Goodfellow, I., Koenig, N., Muja, M., Pantofaru, C., & Takayama, L. (2010). Help me help you: Interfaces for personal robots. Late breaking results of Human-Robot Interaction Conference: HRI 2010, Osaka, JP, 187-188.

Help Me Help You: Interfaces for Personal Robots

Ian J. Goodfellow, Nate Koenig, Marius Muja, Caroline Pantofaru, Alexander Sorokin, Leila Takayama
Willow Garage, Inc.
Menlo Park, California, USA
Email: {goodfellow, nkoenig, mariusm, pantofaru, sorokin, takayama} @willowgarage.com

***Index Terms*—HRI, mobile user interface, information theory**

I. Research Problem and A Proposal

The communication bottleneck between robots and people [1] presents an enormous challenge to the human-robot interaction community. Rather than exclusively focusing on improving robot object learning, task learning, and natural language understanding, we propose also designing interfaces that make up for low communication bandwidth by thoughtfully accounting for the constrained capabilities of robots [2].

People are adept at compensating for communication limitations, changing their communicative strategies for talking to pets, babies [3], foreigners [4], and robots [5]. Communicative accommodation already exists. Thus, instead of requiring robots to perfectly understand natural language, gestures, etc., there is a wide variety of research and design to be done in the space of alternative communicative modalities.

We propose to approach this problem by accounting for limitations in robot abilities and taking advantage of already familiar human-computer interaction models, leveraging a communication model based upon Information Theory. Using this design perspective, we present three different mobile user interfaces that were fully developed and implemented on a PR2 (Personal Robot 2) [6] for task domains in navigation, perception, learning and manipulation.

II. Relevant Theories

We can observe parallels between human robot interaction and the interaction between humans and general complex autonomous systems. Sheridan's taxonomy of complex human-machine systems describes the following sequence of operations: (1) acquire information, (2) analyze and display information, (3) decide on an action, and (4) implement that action [7, p. 61]. This provides the groundwork for identifying the stages at which people and/or robots should lead. In the current projects, the personal robot autonomously completes steps 1, 2 and 4, and the person completes step 3. Thus, the user interface design must address how the robot analyzes and displays its sensor information and world model to the human, and how the human can effectively communicate desired actions to the robot. An analysis of our case studies in Sheridan's framework is displayed in Fig. 1. All of our systems use the trading model of alternately passing control back and forth between human and robot, as opposed to the sharing model of simultaneous control described in [7, p. 63].

Gold proposed using an Information Pipeline model for HRI that is based upon information theory [8], a mathematical model of communication developed for quantifying the amount of information that could be transported through a given channel. Schramm [9] developed a theory of communication that put these ideas into the context of two-way joint communications. This could be helpful when considering the large amount of overhead involved in encoding and decoding messages sent between people and robots.

The focus of the projects in this paper was on designing interfaces that applied this theory to human-robot communication. With a robot encoding messages in a way that humans can understand and humans encoding messages in a way that robots can understand, communication is easy and effective.

III. The Design Space and Three UIs

The personal robot platform used throughout these projects is the PR2, and the robot behaviors are built using the Robot Operating System (ROS) [10]. The PR2 is a mobile robot standing approximately 5 feet high. It drives using casters in its base, has two long arms for manipulation, and has numerous sensors for perception. It is designed to operate in any environment that is American Disabilities Act compliant.

1) Navigation: The first task a personal robot will be asked to do is drive to a given location. The PR2 is capable of navigating a complicated, cluttered environment, but a person needs to tell it where to go. With a labeled map, a user can select a location from a set of options. By installing a PBX server on the PR2 and using DTMF codes, this prototype system allowed users to call the robot to tell it where to go with key presses corresponding to a menu (with items such as "drive to destination"), a language both the user and robot understand. The robot then executes the task and calls the user back to inform them of task completion. See Project 1, Figure 1 for an outline. Although we focused on the task of driving to a destination room, this framework facilitates a variety of interactions such as calling a user when the battery is low.

2) Perception: Another common scenario involves the user asking the robot to fetch objects, e.g., a drink. The location in the house of the general object class of beverages is known, and the PR2 is capable of finding bottles on a flat shelf. However identifying the correct flavor is difficult, especially if the drink selection changes. The robot may have never seen mango juice before, is unable to tell it apart from orange juice, and does not know what name to assign to each bottle. To bridge the gap, the robot photographs the available bottles, remembers their locations, and presents the photos to the user,

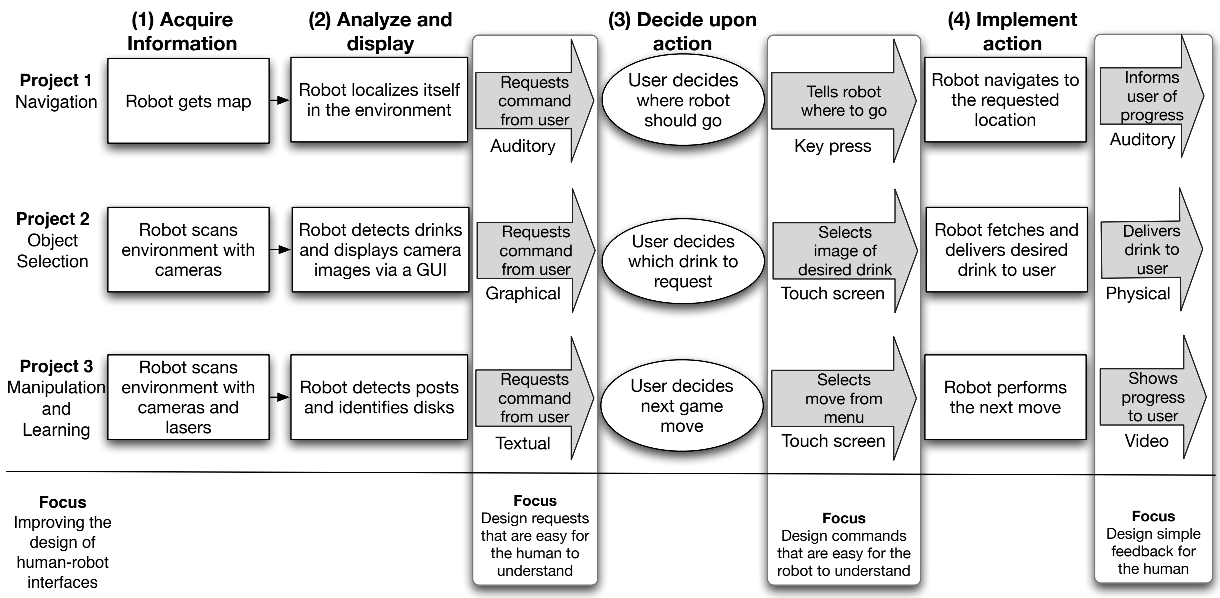

Fig. 1. Summary of the three projects, divided according to Sheridan's four stages of complex human-machine tasks [7].

who then taps the touchscreen to select the desired bottle image. This interface is outlined in Project 2, Figure 1. The key idea is that communication through pictures avoids the need to abstract ideas into language (e.g., [11]).

3) Learning and Manipulation: Although a personal robot deployed in a home will be pre-programmed to do many tasks, there will be unforeseen situations requiring new behaviors. This system allows a person to teach a robot a new hierarchical task, composed of basic pre-programmed actions. The user interface is crucial as the robot is encountering a new situation that it cannot describe. Instead, the robot suggests action primitives that it knows how to perform, guaranteeing that the human instructions are feasible. Thus the human comprehends the robot's capabilities, and formulates the high-level action sequence that the robot lacks. This prototype uses a mobile web interface to provide a selection of known action primitives to the user, and images from the robot's camera to provide feedback, as outlined in Project 3, Figure 1. With this interface, humans instructed a robot on how to solve the Towers of Hanoi puzzle. A user study of twenty end-users showed that this UI design eliminated the need for a long instructional period.

IV. Implications for Design and Future Work

These case studies represent a variety of task domains and system designs, but all UIs leverage existing technologies and established human-computer interactions. From the experience of designing, implementing, and using these systems, we learned several lessons that inform future designs. Start by considering the task, its constraints and available information. For example, since the PR2 knew office numbers and their locations, DTMF codes were sufficient for communication. Second, consider the limitations of the robot to divide work between the human and robot appropriately. Since the PR2 can detect bottles, but not specific juice flavors, it found and photographed the bottles while the customer recognized the images. Finally, design user interfaces that support division of labor. Because the Towers of Hanoi task only required a constrained set of manipulation subtasks, the interface presented a constrained action-object (noun-verb) model of communication instead of natural sentences.

The core requirement for a personal robot is to provide service to its user, and so a key capability for such a robot is to understand and communicate with its user. Through three case studies we have shown how effective communication accommodates for both robot and human abilities, making difficult communication tasks possible with current technologies. By using mobile web and phone technology, all three interfaces are available for already pervasive platforms. The source code for these projects is available at: http://code.ros.org.

References

[1] K. Gold, "An information pipeline model of human-robot interaction," in *HRI*, 2009.

[2] Y. Fernaeus, M. Jacobsson, S. Ljungblad, and L. Holmquist, "Are we living in a robot cargo cult?" in *HRI*, 2009.

[3] D. Burnham, C. Kitamura, and U. Vollmer-Conna, "What's new, Pussycat? On talking to babies and animals," in *Science*, 2002.

[4] R. Scarborough, J. Brenier, Y. Zhao, L. Hall-Lew, and O. Dmitrieva, "An acoustic study of real and imagined foreigner-directed speech," in *International Congress of Phonetic Sciences*, 2007.

[5] A. Batliner, S. Biersack, and S. Steidi, "The prosody of pet robot directed speech: Evidence from children," in *Proc. Speech Prosody*, 2006.

[6] Willow Garage, "Personal Robot 2 (PR2)," www.willowgarage.com.

[7] T. B. Sheridan, *Humans and Automation*. John Wiley, 2002.

[8] C. Shannon and W. Weaver, *Mathematical Theory of Communication*. University of Illinois Press, 2002.

[9] W. Schramm, *The Process and Effects of Mass Communication*. University of Illinois Press, 1955.

[10] M. Quigley, K. Conley, B. Gerkey, J. Faust, T. Foote, J. Leibs, R. Wheeler, and A. Ng, "ROS," in *ICRA*, 2009.

[11] M. Franklin, *The Universal Phrasebook*. Sterling, 2005.