VII. Discussion and Summary

In Sections III–VI, admittance selection conditions for four typical two-PC states each in single-point contact were presented. The strategies and procedures can be used for any combination of two single-contact PCs. In general, for each PC, translational variables δ_{ij} are chosen based on the PC's type. Using translational variables δ_{ij} and orientational variables $(\mathbf{u}, \theta)$, the contact wrenches $\mathbf{w}_i$ and error-measure vector $\mathbf{d}$ are obtained and the interaction of the wrenches from the two PCs is addressed in the calculation of their magnitudes. Then, by (5), the error-reduction function is expressed in terms of $(\delta_{ij}, \mathbf{u}, \theta)$. The function obtained is a polynomial in $(\delta_{ij}, \sin\theta)$ with coefficients being functions of $\mathbf{u}$ and the admittance $\mathbf{A}$. Based on the nature of the contact and the properties of the error-reduction function, selection conditions for the admittance are obtained.

Sections III–VI present results for four of the six combinations of two single-point contact PCs. The remaining two combinations are similar. Note that both the wrench associated with $\{e-e\}$ contact (34) and the wrench associated with $\{\{v-f\}$ contact (22) contain only the linear term in $\sin\theta$. Thus, for $\{f-v,\, e-e\}$ contact, the approach presented in Section V for $\{f-v,\, v-f\}$ contact can be used. Similarly, the approach presented in Section VI for $\{e-e,\, e-e\}$ contact can be used for $\{v-f,\, e-e\}$ contact.

The admittance selection conditions are obtained by geometric and force analysis of each contact state. Redundant coordinates δ_{ij} are used to describe the translational variation. These coordinates are treated as independent variables in a large range (without considering the constraints due to contact). Thus, the conditions obtained are conservative. To make the conditions less conservative, the range of configuration variables considered can be decomposed into a number of nonoverlapping subranges, each addressed with equivalent conditions.

In this paper, a single admittance control law (1) is considered for each contact state. The sufficient conditions obtained impose conditions on the admittance to ensure error reducing motion for the entire contact state. Due to uncertainty in identifying which contact state actually occurs, a single admittance control law (1) *could* be used for all contact cases, if the conditions for all contact states were satisfied simultaneously.

In practice, the selection of an appropriate admittance $\mathbf{A}$ can be formulated as a search routine to find an admittance matrix $\mathbf{A}$ subject to the appropriate conditions. For instance, an optimization procedure can be used to find an admittance matrix for which sufficient conditions such as those described in the paper are used as "constraints." An example illustrating the application of error-reduction conditions to obtain a desired admittance was presented in [5] for a planar case. This procedure applies to the spatial cases as well.

In summary, we have presented conditions for admittance selection of a polyhedral rigid body for force-guided assembly in cases having two point contact PCs. We have shown that, for these cases, the admittance control law can be selected based on its behavior at a *finite* number of configurations. If the error-reduction conditions are satisfied at these configurations, the error reduction conditions will be satisfied for all intermediate configurations.

References

[1] J. M. Schimmels and M. A. Peshkin, "Admittance matrix design for force guided assembly," *IEEE Trans. Robot. Autom.*, vol. 8, no. 2, pp. 213–227, Apr. 1992.

[2] S. Huang and J. M. Schimmels, "Achieving an arbitrary spatial stiffness with springs connected in parallel," *ASME J. Mech. Des.*, vol. 120, no. 4, pp. 520–526, Dec. 1998.

[3] N. Hogan, "Impedance control: An approach to manipulator. Parts I-III," *ASME J. Dyn. Syst., Meas., Control*, vol. 107, no. 1, pp. 1–24, 1985.

[4] S. Huang and J. M. Schimmels, "Sufficient conditions used in admittance selection for force-guided assembly of polygonal parts," *IEEE Trans. Robot. Autom.*, vol. 19, no. 4, pp. 737–742, 2003.

[5] S. Huang and J. M. Schimmels, "Admittance selection for force-guided assembly of polygonal parts despite friction," *IEEE Trans. Robot.*, vol. 20, no. 5, pp. 817–829, Oct. 2004.

[6] S. Huang and J. M. Schimmels, "Spatial admittance selection conditions for frictionless force-guided assembly of polyhedral parts in single principal contact," *IEEE Trans. Robot.*, vol. 22, no. 2, pp. 225–239, Apr. 2006.

[7] J. M. R. Martinez and J. Duffy, "On the metric of rigid body displacements for infinite and finite bodies," *ASME J. Mech. Des.*, vol. 117, no. 1, pp. 41–47, 1995.

[8] D. E. Whitney, "Quasi-static assembly of compliantly supported rigid parts," *ASME J. Dyn. Syst., Meas., Control*, vol. 104, no. 1, pp. 65–77, 1982.

[9] M. A. Peshkin, "Programmed compliance for error-corrective manipulation," *IEEE Trans. Robot. Autom.*, vol. 6, no. 4, pp. 473–482, 1990.

[10] H. Asada, "Teaching and learning of compliance using neural net," *IEEE Trans. Robot. Autom.*, vol. 9, no. 6, pp. 863–867, 1993.

[11] M. Shimizu and K. Kosuge, "An admittance design method for general spatial parts mating," in *Proc. IEEE Int. Conf. Robot. Autom.*, New Orleans, LA, 2004, vol. 4, pp. 3571–3576.

[12] M. S. Ohwovoriole and B. Roth, "An extension of screw theory," *ASME J. Mech. Des.*, vol. 103, no. 4, pp. 725–735, 1981.

[13] J. Xiao and L. Zhang, "Contact constraint analysis and determination of geometrically valid contact formations from possible contact primitives," *IEEE Trans. Robot. Autom.*, vol. 13, no. 3, pp. 456–466, Jun. 1997.

[14] R. M. Murray, Z. Li, and S. S. Sastry, *A Mathematical Introduction to Robotic Manipulation*. Boca Raton, FL: CRC Press, 1994.

Toward a Natural Language Interface for Transferring Grasping Skills to Robots

Maria Ralph and Medhat A. Moussa

***Abstract*—In this paper, we report on the findings of a human–robot interaction study that aims at developing a communication language for transferring grasping skills from a nontechnical user to a robot. Participants with different backgrounds and education levels were asked to command a five-degree-of-freedom human-scale robot arm to grasp five small everyday objects. They were allowed to use either commands from an existing command set or develop their own equivalent natural language instructions. The study revealed several important findings. First, individual participants were more inclined to use simple, familiar commands than more powerful ones. In most cases, once a set of instructions was found to accomplish the grasping task, few participants deviated from that set. In addition, we also found that the participant's background does appear to play a role during the interaction process. Overall, participants with less technical backgrounds require more time and more commands on average to complete a grasping task as compared to participants with more technical backgrounds.**

***Index Terms*—Grasping, human–robot interaction, natural language instruction, skill transfer, user-adaptive robotics.**

I. Introduction

In the last few years, both personal and service robots have become a popular area of study within the robotics community. With

Manuscript received April 30, 2007; revised November 22, 2007. This paper was recommended for publication by Associate Editor C. Breazeal and Editor H. Arai upon evaluation of the reviewers' comments. This work was supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC) and in part by MacDonald, Dettwiler, and Associates (MDA) Space Missions.

The authors are with the Intelligent Systems Laboratory, School of Engineering, University of Guelph, Guelph, Canada (e-mail: mralph@uoguelph.ca; mmoussa@uoguelph.ca).

Digital Object Identifier 10.1109/TRO.2008.915445

current advances, robots are now being seen as assistants in corporate offices [1], aids for the elderly [2], helpers for the disabled [3], and as museum tour guides [4]. Yet there still remain several challenges. One key challenge is how these robots will interact with nontechnical users. Since many everyday users do not have the necessary programming skills to control these complex machines, new interaction modes must be developed. Skills such as grasping will play an increasingly important role as robots are used more regularly in homes and offices. Developing robots that can grasp a variety of objects in an unstructured environment still remains a significant challenge.

A. *Related Work*

There are currently several emerging human–robot interaction modes for applications involving both service and personal robots. Each presents one way of controlling robots, which differs from the more traditional computer-based programming approach typically used in industry.

The first mode is teleoperation and telepresence, which has been used to tightly control a robot's movements. For example, the National Aeronautics and Space Administration's (NASA) Robonaut project [5] incorporates telepresence for conducting tasks in space, while Fong *et al.* [6] use a graphical user interface to control a mobile robot through a cluttered laboratory environment. Although the use of teleoperation and telepresence has produced successful robot control, this mode of interaction still requires time and practice to master.

The second mode is gesturing and imitation. By demonstrating such tasks as spindle assembly [7], robots acquire skills that can later be generalized to other tasks such as changing a paper towel roll in a user's home. Other work [8] explores demonstrating navigation skills to mobile robots, where a robot follows a human teacher as they demonstrate how to traverse a given path around a room. Although this approach is more intuitive for transferring skills between humans and robots, there still remains several challenges including the handling of inaccurate demonstrations and correcting unwanted robot movements.

The third mode is natural language instruction, which has become a common mode of interaction either on its own or as part of a multimodal interaction system. The use of natural language provides an easier communication medium, whereby users do not have to be experts in order to quickly begin commanding a robot. For example, in [9], a group of nonexpert users were asked to instruct mobile robots from one location to another in a miniature town environment.

The last mode is multimodal interaction where the use of both speech and vision systems is typical. This approach has been used for applications such as word–action associations and grasping. For example, the Robota project [10] examines interactions between children and a doll-like robot named Robota. The children explore word–action associations by teaching Robota simple arm movements and the names for different body parts. In [11], multimodal interaction is used to teach a robot a sequence of buttons to be pressed. A robot observes a human teacher press a series of buttons and learns to associate certain buttons pressed with a name and a given order. Multimodal interaction is also being used for grasping tasks [12] where, using both a speech and vision system, a robot is trained by a human teacher to grasp objects using an interactive demonstration approach.

B. *Paper Contribution and Organization*

Although some of the interaction approaches reviewed do incorporate nonexpert users, none have used nonexpert users when teaching robots to grasp. Our goal is to investigate how nonexpert users play a role in transferring grasping skills to a robot. The first step is to develop a communication language specifically focused on grasping that translates nonexpert users natural language instructions to appropriate robotics commands that a robot can execute. Once the robot executes the commands successfully, it can learn from its experience how to grasp other objects [13], [14]. This mimics how humans train their young children to grasp and manipulate various objects.

This paper's contribution is a study where a group of nonexpert users, who have never programmed a robot before, use natural language to instruct a robot arm to grasp several small everyday objects. Our interest is to observe the type of language used, how it was used, and the impact of the user's background and experience on the command preferences and usage. This paper is a significant expansion over earlier work [15], where only the impact of the user's background was briefly discussed. As far as we know, this is the first study of its kind focused on using natural language by nonexpert users for grasping tasks.

The paper is divided into the following sections. Section II outlines details of the study. Section III presents the first set of results focusing on the participants communication language used for commanding the robot. Section IV reports the second set of results focusing on the impact of the participant's background and command selection. Section V concludes our discussion and outlines future work.

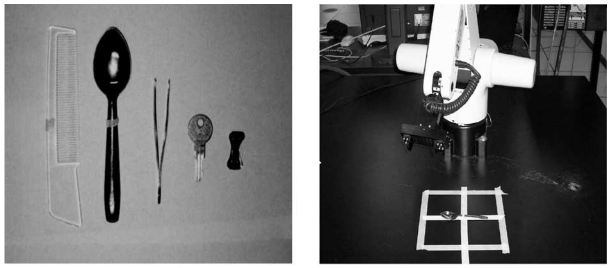
II. Study Details

A. *Experimental Setup Design*

The setup for this study followed the typical stages used in usability studies [16]. First, a test plan was developed that outlined the objective of the study and what metrics would be recorded. Second, questionnaires were distributed in order to select the appropriate group of participants. This was followed by the preparation of relevant testing material such as how participant data would be recorded, pretest and posttest questionnaires, participant consent forms, and task outlines. The test was then conducted with each participant, and a debriefing session was held to gather additional relevant information. The experimental protocol was reviewed and approved by the Ethics Board at the University of Guelph.

Furthermore, we conducted an earlier separate pilot study that had only four participants, but with a similar experimental setup to the one discussed here [17]. Lessons learned from that previous study helped fine tune the final setup of the current study. For example, we added a practice session for each participant and provided explanations of basic robot movements and primitive commands. We also introduced the think-aloud method [18] to encourage participants to verbalize their thought process throughout the experiment.

Participants were asked to command a five-degree-of-freedom CRS A255 human-scale robot arm, equipped with a two-fingered parallel gripper, to grasp five small objects using any of the primitive commands listed in Table II. Once grasped, objects were lifted upward and away from the table surface up to a height decided on by each of the participants. The order of the objects presented to participants was a comb, followed by a spoon, tweezers, a key, and a tensor bandage clip, as shown in Fig. 1(a). The objects used in the study were deliberately selected from items typically found in a home environment, and were considered a challenge to grasp. This contrasts with typical grasping studies, where "laboratory friendly" objects such as cones, balls, and blocks are typically used. Participants were asked to grasp objects in an order that was based primarily on size. The largest object in the set (i.e., the longest), the comb, was selected as the first object for users to grasp. The final object in the set was the tensor bandage clip, which was the smallest and most challenging of the group. This order provided participants with a relatively easier object to grasp first, where the level of difficulty gradually increased for subsequent objects presented. Objects were positioned within a marked off area, as shown in



(a) (b)

Fig. 1. Objects set and pose used during interaction sessions.

Fig. 1(b). During each session, the participant's language, sequence of commands, number of commands, and time to complete the task were recorded.

B. Participants

The participant group consisted of 15 participants ranging in age, occupation, education level, and experience with technology, as shown in Table I. Participants were evenly distributed into one of the three participant classification groups (beginner, intermediate, and advanced), resulting in five participants per group. Participant group assignments were based on a background/screening questionnaire collected from a pool of potential participants. The beginner group was considered the least technical of the three groups. Participants did not use computers or alternative interaction devices on a regular basis, expressed anxiety about using new technology, and were overall more intimidated to interact with robots. The intermediate group used computers on a regular basis at work, and occasionally at home. These participants were interested in interacting with robots, had some experience with alternative interaction devices, and expressed less anxiety about using new technology. The advanced group was considered the most technical of the three groups. This group used computers both at home and at work on a regular basis. Other interactive devices were used more often than the other two groups and their interest in using robots or any new technology was very high. It is important to note that although some participants hold university degrees, in some cases, these individuals were still not considered computer experts or experts with technology.

Participants were informed of their involvement in a robotics experiment; however, details of the experiment were withheld until participants entered the laboratory facility. The first object, the comb, was used as a practice session for all the participants. The practice session did not have a time limit and participants were encouraged to experiment with the set of primitive commands provided. A time limit was also not placed on participants for the remainder of the objects in the study. On average, participants required just under 45 min approximately to complete grasping all five objects. The position of the robot and participant during each of the interaction sessions is shown in Fig. 2. During interaction sessions, participants spoke directly to the robot and maintained eye contact with the robot's end-effector throughout most of the interaction session. A Wizard of Oz type approach [19], [20] was used for the speech interface, where an operator selected commands from a GUI, based on the participant's language. In this study, this approach is preferable to using a fully integrated speech recognition system, since restrictions would have been placed on what the user could and could not say based on the speech systems lexicon of commands. By removing this potential roadblock, the user's dialog opens up, allowing language for existing and new commands to be used.

C. Operator Details

The operator received 2 months of initial training with the robot's GUI prior to commencing the study. Every effort was taken to ensure that the operator did not influence each participant's choice of language or commands used during the study. The operator only provided feedback to participants about the robot's limitations (i.e., maximum range of motion reached), the current status of the robot (i.e., busy, ready for next command), and prompted participants for more information when new natural language commands were encountered. In this case, the operator prompted participants to define these new commands by an equivalent series of primitive commands from the list in Table II. For example, instructions such as "level off" found in Table III were broken down into either a "tilt up" or "tilt down" command, which was used to align the robot's end-effector at a perpendicular angle to the table surface. In some cases, such as the "bend your arm over" command also shown in Table III, the robot was not capable of executing this type of complex command, since it required the coordination of several simple commands. In this event, the command was recorded with the user's expected direction of motion; however, the robot did not execute the command due to limitations in both the robot's range of motion and the GUI used. The operator informed participants of this limitation and requested an alternative command be issued.

D. Primitive Command Set

As illustrated in Table II, the primitive command set consisted of 12 commands, 11 of which were considered simple, more predictable commands. The Move Closer command was used as the only complex command choice. Complex commands involve more than one simple command. For example, combining Move Down with Move Forward results in a complex command. The Move Closer command produces a final gripper position closer to the object's geometric centre. The orientation of the gripper, however, does not change. This command was included to test how often participants used commands they were less familiar with.

A standard scalar equation with a coefficient of motion f (1) was used to determine the relative distance between the tool frame and the object frame.

$$d = \sqrt{(f(x_1 - x_2))^2 + (f(y_1 - y_2))^2 + (f(z_1 - z_2))^2} \qquad (1)$$

The coefficient of motion f was used to ensure the proportionality between the robot's degree of motion and its location with respect to the object. In other words, the farther apart the tool and object frames, the larger the translation. Likewise, the closer the tool and object frames, the smaller the translation. Possible values for f included 0.1, 0.3, 0.6, and 0.9 for translations toward the object and 1, 3, 6, and 9 for translations away from the object once the object was grasped. This equation was used to map participant expectations to the executed actuator motion as closely as possible.

The coefficient of motion f was initially set to a default value of 0.3. This initial setting was changed according to the participant's language. For example, if descriptive terms such as "a lot" were used, then f was set to 0.9. Once the factored distance d was calculated, this value was used as the robot's translation for any of the primitive commands provided. However, instructions for orienting the gripper still used standard degree values ranging from 1° to 90°. Each command was executed as one continuous motion, moving the robot's arm from one location to another within the robot's workspace. Following failed grasp attempts in which the object moved, the object was repositioned back to its original pose before another grasp attempt was made.

TABLE I
STUDY PARTICIPANTS

User	Age	Occupation	Education	Major	Classification Group
1	25	Student	University	General Science	Intermediate
2	65	Retired Warehouse Supervisor	High School	-	Beginner
3	31	Lease Administrator	College	Nursing	Intermediate
4	31	Program Coordinator	University	Business	Intermediate
5	25	Student	University	General Science	Intermediate
6	33	Programmer	University	Mathematics	Advanced
7	34	Programmer	University	Engineering	Advanced
8	33	Programmer	University	Computer Science	Advanced
9	32	Systems Analyst and Programmer	University	Business and Computers	Advanced
10	60	Retired Nurse	College	Nursing	Beginner
11	31	Teacher Librarian	University	Education	Intermediate
12	65	Retired Office Manager	High School	-	Beginner
13	59	Bank Teller	High School	-	Beginner
14	33	Programmer	University	Business and Computers	Advanced
15	58	Office Clerk	University	Psychology	Beginner

Fig. 2. Interaction session.

III. RESULTS I: COMMUNICATION LANGUAGE

All participants initiated their own commands without any help from the operator, and were able to grasp all five objects using the supplied primitive command set. However, there were other interesting findings.

A. Natural Language Mapping of Primitive Commands

During each experiment, common terms emerged for the primitive commands provided, as shown in Table II. Language such as "lift up," and "move outward" were commonly used to refer to the Move Up and Move Forward primitive commands. Participants also used objects in the laboratory to give the robot some sense of orientation. Commands such as Move Left became "move closer to me," likewise Move Right became "move away from me" or "move the other way." Participants also viewed the robot as a human arm, using terms similar to those used to explain new skills between people. Language such as "straighten your arm out" was commonly used in place of the Move Forward command.

B. New Commands

The study also revealed new commands, as shown in Table III, from participant feedback and interaction sessions. New commands were generated by participants without the help of the operator. In this case, participants used language typically found in human–human interactions to instruct the robot. The "turn sideways" command, for example, was used to orient the gripper by some θ value either perpendicular or parallel to the table surface. Language such as "point to the object" was also used to position the robot's gripper directly over the object.

Participants also provided feedback, encouragement, and verbalized disappointment at commands that did not meet their expectations. Feedback from participants such as "you're almost there," and "you've got it" were used to give the robot some sense of its location and the success of the grasp attempted. Participants also provided encouragement such as "good" and "that's it" to the robot in much the same way as children are encouraged during the learning of a new skill. If the robot did not perform as expected, participants verbalized their disappointment to the robot through language such as "oh that's wrong." In some cases, the robot moved much further than participants expected. In this case, participants used language such as "go back to where you were" to command the robot to return to its previous location. This type of command was recorded, but not executed by the robot.

C. Command Translation and Implicit Assumptions

While mapping various commands to natural language, we observed that participants often made implicit assumptions about the robot's behavior. Commands such as Move Down, for example, were issued with the assumption that the robot would intuitively know how far to move. During the practice session, participants issued a Move Down command to the robot, then sat staring at the robot expecting an additional downward motion to take place. When asked what they expected for the Move Down command, participants responded that they "expected the robot to move all the way down to the table surface in one motion." Eventually, participants included terms such as "a little" or "a lot" with commands issued. Typical descriptive terms used by participants for translating and rotating the robot's tool frame are shown in Table IV. Values for f were used as part of (1), whereas θ values were used directly by the command itself. Since values shown in Table IV only require a target position to be known *a priori*, this allows for a more generalized approach, which can be used for other types of robots and tasks.

Participants also chose to leave out commands altogether when issuing several of the same commands in sequence. For example, commands such as Move Forward and Move Down were followed by languages such as "again" or "keep going." Assumptions were also made concerning locations for objects in the laboratory. Participants assumed that the objects located in the laboratory were known to the robot *a priori*. This assumption produced language that included these locations as landmarks for positioning the robot's gripper. The robot was also viewed as a machine, where participants were unsure of what the robot knew *a priori*. This resulted in simple direct language, which was restricted to short-contained instructions. Assumptions were also made regarding the robot's end-effector. Participants viewed a 180° rotation instead of a 360° rotation of the end-effector as a full turn.

TABLE II
COMMON NATURAL LANGUAGE INSTRUCTIONS

No.	Primitive Commands	Natural Language	Robot Action
1	Move Up	lift/ raise (up), come up as far as you can	translate along +Z world axis
2	Move Down	come/ go down far as you can, touch the table/ surface below, lower your arm	translate along -Z world axis
3	Move Right	go away from me/ the other way	translate along +Y world axis
4	Move Left	come closer to me/ this way	translate along -Y world axis
5	Move Forward	push/ straighten/ stretch your arm out, move out(ward)	translate along +X world axis
6	Move Backward	pull your arm back into your body, move in(ward)	translate along -X world axis
7	Tilt Up	tilt forward/ back, move your wrist/ hand up and back/ forward	pitch up tool frame
8	Tilt Down	tilt forward/ back, point all the way down (to the table/ floor/ ground), move your wrist/ hand down and back/ forward, drop/relax your wrist	pitch down tool frame
9	Rotate	turn your wrist/ hand around	roll tool frame
10	Close	close your fingers/ hand, close jaws, catch/ grab it, grasp it	close grippers
11	Open	open your fingers/ hand, open jaws, release, let go	open grippers
12	Move Closer	go/ get closer to object	translate towards geometric centre of object (world frame)

TABLE III
NEW LANGUAGE

No.	Type	Natural Language	Robot Action
1	New Cmd	point to the object, look at the object	pitch down tool frame towards geometric centre of object
2	New Cmd	bend your arm over	combine Move Forward, Move Down and Tilt Down commands
3	New Cmd	[command] slowly	vary speed of robot's movement
4	New Cmd	pick it (object) up	combine Close and Move Up commands
5	New Cmd	level off, make it flush, turn so that you are sideways/ longways	roll or pitch tool frame so approach axis (X) is either perpendicular or parallel to table/ surface
6	Feedback	you're right (on top)/ near the object now	
7	Feedback	(you're) almost there, getting close now	
8	Feedback	you've got it/ (the object) now, oh you (just) missed it, that's it, okay, good	
9	Corrective	(you've gone)/ that's too much/ too far/ too little	
10	Corrective	oh that's right/ wrong	
11	Corrective	go back to where you were	

TABLE IV
DESCRIPTIVE LANGUAGE AND VARIABLE ASSIGNMENTS

No.	Variable	Values	Language
1	f	0.1	tiny bit, little bit, slightly
2	f	0.3 (default)	again, continue, keep going, once more, little more, more than slightly, move (no description)
3	f	0.6	(some) more, further, twice as much (as default amount)
4	f	0.9	a lot, large amount, three times as much (as default amount)
5	θ	1°	a tiny bit, very slight
6	θ	11°	a little bit, slightly
7	θ	22°	(little, some) more, further
8	θ	45° (default)	again, continue, keep going, once more, move (no description)
9	θ	90°	much more, half way, quarter turn, turn a lot, perpendicular/ parallel to now, horizontal/ vertical to now, turn all the way up/ down

Since a parallel gripper was attached as the grasping tool, a 180° rotation appeared to produce the same gripper configuration as the robot's initial gripper configuration. This produced language such as "turn half way" for 90° rotations.

D. Misinterpreted Commands

Although commands from the primitive command set were explained to participants prior to commencing the experiment, some participants still misinterpreted how certain commands would be executed. The Move Down command, for example, was interpreted as the Tilt Down command, pitching down the tool frame instead of moving the entire arm downward. Other misinterpreted commands included Move Up and Tilt Up. Again, the tool frame was thought to pitch upward instead of moving the entire robot's arm upward. There was also some confusion regarding the Rotate command. Here, participants interpreted the command as a rotation of the robot's "body", instead of a roll of the tool frame around its approach axis (X). Suggestions for changing the command included using the word "turn" such as "turn your wrist around."

E. Perceptions

Based on the participant's seating arrangement, some participants viewed certain commands differently. Move Forward, for example, was seen as the Move Left command based on the participant's perspective, since participants were seated at a 90° angle to the left of the robot. Likewise, commands such as Tilt Down and Tilt Up were viewed differently depending on the orientation of the gripper. At 0°, the tool frame's approach axis (X) is parallel to the table surface below. Once the gripper is tilted upward or downward to within $+/-45^\circ$, the command is perceived as a tilt down or tilt up movement. The participant's perception changes once the tool frame moves beyond $+/-45^\circ$. From here, the Tilt Up command becomes "tilt up and back," and the Tilt

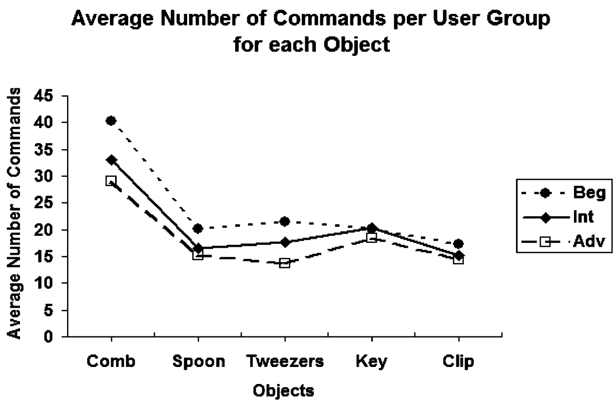


Fig. 3. Average number of commands recorded per object for beginner, intermediate, and advanced users.

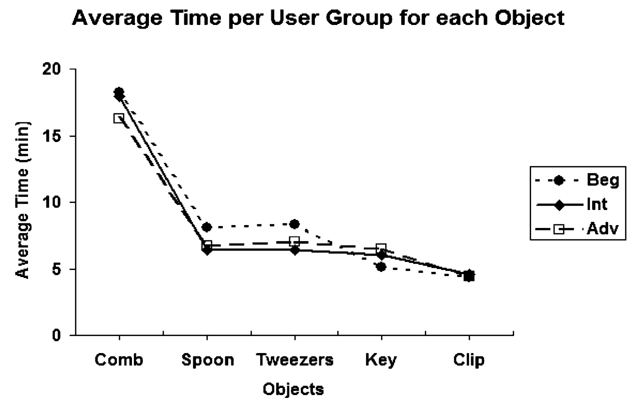


Fig. 4. Average time recorded per object for beginner, intermediate, and advanced users.

Down command becomes "tilt down and back." Likewise, once the tool frame moves beyond $+/-90°$, Tilt Up and Tilt Down commands become "tilt up and forward" and "tilt down and forward."

F. Participant Behavior Observed

During the course of the study, it was found that most participants eventually developed their own individual language for moving the robot into a desired position and orientation. Once established, usually by the second or third object, participants seldom deviated from this established instruction set.

Some participants also found the use of natural language instructions to be a challenge for certain movements such as orienting the robot's tool frame. Trying to rotate the end-effector by a desired θ value using ordinary language, other than degrees, was a difficult task. Other participants found translating the robot's tool frame to be more difficult. Issuing a Move Down command, for example, was straightforward; however, finding the descriptive language to indicate how far the robot should move was more of a challenge.

All participants used a "stop" command initially in order to control the robot's motion more closely. Participants issued a command such as Move Down followed by "stop now" when the robot reached a desired location. Participants also used metrics to establish tighter control of the robot's movements. Language such as "move down 1 inch" was used in order for participants to predict the robot's behavior.

Several participants also chose to personalize the robot. Unique names such as "Robin," "Rachel," "Max," and "Sandpiper" were assigned to the robot. In most cases, the robot's new name was used prior to the command. Language such as "Robin, I want you to move forward," was used to clearly direct the current instruction to the robot.

IV. Results II: Participant's Background and Command Selection

Answers from pretest and posttest questionnaires along with information gathered from each interaction session were collected and analyzed.

A. General Observations

We tracked the amount of time and the total number of commands used to grasp each object as performance indicators. Overall, Figs. 3 and 4 suggest that the inclusion of a practice session leads to a significant improvement for both the number of commands and times recorded by participants from the first object (i.e., the comb) to the second object (i.e., the spoon) used in the study. Furthermore, all participants, regardless of background, experienced difficulty with one or more objects in the object set. Of the five objects used, the key appeared to be the most problematic one, with most participants experiencing declines in individual performance. This was followed by the tweezers, the tensor clip, and the spoon.

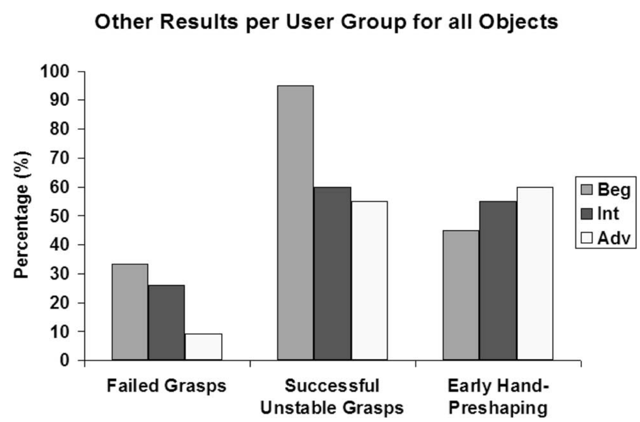


Fig. 5. Other results for all objects for beginner, intermediate, and advanced users.

Some participants chose to align the gripper with the handle of the object on the table, much the same way as *they* would retrieve these types of objects. For example, most participants grasped the spoon by its handle, instead of by its round concave section, which has a larger surface area. This approach suggests that participant preferences and experience do play a role in grasp choices. Participants also chose gripper configurations, which resulted in unstable grasps, as shown in Fig. 5. Unstable grasps are grasping configurations that result in an object moving once the gripper fingers touch the object due to linear or angular misalignment. Linear misalignment occurs when the gripper is not centered directly over the object. Angular misalignment takes place when the gripper is not positioned at a perpendicular angle relative to the object's main axis. In both the cases, the object either moves during a grasp attempt or is grasped using a small surface area, increasing the likelihood for slippage to occur afterward. Some participants also tried to scoop the practice object, the comb, by tilting and attempting to close the gripper on the object. Once this approach proved unsuccessful, alternative gripper configurations were successfully found.

B. Comparison of Participant Groups

1) Beginner Group: This group appeared to approach the task with the most caution, asking questions and gathering as many details as possible. All participants in this group requested a "slow" movement when the robot approached the object. Participants needed to be in complete control of the robot to ensure unwanted movements could be contained. Some beginner participants in the study also grasped objects using a unique gripper configuration. Typically, following a pitch down of the tool frame by -90°, the tool frame is then rolled 90° to prepare for a grasp attempt. However, some participants on occasion chose instead to rotate the tool frame between 0° and 90°, choosing, for example, to rotate the gripper by 68° and 77°. Participants believed that this type of angular alignment produced a more stable grasp for certain objects.

This group was also the only group that did not use the Move Closer command. Participants in this group either did not realize that this command was available or expressed feeling uncomfortable using instructions that were less predictable. Since the robot's final location could not be visualized, participants felt that there was a potential for the robot to move too close to the object. Overall, participants in this group appeared to use a trial and error approach the longest in order to establish a clear mental model of the robot's capabilities. Figure 5 shows that this group also produced the most number of failed grasp attempts, the highest number of unstable grasps, and performed early hand preshaping less often than the other two groups in the study.

2) Intermediate Group: This group was less apprehensive when controlling the robot than the beginner group. Participants did not require significant feedback from the operator and asked fewer questions during each interactive session. They also appeared to use a trial and error approach for a shorter period of time as compared to the beginner group, and produced the second highest number of failed grasp attempts and unstable grasp choices, as shown in Fig. 5.

3) Advanced Group: Some participants in the advanced group experimented with the robot's limitations. In some cases, participants grasped an object, then commanded the robot to bring the object to them. In this case, commands in the primitive command set were used to move objects from the table toward the participant's outstretched hand. This additional step suggests task planning by the advanced group.

This group, for the most part, did not deviate from the list of primitive commands provided. Most participants in this group did not ask many questions to further clarify either the given commands or the task in general. Questions asked were of a more technical nature with more interest placed on the robot's physical limitations. This group also appeared to be the most methodical of the three groups. Figure 5 shows that the advanced group recorded the lowest number of failed grasp attempts and unstable grasps of all the three groups. This group also appears to perform early hand preshaping slightly more often than the other two groups in the study.

C. Participant Feedback

Some participants expressed feeling a steep learning curve during the initial stages of the study. Although a practice session was provided, participants still felt that they needed to go through all five objects before feeling more confident in commanding the robot. For most participants, initial opinions changed from robots being intimidating to not being intimidating following individual interaction sessions. Views on programming robots, however, did not change for the most part. Most participants still felt that it was a challenge to effectively program the robot even after a successful interaction session. Some participants also stated that they would more likely use complex commands further away from the object to get closer to the object, but would still use the simple command set once the gripper was positioned closer to the object. These participants indicated that they would consider using complex commands more often if the movements were predictable, but stated that a collection of simple more straightforward commands would still be preferred.

V. Discussion and Future Work

Several important findings were revealed during the course of the study. First, while the study had a wide range of participants and objects, most participants were capable of finding a set of communication commands to accomplish the grasping tasks. This set was developed over time once the user became more comfortable with the robot. The set used by most participants consisted of a small number of simple commands, not advanced, more "intelligent" commands. This is a significant result since one of the major issues of human–robot interaction (HRI) is the parsing of the user's language into meaningful commands. If the user, as this study shows, prefers a small set of simple commands with limited impact, then designing a speech recognition system, for at least transferring grasping skills, could be much easier. Another interesting finding is the fact that users tend to use commands that they have experience with and have worked with before, regardless of the availability of more capable commands. It would be interesting to see if a long-term interaction study will support this finding, particularly as users transition from novice to expert users.

The study had several quantitative results such as: 1) the number of commands issued; 2) the total interaction times recorded; 3) percentage of failed grasp attempts; 4) percentage of unstable grasps chosen; and 5) percentage of users performing early hand preshaping. Furthermore, it also revealed that the user's background does play some role in the interaction process. Yet, despite these differences between the three groups, the participant's background does not, however, impact his/her ability to successfully instruct the robot to grasp various objects. To our knowledge, this is the first time that a study focused on grasping, and nonexpert users has shown this important finding.

Can these findings be generalized to other more complex grippers and robots? The gripper used in this study is a basic two-fingered parallel gripper. The grasps produced by this gripper (precision or tip grasps) can be easily duplicated by more complex multifingered hands. The objects grasped are small objects that are difficult to grasp. Thus, it is reasonable to say that the results in this study can be generalized to more complex grippers and objects as long as the required grasp is a precision or tip grasp, which one can say is the most common grasp for pick and place tasks. The degrees of freedom of the robot played no role in these findings, since all of the commands used were high-level commands.

A. Grasping Skill Transfer

Grasping is a two-stage operation. The first stage consists of reaching toward the object while orienting the hand in a particular pose, while the second stage starts once the fingers contact the object and the hand closes. The initial approaching hand pose is dependent on the grasping task and on the experience of the person grasping the same or similar objects before. Thus, in robotics, a proper grasping configuration must take into consideration the functionality of the object and the object's use, even when using multifingered hands. This study's aim is to develop a natural language interface that allows nontechnical users to command a robot to reach and grasp an object in the *same way as they would*. The resulting command sequence and final pose implicitly represents each user's grasping skills and preferences. If the robot can now learn from these interaction sessions of how to reach for and grasp other objects, this will lead to developing user-adaptive robots. But, much work remains. We have collected 60 records, where each represents the command sequence of one user grasping one object. These records along with other sensory data, like vision and tactile data, can then be used as input data to a learning system that can learn how to grasp objects from actual grasping experiments. Moussa [13], [14] tested such a grasping system in a simulated environment using an object set that included 28 everyday objects with very promising results. But, this work did not include any user-specific grasping configurations or preferences.

Our next step is to enhance this natural language interface by adding a learning component. Using the data collected, we plan to explore finding patterns of commands that could be used to predict the user's next command and the type of object grasped, thus leading to faster interaction between the user and the robot. Our view is that this could eventually lead to meta grasping commands like "grasp this object like you grasp a comb."

ACKNOWLEDGMENT

The authors would like to thank the participants who volunteered for the study.

REFERENCES

[1] H. Asoh, Y. Motomura, F. Asano, I. Hara, S. Hayamizu, K. Itou, T. Kurite, T. Matsui, N. Vlassis, R. Bunschoten, and B. Krose, "Jijo-2: An office robot that communicates and learns," *IEEE Intell. Syst.*, vol. 16, no. 5, pp. 46–55, Sep./Oct. 2001.
[2] J. Pineau, M. Montemerlo, M. Pollack, N. Roy, and S. Thrun, "Towards robotic assistants in nursing homes: Challenges and results," *Robot. Auton. Syst.*, vol. 42, no. 3–4, pp. 271–281, Mar. 2003.
[3] H. Huttenrauch, A. Green, K. Severinson-Eklundh, L. Oestreicher, and M. Norman, "Involving users in the design of a mobile robot," *IEEE Trans. Syst., Man Cybern.*, vol. 34, no. 2, pp. 113–124, May 2004.
[4] W. Burgard, A. Cremers, D. Fox, D. Hahnel, G. Lakemeyer, D. Schulz, W. Steiner, and S. Thrun, "Experiences with an interactive museum tour-guide robot," *Artif. Intell.*, vol. 114, no. 1–2, pp. 3–55, 1999.
[5] W. Bluethmann, R. Ambrose, M. Diftler, S. Askew, E. Huber, M. Goza, F. Rehnmark, C. Lovchik, and D. Magruder, "Robonaut: A robot designed to work with humans in space," *Auton. Robots*, vol. 14, no. 2–3, pp. 179–197, 2003.
[6] T. Fong, C. Thorpe, and C. Baur, "Robot, asker of questions," *Robot. Auton. Syst.*, vol. 42, no. 3–4, pp. 235–243, 2003.
[7] J. Chen and A. Zelinsky, "Programming by demonstration: Coping with suboptimal teaching actions," *Int. J. Robot. Res.*, vol. 22, no. 5, pp. 299–319, May 2003.
[8] M. Nicolescu and M. J. Mataric, "Linking perception and action in a control architecture for human–robot domains," presented at the 36th Hawaii Int. Conf. Syst. Sci. (HICSS), Hawaii Island, HI, Jan. 2003.
[9] S. Lauria, G. Bugmann, T. Kyriacou, and E. Klein, "Mobile robot programming using natural language," *Robot. Auton. Syst.*, vol. 38, no. 3–4, pp. 171–181, 2002.
[10] A. Billard, K. Dautenhahn, and G. Hayes, "Experiments on human–robot communication with Robota, an imitative learning and communicating doll robot," presented at the Socially Situated Intell. Workshop (SAB 1998), Zurich, Switzerland, Aug. 1998.
[11] C. Breazeal, A. Brooks, J. Gray, G. Hoffman, C. Kidd, H. Lee, J. Lieberman, A. Lockerd, and D. Mulanda, "Humanoid robots as cooperative partners for people," *Int. J. Humanoid Robots*, vol. 1, no. 2, pp. 1–34, 2004.
[12] P. McGuire, J. Fritsch, J. J. Steil, F. Rothling, G. A. Fink, S. Wachsmuth, G. Sagere, and H. Ritter, "Multi-modal human–machine communication for instructing robot grasping tasks," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS '2002)*, Lausanne, Switzerland, pp. 1082–1088.
[13] M. A. Moussa, "Combining expert neural networks using reinforcement feedback for learning primitive grasping behavior," *IEEE Trans. Neural Netw.*, vol. 15, no. 3, pp. 629–638, May 2004.
[14] M. Moussa, "Categorizing arbitrarily shaped objects based on grasping configurations," *Robot. Auton. Syst.*, vol. 54, pp. 858–863, 2006.
[15] M. Ralph and M. Moussa, "On the effect of the user's background on communicating grasping commands," in *Proc., ACM Conf. Human-Robot Interaction (HRI '2006)*, Salt Lake City, UT, pp. 353–354.
[16] J. Rubin, *Handbook of Usability Testing*. Hoboken, NJ: Wiley, 1994.
[17] M. Ralph and M. Moussa, "Human-robot interaction for robotic grasping: A pilot study," in *Proc., IEEE/RSJ Int. Conf. Int. Robots Syst. (IROS '2005)*, Edmonton, Canada, pp. 2146–2151.
[18] M. van Someren, Y. Barnard, and J. Sandberth, *The Think Aloud Method: A Practical Guide to Modelling Cognitive Processes*. London, U.K.: Academic, 1994.
[19] J. Gould, J. Conti, and T. Hovanyecz, "Composing letters with a simulated listening typewriter," *Commun. ACM*, vol. 26, no. 4, pp. 295–308, Apr. 1983.
[20] J. Kelley, "An empirical methodology for writing user-friendly natural language computer applications," in in *CHI '83: Proc. SIGCHI Conf. Human Factors in Comput. Syst.*, Boston, MA: ACM Press, pp. 193–196.

Disassembly Path Planning for Complex Articulated Objects

Juan Cortés, Léonard Jaillet, and Thierry Siméon

***Abstract*—Sampling-based path planning algorithms are powerful tools for computing constrained disassembly motions. This paper presents a variant of the Rapidly-exploring Random Tree (RRT) algorithm particularly devised for the disassembly of objects with articulated parts. Configuration parameters generally play two different roles in this type of problems: some of them are essential for the disassembly task, while others only need to move if they hinder the progress of the disassembly process. The proposed method is based on such a partition of the configuration parameters. Results show a remarkable performance improvement as compared to standard path planning techniques. The paper also shows practical applications of the presented algorithm in robotics and structural bioinformatics.**

***Index Terms*—Articulated mechanisms, disassembly paths, molecular interactions, path planning algorithms, robotic manipulation.**

I. INTRODUCTION

This paper[1] addresses the problem of automatically computing motions to disassemble objects. The problem can be formulated as a general path planning problem [2], [3] (see Section III). Indeed, path planning concepts and algorithms have been applied to solve different instances of the (dis)assembly planning problem (see Section II). The instance treated in this paper considers two objects, with

Manuscript received September 11, 2007; revised November 30, 2007. This paper was recommended for publication by Associate Editor O. Brock and Editor H. Arai upon evaluation of the reviewers' comments. This work was supported in part by the European Community under Contract IST 045359 "PHRIENDS," in part by the Région Midi-Pyrénées under Project "ALMA" of the Institut des Technologies Avancées en Sciences du Vivant (ITAV), and Project "AMOBIO" in the framework of the Communauté de Travail des Pyrénées (CTP), and in part by the French National Agency for Research (ANR) under Project "NanoBioMod."

The authors are with the Laboratoire d'Architecture et d'Analyse des Systmes, Centre National de la Recherche Scientifique (LAAS-CNRS), Université de Toulouse, 31077 Toulouse, France (e-mail: jcortes@laas.fr; ljaillet@laas.fr; nic@laas.fr).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TRO.2008.915464

[1]A preliminary version of this paper is appeared in [1].