to brackets (Fig. 2), flat surfaces such as the plywood panels can be equipped with adhesive fiducial markers to ease perception.

HRC model set was for it to be easily deployed to a variety of robotics platforms and collaborative setups. To this end, a number of considerations were taken into account. In the following, we highlight the design choices behind the HRC model set to ease perception, manipulation and robot control.

Although a wide range of state of the art algorithms can be exploited in HRC experiments, we designed the model set to comply with minimal requirements. For what concerns the perception problem, our only prerequisite is the availability of standard RGB cameras. As detailed in Section VI, we provide a marker-based software to detect flat parts such as the plywood panels, if conveniently equipped with proper markers (cf. Fig. 3). For dowels, we have created a marker made up of a flat piece of plywood and a cable clip that attaches to them (see Fig. 5a). Lastly, brackets can be printed with a unique marker identifier that takes advantage of the same technique, as shown in Fig. 2b.

The design of the model set also takes into consideration standard hardware in terms of manipulation and grasping. With the only exception of the screws—that are too small for most robotic platforms to date—it is possible to individually pick up the totality of the parts that compose the model set. Flat surfaces are particularly suitable for vacuum grippers; if a vacuum gripper is not available, they can be vertically positioned in order to facilitate pickup by a standard parallel gripper. Dowels and brackets offer the same type of affordance to a robot, as they are suitable for e.g. parallel grippers that can grasp by aligning with the longitudinal axis. Furthermore, it is always possible to conveniently locate smaller objects such as screws into bigger containers that are easier to manipulate, such as the boxes used in Fig. 5b.

Lastly, to overcome the limitations of standard collaborative robots' payloads and ease experiment design, we favored

a scaled down furniture design over e.g. commercially available solutions. We employed two strategies: i) reduce the size of each final design to only one fourth of what it would be as a real furniture item; ii) use lightweight components. This latter aspect led to the choice of plywood among a variety of possible materials, and weight considerations when designing the brackets. As a consequence of this, the heaviest bracket weighs only 18 g and the heaviest plywood panel used in the assembly is 127 g. Importantly, even the entertainment console has a weight of only 1.3 kg in total (cf. Table II)—well below the payload limit of most robot platforms.

B. Impact of 3D Printing on Dissemination

3D printing is a quickly evolving technology that facilitates distribution of design; if compared to common manufacturing methods, it allows to contain costs when small batches and multiple iterations are required [47], [15]. Both these features are essential to support the level of dissemination that the HRC model set targets. As mentioned in Section III-A, we open-sourced the design of all the brackets, with the goal of encouraging external contributions. They are available online alongside detailed instructions on how to purchase all the components of the model set, and how to assemble the prototypical configurations. Importantly, the brackets can be printed with any commercially available 3D printer. In order to guarantee that plywood and dowels snap into brackets without being too loose, an additive technique that employs soluble material and a layer height of 0.15 mm are recommended. For the purposes of this work, we employed a fused deposition modeling technique with ABS+ plastic. While the first solution guarantees clean inner edges on the inside of the brackets, ABS+ plastic can withstand multiple utilizations and structurally support the weight of each final configuration.

C. Scalability

An important asset of the model set is that its modular nature allows for reproducibility, reuse, and scalability. For instance, the wooden parts, in particular brackets and dowels, can be reused in numerous assemblies of varying complexity. Indeed, all the prototypical designs described in Table II are characterized by a high degree of reusability: each assembly shares at least 60% of its components with other objects. For simpler designs such as the table, the overlap increases to 100%, which means that the table can be constructed entirely from parts used to create a chair or a shelf. This overlap in terms of components allows to deploy a variety of construction tasks—and hence experimental scenarios—without the need to purchase additional hardware. This enables to explicitly treat the complexity of the assembly (for example in terms of number of distinct parts) as an experimental variable, that can be tuned by combining the same components into different pieces of furniture. Finally, distributing the 3D models of the brackets enables third parties to adapt the furniture designs to specific research spaces and robot capabilities, or to extend the model set to investigate novel scenarios. For example, in order to create

a full-scale chair, one can print the brackets at 4 times their size, use plywood that is $\frac{1}{2}$ in in thickness, dowels that are 2 in in diameter, and #10 screws in 1 in and 2 in lengths.

D. Cost Considerations

Accessibility of the model set is guaranteed by keeping costs reasonably low. As mentioned in Section III-A, most of the parts that compose the model set can be easily retrieved online or at hardware stores for a small price. For the custom designed brackets, we provide an estimate of their cost using the SolidWorks® cost evaluation for 3D printing, which includes the cost of outsourcing the printing if an appropriate device is not available. Results are shown in Table II: for what concerns the majority of the proposed configurations (table, chair, and shelf), it is possible to purchase the totality of the hardware from under \$100; the only outlier being the entertainment console, which is the extreme case composed of 136 different parts, and has a cost of \$300.

V. HUMAN-ROBOT INTERACTION DESIGN

The model set allows for a series of assembly tasks to be performed in collaboration between a human and a robot. Particular attention has been put into providing multiple dimensions of variability in the design of experiments that can cover the broad range of applications HRC research. Not only does it enable a large spectrum of experiments, but it explicitly exposes *experimental variables* for finer evaluation of the algorithms. Firstly, in order to be as close as possible to real collaborative environments, we impose the requirement of needing at least one tool for the assembly. As detailed in Section III-A, only one type of screwdriver is needed by default since the screws are all of the same type, but the user is free to adapt the design to allow for multiple tools. In order to succeed in the construction of each of the objects detailed in Section III-B, the following is a list of the necessary actions: i) pre-drilling screw holes into plywood panels; ii) retrieving components; iii) retrieving tools; iv) mounting parts together; v) holding parts in place to support screwing; vi) screwing parts; vii) cleaning up the workplace. An important asset of the proposed model set is that the combination of skills required for each task—screwing, holding, bringing parts—integrates well with the different capabilities of robot and human. This allows for easy incorporation of robot participation and support into a task that would be difficult for a human to complete alone. For example, human participants might find it particularly difficult to screw a bracket onto a dowel without the support of the robot, who can help by holding the dowel in place (cf. Fig. 1). That is, a potential experiment can force the two partners to a certain level of interaction *by design*.

Further, it is possible to tune role assignment and task allocation between peers, if needed. This is a direct consequence of the design choices that led to the proposed model set (cf. Section IV-A), which is composed of a series of parts that a standard robotic platform can interact with. That is, although some tasks pertain to the human exclusively because of their superior manipulation and perception skills, there exists some

degree of overlap in terms of actions that can be performed by both agents. Consequently, it is possible to devise an experiment in which there is a clear compartmentalization of the roles the two partners have. Similarly to [38], [40], [1], specific actions (e.g. bringing components) can be assigned to the robot, whereas the human can be entrusted to perform only the actions that require fine manipulation skills (e.g. screwing). Conversely, the respective areas of competence can be artificially mixed and the two agents can be forced to compete for the same resource [32].

Lastly, the modularity of the model set allows granular control on the amount of information available to the human and the robot. One of the issues the majority of works in the literature face is that, as of now, the human is far superior to the robot in terms of both perception and manipulation skills. This dissymmetry of skills and available information between the two partners fundamentally limits the amount of possible interactions allowed in standard HRC experiments. Without diverging from the scope of this work, it is possible to customize the model set in order to balance out the information asymmetry, and consequently favor the robot. For example, brackets can be customized to be compatible with multiple plywood depths and dowel widths, and their detection can become particularly challenging for a human participant. Conversely, this granular information can be implicitly embedded in the perception system of the robot—via e.g. the fiducial markers (Section VI). That is, it becomes easier for the robot to pick up the correct dowel for a specific bracket, increasing the efficacy and effectiveness of the robot platform in the task domain. Similarly, it is also possible to obfuscate the task goal by for example providing a set of components that are not needed for the assembly and neither partners (or only one of them) know how to use. These use cases are of particular interest for the study of collaboration with ambiguity in the task structure. As mentioned, the modular structure of the furniture set also makes task complexity an experimental variable that allows the experimenter to modulate the total number of parts involved in the final construction.

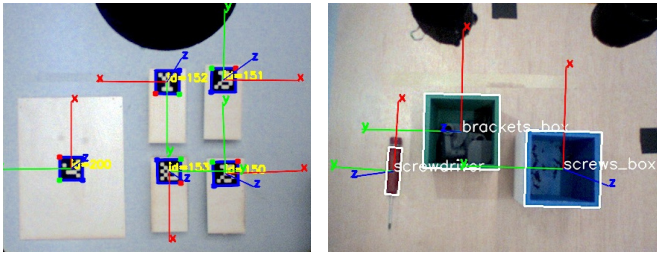
Evaluation of human-robot collaborations: Standardization of human-robot collaborative interactions is currently lacking in the field. Although benchmarking is out of the scope of this work, the proposed model set constitutes a first step toward achieving a common ground for repeatable, reusable algorithms. In related work [1], [45], [46], [48], examples evaluations in HRC settings are presented; they provide insights on how to set the baseline on the collaborative efficiency of the proposed algorithms, as well as how to compare these algorithms across several conditions. In particular, two metrics are introduced: i) the ratio of *valid actions* and *robot errors* as a criteria for robot performance—useful to evaluate the robot w.r.t. a canonical or reference behavior; ii) the *completion time* of the overall assembly to provide an objective evaluation of the interaction. Additionally, one can also consider the human perception of the robot (e.g. through questionnaires) or evaluate the cognitive load on the human.

VI. ACCOMPANYING SOFTWARE

Our model set is intended as an inexpensive, accessible, reproducible, and scalable resource for easier development of HRC experiments. In conjunction with this reference set, we have also developed a complementary collection of software tools and libraries compatible with the Robot Operating System (ROS, [49]), the most used middleware to date for robotics research and industry applications. The accompanying software has been released under the *LGPLv2.1* open-source license, and it is freely available on GitHub². Its main purpose is to facilitate prospective users with the adoption of the model set in collaborative experiments. As such, the use of the software libraries is advantageous in contexts that are similar to what is generally envisioned by the authors. For the purposes of this work, we have developed: i) a collection of robot-independent perception tools tailored to the constituent parts of the model set; ii) a robot controller that exposes a set of high-level actions to the user. Collectively, they target a generic human-robot collaborative setup, with the aim of equipping a standard robot with enough skills for it to effectively locate, manipulate, and interact with parts and tools. It is worth noting how the software presented in this work has been implemented on a Baxter Research Robot, a widely used platform for research in HRC (cf. Fig. 1). Nonetheless, it is easy to customize to any robotic platform that shares similar hardware features and is provided with a kinematic model in URDF form.

As mentioned in Section IV-A, a system based on fiducial markers is provided to support 6D localization of the majority of components of the model set. Such markers can be directly engraved in the brackets during printing, or taped to the plywood panels. For dowels, we provide a marker support that clips to them. The perception system is based on ARuco [50], an OpenCV-based library able to generate and detect fiducial markers; it is depicted in Fig. 5. Additionally, the robot should also be able to interact with smaller objects and tools—and in general objects that are difficult to tag with a marker. To this end, we developed an additional HSV-based color segmentation algorithm, that provides similar 6D localization features as the ARuco based system. Importantly, the HSV software requires minimal information to operate—namely the color of the objects to track and their physical sizes in the real world. As shown in Fig. 5b, the proposed algorithm is capable of successfully identifying and locating a tool (the screwdriver), and containers that store smaller parts such as brackets and screws. Provided a robot-camera calibration system is available, by coupling the feedback from either the ARuco-based system or the HSV-detection algorithm with the robot’s kinematic model it is then possible to track objects into the 3D operational space of any robotic platform. Importantly, both perception systems run at the standard frequency for robot camera systems, that is 30 Hz. This proves useful in order to achieve precise localization and grasping, in particular if the camera is

² github.com/scazlab/human.robot.collaboration hosts both the source code for the software library, and some example applications.



(a) Dowels and plywood panels. (b) Screwdriver, brackets and screws.

Fig. 5. Snapshots of the left and right camera views from the Baxter's end effectors, that show the fiducial marker algorithm (Fig. 5a) and the HSV-based detection system (Fig. 5b) in action.

positioned on the end-effector. In such settings, it is possible to implement visual servoing with a feedback loop that is close to the interaction itself, i.e. with an estimation that continuously improves as the robot gets closer to the target. Please refer to the accompanying video (also available at <https://youtu.be/09Zflg7ZzKU>) for an example of the robot interacting online with an human participant.

On top of the perception algorithms highlighted above, we implement a control library for high-level actions. The goal is to provide a set of basic capabilities for the robot to interact with the components, thus lowering the barrier to entry for new users of the model set. Similarly to the approach used for the perception software, our goal is to provide a set of robot skills that use minimal equipment and do not exploit complex machinery, such as room-scale RGB-D sensing or 3D tracking. For this reason, we implemented the control software on the Baxter robot, using the default hardware and software SDK (see Fig. 1). In the context of this work, we developed two state-less position controllers (one for each of the robot's arms), able to interface with the user through a library of high-level actions (in the form of ROS services). These actions can range from single-arm commands—such as ‘pick screwdriver’—to complex bi-manual trajectories—e.g. ‘handover dowel from left to right arm’—or instructions that require physical interaction with the human—such as ‘hold plywood panel’ (to support screwing).

Lastly, a number of context-based, multi-modal communication channels are provided to support human-robot interactions. The exposed layers are: i) a web interface for high-level action management; ii) a feedback channel on the Baxter's head display; iii) a text-to-speech and a speech-to-text system for voice interactions; iv) a physical interaction layer based on force feedback; v) an error/emergency channel to send error messages to the robot during operation. It is worth noting how the utilization of these interaction layers is at the discretion of the experimenter and is contingent on the specific requirements that each experiment entails; nonetheless, they are deemed a relevant component aimed at supporting HRC experiments and complementing the basic perception and control skill-set of the robot. The model set (and its accompanying software) has been already adopted in related work [1], [45], [46], [48].

VII. DISCUSSION

In this paper, we introduce a model set for human–robot collaboration that is intended to bring HRC research closer to standards for replicability. We provide references and 3D models of the parts composing the set, as well as open-source software for robot perception and human interaction. In particular, we chose to implement the task set as an assembly task, which is close to collaborative manufacturing domains and covers the majority of previous work in terms of task complexity and robot actions (cf. Table I). We expect our contribution to simplify the design and deployment of future experiments in the field, and to form a language around which to establish reference tasks for HRC research.

The ability to accurately compare contributions to the field is indeed essential for further advancing HRC research. With this in mind, we intend for the proposed model set to foster the development of greater standardization. In particular, we provide examples of how to evaluate collaboration performance and expect future works to leverage our contribution to converge toward benchmarks for collaborative domains, and most notably to establish reference measures of HRC task complexity. Importantly, the range of assembly tasks conceivable by the set will not necessarily cover all the potential ramifications of the field. Nevertheless, we are confident that it will prove useful to the rapid, cost-effective development of future experiments. Moreover, we expect future works to identify unforeseen pitfalls and either extend the set or identify new complementary task domains. Finally, a noteworthy direction for future work is to integrate and compare independent contributions under the same experimental setup. Such efforts will enable thorough assessment of their interoperability, scalability, potential for applications, and eventual limitations.

REFERENCES

- [1] A. Roncone, O. Mangin, and B. Scassellati, “Transparent Role Assignment and Task Allocation in Human Robot Collaboration,” *Robotics and Automation (ICRA), IEEE International Conference on*, 2017.
- [2] R. A. Knepper, T. Layton, J. Romanishin, and D. Rus, “IkeaBot: An autonomous multi-robot coordinated furniture assembly system,” in *IEEE International Conference on Robotics and Automation*, 2013.
- [3] S. Tellex, R. Knepper, A. Li, D. Rus, and N. Roy, “Asking for help using inverse semantics,” in *Robotics: Science and Systems*, 2014.
- [4] L. Rozo, S. Calinon, D. Caldwell, P. Jiménez, and C. Torras, “Learning Collaborative Impedance-Based Robot Behaviors,” 2013.
- [5] B. Calli, A. Walsman, A. Singh, S. Srinivasa, P. Abbeel, and A. M. Dollar, “Benchmarking in Manipulation Research: Using the Yale-CMU-Berkeley Object and Model Set,” *IEEE Robotics Automation Magazine*, vol. 22, no. 3, pp. 36–52, Sept 2015.
- [6] B. Cohen, I. A. Şucan, and S. Chitta, “A generic infrastructure for benchmarking motion planners,” in *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Oct 2012, pp. 589–595.
- [7] M. Moll, I. A. Sucan, and L. E. Kavraki, “Benchmarking Motion Planning Algorithms: An Extensible Infrastructure for Analysis and Visualization,” *IEEE Robotics Automation Magazine*, vol. 22, no. 3, pp. 96–102, Sept 2015.
- [8] R. Madhavan, R. Lakaemper, and T. Kalmar-Nagy, “Benchmarking and standardization of intelligent robotic systems,” in *2009 International Conference on Advanced Robotics*, June 2009, pp. 1–7.
- [9] C. Sprunk, J. Röwekämper, G. Parent, L. Spinello, G. D. Tipaldi, W. Burgard, and M. Jalobeanu, “An Experimental Protocol for Benchmarking Robotic Indoor Navigation,” in *The 14th International Symposium on Experimental Robotics*, 2014.

- [10] I. Iossifidis, G. Lawitzky, S. Knoop, and R. Zöllner, *Towards Benchmarking of Domestic Robotic Assistants*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, pp. 403–414.
- [11] T. Wisspeintner, T. van der Zant, L. Iocchi, and S. Schiffer, “RoboCup@Home: Scientific Competition and Benchmarking for Domestic Service Robots,” *Interaction Studies*, vol. 10, no. 3, 2009.
- [12] V. Groom and C. Nass, “Can Robots Be Teammates? Benchmarks in Human-Robot Teams,” *Interaction Studies*, vol. 8, no. 3, 2007.
- [13] Willow Garage. PR2, Personal Robot 2.
- [14] G. Metta, L. Natale, F. Nori, G. Sandini, D. Vernon, L. Fadiga, C. von Hofsten, K. Rosander, M. Lopes, J. Santos-Victor, A. Bernardino, and L. Montesano, “The iCub humanoid robot: An open-systems platform for research in cognitive development,” *Neural Networks*, vol. 23, no. 89, pp. 1125 – 1134, 2010.
- [15] M. Lapeyre, P. Rouanet, and P.-Y. Oudeyer, “The poppy humanoid robot: Leg design for biped locomotion,” in *Intelligent Robots and Systems (IROS), IEEE/RSJ International Conference on*, 2013.
- [16] The Amazon Picking Challenge.
- [17] The DARPA Robotics Challenge.
- [18] F. Amigoni, A. Bonarini, G. Fontana, M. Matteucci, and V. Schiaffonati, “Benchmarking Through Competitions,” in *European Robotics Forum Workshop on Robot Competitions: Benchmarking, Technology Transfer, and Education*, Lyon, France, 2013.
- [19] L. H. De Mello and A. C. Sanderson, “AND/OR graph representation of assembly plans,” *IEEE Transactions on Robotics and Automation*, vol. 6, no. 2, pp. 188–199, 1990.
- [20] K. Kosuge, M. Sato, and N. Kazamura, “Mobile robot helper,” in *IEEE International Conference on Robotics and Automation (ICRA’00)*, vol. 1. IEEE, 2000, pp. 583–588.
- [21] C. Breazeal, G. Hoffman, and A. Lockerd, “Teaching and working with robots as a collaboration,” in *Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems-Volume 3*. IEEE Computer Society, 2004, pp. 1030–1037.
- [22] G. Hoffman and C. Breazeal, “Collaboration in human-robot teams,” in *In Proc. of the AIAA 1st Intelligent Systems Technical Conf.*, 2004.
- [23] P. J. Hinds, T. L. Roberts, and H. Jones, “Whose Job is It Anyway? A Study of Human-Robot Interaction in a Collaborative Task,” *Hum.-Comput. Interact.*, vol. 19, no. 1, pp. 151–181, June 2004.
- [24] K. Kobayashi and S. Yamada, “Human-robot cooperative sweeping by extending commands embedded in actions,” in *Intelligent Robots and Systems, 2005.(IROS 2005). 2005 IEEE/RSJ International Conference on*. IEEE, 2005, pp. 2977–2982.
- [25] A. Müller and M. Beetz, “Designing and implementing a plan library for a simulated household robot,” in *In Cognitive Robotics: Papers from the AAAI Workshop*. Citeseer, 2006.
- [26] M. Beetz, F. Stulp, B. Radig, J. Bandouch, N. Blodow, M. Dolha, A. Fedrizzi, D. Jain, U. Klank, I. Kresse, A. . Maldonado, Z. Marton, L. Mosenlechner, F. Ruiz, R. B. Rusu, and M. Tenorth, “The assistive kitchen – A demonstration scenario for cognitive technical systems,” in *International Symposium on Robot and Human Interactive Communication (RO-MAN 2008)*. IEEE, 2008, pp. 1–8.
- [27] G. Hoffman and C. Breazeal, “Cost-Based Anticipatory Action Selection for Human Robot Fluency,” *IEEE Transactions on Robotics*, vol. 23, no. 5, pp. 952–961, Oct 2007.
- [28] M. Pardowitz, S. Knoop, R. Dillmann, and R. D. Zöllner, “Incremental learning of tasks from user demonstrations, past experiences, and vocal comments.” *IEEE transactions on systems, man, and cybernetics. Part B, Cybernetics : a publication of the IEEE Systems, Man, and Cybernetics Society*, vol. 37, no. 2, pp. 322–32, April 2007.
- [29] P. F. Dominey, A. Mallet, and E. Yoshida, “Real-time spoken-language programming for cooperative interaction with a humanoid apprentice,” *International Journal of Humanoid Robotics*, vol. 6, no. 02, 2009.
- [30] M. Cakmak, C. Chao, and A. L. Thomaz, “Designing interactions for robot active learners,” *IEEE Transactions on Autonomous Mental Development*, vol. 2, no. 2, pp. 108–118, 2010.
- [31] J. Shah and C. Breazeal, “An empirical analysis of team coordination behaviors and action planning with application to human-robot teaming,” *The Journal of the Human Factors and Ergonomics Society*, vol. 52, no. 2, pp. 234–245, 2010.
- [32] J. Shah, J. Wiken, B. Williams, and C. Breazeal, “Improved human-robot team performance using chaski, a human-inspired plan execution system,” in *Proceedings of the 6th international conference on Human-robot interaction*. ACM, 2011, pp. 29–36.
- [33] S. Tellex, T. Kollar, S. Dickerson, M. Walter, A. Banerjee, S. Teller, and N. Roy, “Understanding natural language commands for robotic navigation and mobile manipulation,” in *AAAI Conference on Artificial Intelligence*. AAAI Publications, 2011.
- [34] R. Wilcox, S. Nikolaidis, and J. Shah, “Optimization of Temporal Dynamics for Adaptive Human-Robot Interaction in Assembly Manufacturing,” in *Proceedings of Robotics: Science and Systems*, 2012.
- [35] J. Dumora, F. Geffard, C. Bidard, T. Brouillet, and P. Fraisse, “Experimental study on haptic communication of a human in a shared human-robot collaborative task,” in *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Oct 2012, pp. 5137–5144.
- [36] R. A. Knepper, D. Ahuja, G. Lalonde, and D. Rus, “Distributed assembly with and/or graphs,” in *Workshop on AI Robotics at the Int. Conf. on Intelligent Robots and Systems (IROS)*, 2014.
- [37] B. Hayes and B. Scassellati, “Discovering task constraints through observation and active learning,” in *Intelligent Robots and Systems (IROS), 2014 IEEE/RSJ International Conference on*, 2014, pp. 4442–4449.
- [38] K. P. Hawkins, S. Bansal, N. N. Vo, and A. F. Bobick, “Anticipating human actions for collaboration in the presence of task and sensor uncertainty,” in *Robotics and Automation (ICRA), 2014 IEEE International Conference on*. IEEE, 2014, pp. 2215–2222.
- [39] A.-L. Vollmer, J. Grizou, M. Lopes, K. Rohlfing, and P.-Y. Oudeyer, “Studying the Co-Construction of Interaction Protocols in Collaborative Tasks with Humans,” in *The Fourth Joint IEEE Int. Conf. on Development and Learning and on Epigenetic Robotics*, 2014.
- [40] N. Gopalan and S. Tellex, “Modeling and Solving Human-Robot Collaborative Tasks Using POMDPs,” in *RSS Workshop on Model Learning for Human-Robot Communication*, 2015.
- [41] S. Nikolaidis, P. Lasota, R. Ramakrishnan, and J. Shah, “Improved human-robot team performance through cross-training, an approach inspired by human team training practices,” *The International Journal of Robotics Research*, vol. 34, no. 14, pp. 1711–1730, 2015.
- [42] C.-M. Huang, M. Cakmak, and B. Mutlu, “Adaptive Coordination Strategies for Human-Robot Handovers,” in *Proceedings of Robotics: Science and Systems*, Rome, Italy, July 2015.
- [43] B. Hayes and B. Scassellati, “Autonomously Constructing Hierarchical Task Networks for Planning and Human-Robot Collaboration,” in *International Conference on Robotics and Automation (ICRA)*, 2016.
- [44] M. Toussaint, T. Munzer, Y. Mollard, L. Y. Wu, N. A. Vien, and M. Lopes, “Relational activity processes for modeling concurrent cooperation,” in *2016 IEEE International Conference on Robotics and Automation (ICRA)*, May 2016, pp. 5505–5511.
- [45] O. Mangin, A. Roncone, and B. Scassellati, “How to be Helpful? Implementing Supportive Behaviors for Human-Robot Collaboration,” *arXiv preprint arXiv:1710.11194*, 2017.
- [46] J. Brawer, O. Mangin, A. Roncone, S. Widder, and B. Scassellati, “Situated Human-Robot Collaboration: predicting intent from grounded natural language,” in *Intelligent Robots and Systems (IROS), 2018 IEEE/RSJ International Conference on*. IEEE, Oct 2018.
- [47] R. R. Ma, L. U. Odhner, and A. M. Dollar, “A modular, open-source 3d printed underactuated hand,” in *Robotics and Automation (ICRA), 2013 IEEE International Conference on*. IEEE, 2013, pp. 2737–2743.
- [48] E. Grigore, A. Roncone, O. Mangin, , and B. Scassellati, “Preference-Based Assistance Prediction for Human-Robot Collaboration Tasks,” in *Intelligent Robots and Systems (IROS), 2018 IEEE/RSJ International Conference on*. IEEE, Oct 2018.
- [49] M. Quigley, K. Conley, B. P. Gerkey, J. Faust, T. Foote, J. Leibs, R. Wheeler, and A. Y. Ng, “ROS: an open-source Robot Operating System,” in *ICRA Workshop on Open Source Software*, 2009.
- [50] S. Garrido-Jurado, R. Muñoz-Salinas, F. Madrid-Cuevas, and M. Marín-Jiménez, “Automatic generation and detection of highly reliable fiducial markers under occlusion,” *Pattern Recognition*, vol. 47, no. 6, pp. 2280 – 2292, 2014.