

Magic Cards: A Paper Tag Interface for Implicit Robot Control

Shengdong Zhao^{1,3}, Koichi Nakamura^{1,2}, Kentaro Ishii¹, Takeo Igarashi^{1,2}

¹JST ERATO
Tokyo, Japan

²Department of Computer Science
The University of Tokyo

³Department of Computer Science
National University of Singapore

zhaosd@comp.nus.edu.sg; {nakamura, kenta}@designinterface.jp; takeo@acm.org

ABSTRACT

Typical Human Robot Interaction (HRI) assumes that the user explicitly interacts with robots. However, explicit control with robots can be unnecessary or even undesirable in certain cases, such as dealing with domestic services (or housework). In this paper, we propose an alternative strategy of interaction: the user implicitly controls a robot by issuing commands on corresponding real world objects and the environment. Robots then discover these commands and complete them in the background. We implemented a paper-tag-based interface to support such implicit robot control in a sensor-augmented home environment. Our initial user studies indicated that the paper-tag-based interface is particularly simple to use and provides users with flexibility in planning and controlling their housework tasks in a simulated home environment.

Author Keywords

Implicit Human Robot Interaction, Magic Cards, Paper-tag-based interface, Domestic service.

ACM Classification Keywords

H5.2. [Information interfaces and presentation] User Interfaces -Interaction styles; I.2.9 [Computing Method] Robotics -Commercial robots and applications.

INTRODUCTION

It has long been our dream to have things done "magically". In the West, the Brothers Grimm told a famous fairy tale about an honest and hardworking shoemaker getting magical help from two little dwarfs with his cobbling while sleeping [9]. In the East, an ancient Chinese story fantasized about a girl hidden inside an oyster shell, completing housework while her host was away [26].

The fairy tales may still sound like dreams, but recent advances in robot technologies have brought us closer to achieving this vision, especially when dealing with domestic services. Robots have already walked out of science fiction and factories, and have started to enter our

homes. For example, the vacuum robot, Roomba, has been on the market for seven years and reached millions of users worldwide [25]. Other domestic-service robots including Scooba (mopping), Robomower (lawn mowing), Dirt Dog (garage cleaning), Dressman (ironing), and Paro (nursing) have also appeared in the market place [25].

Although progress has been encouraging, especially regarding the positive user response generated while using Roomba [7, 8], we still have not reached the vision described earlier. Typical interfaces for robot control assume that users explicitly interact with robots. We acknowledge that explicit interactions between humans and robots are essential, especially when working with entertainment and social robots. However, in some other cases, such as dealing with domestic housework, having the tasks done is the ultimate goal, rather than interacting with the robots.

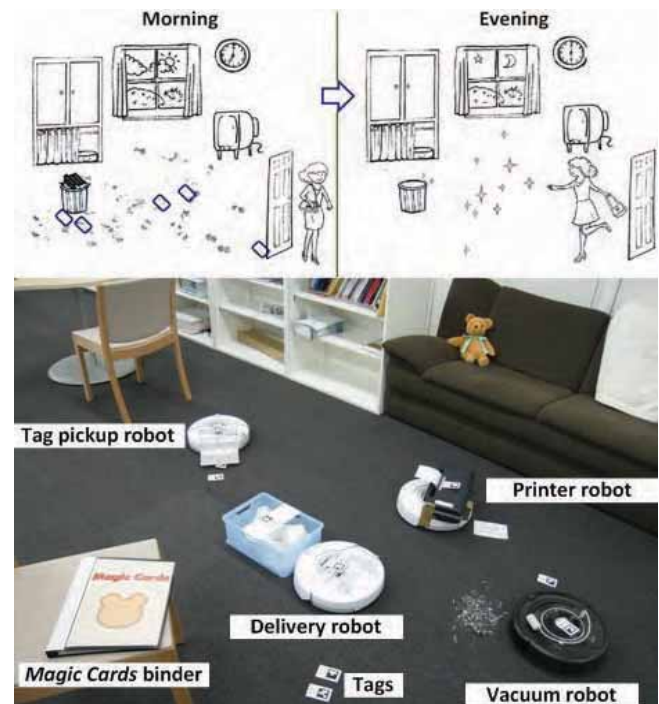


Fig. 1: Top of the figure envisions a user scenario of the proposed system with the *Magic Cards* interface. Bottom of the figure shows an implementation of the system with various robots in a simulated home environment.

This motivated us to investigate a new robot control paradigm where users *only* focus on task planning with real

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

CHI 2009, April 4–9, 2009, Boston, Massachusetts, USA.
Copyright 2009 ACM 978-1-60558-246-7/09/04...\$5.00.

world objects and the environment instead of interacting with robots. The planned tasks are later discovered and completed by robots in the background.

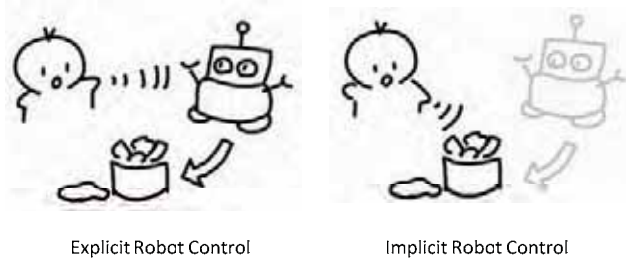


Fig. 2: Explicit robot control (left) and implicit robot control (right). Users explicitly interact with robot in traditional HRI approaches. In contrast, users only focus on tasks while a robot works in background in our approach.

We envision several advantages of this new approach. First, it is simpler to only deal with “tasks”. In traditional HRI approaches, the user would have needed to interact with both “robots” and “tasks”, while in our approach, “tasks” is the only concern.

Second, compared to robot technology, daily housework tasks have a much longer history and are relatively stable. Less change is likely to happen to the nature of tasks. In contrast, robot technology is still immature and rapidly evolving. Interactive techniques designed for specific robots are more likely to become obsolete in the future. In other words, focusing on tasks can better *encapsulate* users from pursuing ever-changing technology, leaving the enabling technology to evolve in the background without any active involvement by users.

Third, our approach encourages asynchronous operation: users assign tasks and have them completed later. Many of us spend significant amounts of time at work, in school, or enjoying various activities outside our homes. Ideally, tasks can be completed while we are away, sleeping, or in a different room, which leads to *magical* experiences.

To explore the above vision, we designed a simple paper-tag-based programming interface to describe daily housework tasks such as vacuuming or delivering an object to a specific location. Users place paper tags on corresponding real-world objects and the environment to assign tasks. A vision based tracking system then detects users’ instructions and controls the robot to complete these tasks in the background (Fig. 1). Paper tags are selected because they are tangible, simple to use, and “computer- and screen-less”, which allows the proposed solution to better reach the numerous housewives, elderly, and disabled users who are unfamiliar with (or even intimidated by) computers.

In our initial prototype, all robots are customized from iRobot Create and are currently only capable of completing tasks on the ground level; therefore, only limited types of tasks are supported. However, since the complexity of the system is hidden from users, our system can be easily

extended with more sophisticated robots which are capable of complex tasks in the future without much change to the *Magic Cards* interface. In the following sections, we first present several usage scenarios to provide an overview of how the system works from the users’ point of view. We then describe the related work, followed by a detailed explanation of the paper-tag user interface, the overall system architecture and its implementation, and the results of the various user studies before concluding the paper.

USAGE SCENARIO

Profile • John and Mary are a working couple without children. During weekdays, John and Mary wake up early and leave for work at 7 am. In the evening, Mary usually arrives home around 6:30 pm while John does not return until 8 to 9 pm. Mary is responsible for most of the housework in their two-bedroom apartment. Maintenance tasks, such as vacuuming and trash management, are typically done during the weekdays while major cleaning is left to weekends.

Setup • Now, let us see how the robot housework system with an interface of *Magic Cards* can assist their lives. The installation process is handled by professionals and involves mounting cameras on the ceiling and setting up the software. Object tags are placed on certain objects Mary wants the system to uniquely identify and handle, such as trash bins. Finally, she receives a set of paper tags organized in a binder to give instructions to the system with a simple tutorial on how to use it.

First day • In the evening, the system is ready. Mary starts to use it to plan housework tasks for the next day. One task she wants done tomorrow is vacuuming, since the installation process left dust and dirt behind. To have the robots carry out the task, she places a “Vacuum this room” tag and a “Start at noon” tag in each of the two rooms. In addition, she needs the trash bin to be placed at the door before 8 am the following morning for pickup. However, since more trash will be produced in the morning, she does not want it to be moved now. She also wants the empty trash bin to be returned to the original location in the evening. To do that, she places a “Take me to ‘destinations’” tag beside the trash bin, the “First destination” tag near the apartment door, and a “Start at 7:30 am” tag next to the deliver tag to specify the time for the act of delivery. The “Second destination” tag and a “Start at evening” tag are placed at the current location of the trash bin.

The next day, Mary goes to work at 7 am as usual, and when she returns home in the evening, she finds that the tags have been removed. The trash has been emptied and returned to its original location, and both rooms appear clean (Fig. 1).

Another day • One day, Mary is alone at home, plans to watch a movie in the living room, and she decides she wants the bedroom to be cleaned while she is watching the movie. She places a “Vacuum this room” and a “Do this

first" tag beside it. She then places a "Mop this room" tag and a "Do this second" tag on the ground. There is one especially dirty spot in the bedroom, so she places a "Mop this location" tag at that spot to ensure that location will be cleaned multiple times. She also has an expensive vase near the corner of the room and she is afraid that the robot's movements may accidentally break it, so she places an "Avoid this object" tag next to it. Finally, she places a "Start activities now" tag on the floor, and starts to watch the movie in the other room (Fig. 3).



Fig. 3: Visual illustration of user scenario for another day.

Compound task • Using the system for a while, Mary found some sets of tasks appeared repeatedly, which could be reused. To do that, after planning a set of tasks such as vacuuming first and then mopping, she places an additional "Memorize this set of tasks for future reuse" tag, and when these tasks are finished, a new "Memorized set of tasks" tag representing the combination of these tasks appears (this is produced automatically by a printer robot) and is ready to be reused for similar occasions in the future.

The above highlights the basic usage scenarios for our system and how it can improve people's lives. The most important aspect is that the system allows the robot's services to be fully controlled (what, when, and how the tasks are done) while the complexity of the system is hidden in the background. No menus, buttons, or even displays are needed. Placing paper tags at appropriate locations and housework will be done automatically.

RELATED WORK

While our research lies in the area of HRI, our motivation was rooted in the fact that most of us spend significant amounts of time on housework. According to the National Survey of Families and Households, the average woman spends 28 hrs per week on housework while the average

man spends 18 [2]. Excessive housework also has various objectionable effects [11, 17]. To most people, housework is still a tedious necessity rather than an enjoyable pastime.

The emergence of domestic service robots, such as the vacuum robot (Roomba) and other types of robots from iRobot (www.iRobot.com), points to a start in solving this problem in the future. Various studies in the HRI community focusing on Roomba and domestic service robots [7, 8, 25] have pointed out the potential of these robots as well as the challenges of introducing them into our homes. However, "homes are not simply designed to accommodate service-robot technologies, rather, domestic-service robots need to be designed to 'artfully integrate' with structures and practices of the home" [8, 27], which motivated the approach we took.

Both conceptually and technologically, our approach was inspired by the vision and practices of several fields in Human Computer Interaction (HCI), including "Ubiquitous Computing" [28], "Augmented Reality" (e.g., [5, 13, 19]), Tangible User Interface (TUIs) [12], and various work on paper-based interfaces [10, 13, 20, 23, 29].

"Ubiquitous computing" (UbiComp) [28] envisions a paradigm of "calm" computing that pushes computers into the background and attempts to make them invisible, and has led to a tremendous amount of research on context aware and location aware computing (e.g., [18]) in recent years. Recently, efforts have been made to introduce UbiComp into the home environment (e.g., Georgia Tech's aware home [14]). Due to the special role of the home environment and housework nature of these tasks, we felt it was even more important to shift away from the traditional HRI approach where humans directly interact with robots and move robots and sensors into the background.

To accomplish this vision, we leveraged tools and results from the Augmented Reality (AR) community, which has a long history of integrating real-world environments with computational media. Several AR programming libraries, such as the open source package of ARToolkit [1] and robust marker technology (e.g., Rekimoto and Ayatsuka's CyberCode [19] and Fiala's ARTag [5]) allowed us to build a system to support the kind of interactive style.

In designing interaction with the system, we were inspired by work on TUIs [12], which uses rich physical affordance of real world objects to assist interaction. In particular, the extensive work on using paper as a computer interface [10, 13, 20, 23, 29] led us to use paper cards as our interactive media.

Two projects in particular have influenced our design of the overall system: Collaborage [16] and Zombie board [22] from Xerox PARC, which themselves have been inspired by the earlier work of Insight Lab [15] and BrightBoard [24]. These systems use vision technology to detect either paper tags or diagrams drawn in ink to control the system.

Although these projects focused on whiteboard interactions, similar approaches can be adopted for our system.

While various work has inspired us, our research is nevertheless different. Compared with traditional *robot-centric* HRI approaches, our interaction focuses on tasks. Instead of interacting with robots, we use paper cards to interact with tasks and objects. The complexity of robots and the system are hidden in the background.

When compared with UbiComp, AR, and TUI, our research can be distinguished by using real-world artifacts to manipulate real-world objects and tasks. This is in contrast to the conventional approach of using specialized input devices (such as a mouse and keyboard) to manipulate objects in the virtual world (desktop computing), or using real-world environments or objects to augment or manipulate the digital world or digital objects (e.g., AR and TUI). Although the interaction regarding input in our approach is similar to both AR and TUI, the output or outcome of the system is different: instead of moving and manipulating bits and bytes, we move real objects using mobile robots.

Task-centric approach in tele-operation robots [4, 6] shares the goal of freeing the users from low-level control with our work. However, these systems still require the user to explicitly interact with the remote robot using a control console in front of the user. It assumes full attention of the operator during the task execution. Our goal is to make the user unaware of the robot as a physical entity. The user only leaves a command in the environment and tasks will be executed in the absence of the user.

PAPER TAG BASED INTERFACE

Several factors influence our design of the paper tag interface: 1) Since our target users include housewives, this requires the interface to be simple enough for them to understand and use. 2) To cater the diverging needs of individuals, the interface also needs to be flexible enough to allow planning tasks in a variety of contexts. 3) To allow easy incorporation of future robotics technologies while minimize unnecessary learning efforts from users, the interface needs to be stable and extensible.

Considering all these factors, we designed a paper-tag-based interface inspired by the grammar of natural languages (such as English). Paper is used because of its tangible nature and long history of use, which makes it easy to understand and be manipulated. English as a natural language has long been used by people to communicating about daily tasks. In English grammar, tasks are generally described using the subject, verb, object, and modifiers (e.g., in a simple task such as “Mary vacuums this room at noon”, the subject here is “Mary”, the verb is “vacuum”, the object is “this room’s floor”, and the modifier is “at noon”).

However, to focus on tasks and objects, the subject in the task description is intentionally avoided. Only “verb”, “object”, and “modifiers” are used in our approach and are

translated into *action tags*, *object tags*, and *modifier tags*, respectively. Under the new approach, the subject “Mary” from the previously task is omitted from the task description and the entire task is described as “this room needs to be vacuumed by noon” instead.

Tag design

The design of the layout of the cards is similar to that of Collaborage [16]. The paper tags in our system serve a dual purpose: they both inform humans so that they know how to assign tasks easily and accurately, and the system itself about the nature of the tasks and where exactly these tasks are located. Natural language and intuitive image icons are used to communicate with humans while 2-D id-markers are used to instruct the system. There is an example of this design in Fig. 4.

Note that 2-D planar id-markers are not the only solution to providing identity, position, and orientation information on objects as well as that on robots to the system. Other methods, including magnetic fields, radio, active LED, and laser beacons can also be used. Nevertheless, passive marker patterns that can be detected by computer vision were chosen because they are the most simple and inexpensive.

We used proprietary 2-D planar id-markers, which were very similar to those in earlier work such as CyberCode [19], and ARTag [5]. A marker consisted of a 3x3 black and white matrix pattern within a black border with a white margin around it (Fig. 4 top left). Each marker was about 5 x 5 cm, which we managed to recognize each stably using a 960x720 resolution ceiling camera (2.5 m high) covering a 2.5 x 2 m region on the floor. Our system could uniquely identify 120 patterns with their orientations. We observed that it worked robustly in various illumination environments, sufficient for our initial study.

Type of tags

Fig. 4 summarizes all tags supported by our current design. In addition to the previously-mentioned *action tags*, *object tags*, and *modifier tags*, a few special purpose tags are used to provide additional functions for the system. For convenience, in addition to organize tags into types, they were further divided into functional sets within types. Each set was labeled with a number in Fig. 4. A total of 14 functional tag sets are listed. We will use phrases of “set” + “a number” to refer to a particular tag set in our text (e.g., *set 1* refers to the “Vacuum this room” tag). Different types of tags were placed on different pages and highlighted with different colors. To make it easier for users, tags were organized in a binder.

Action tags

Ideally, action tags can include all types of housework such as ironing, folding, hanging clothes, cooking meals, washing dishes, trimming plants in the garden, and cleaning the toilet. However, many of these tasks are difficult today even for the most advanced robots. Since the purpose of this project was to explore the paper-tag-based implicit

robot control interface concept, we chose to implement and support several of the tasks (*vacuuming, mopping, delivery, and avoidance*) in our exploratory design. We envision that new robots with new action tags can easily be incorporated into the system.

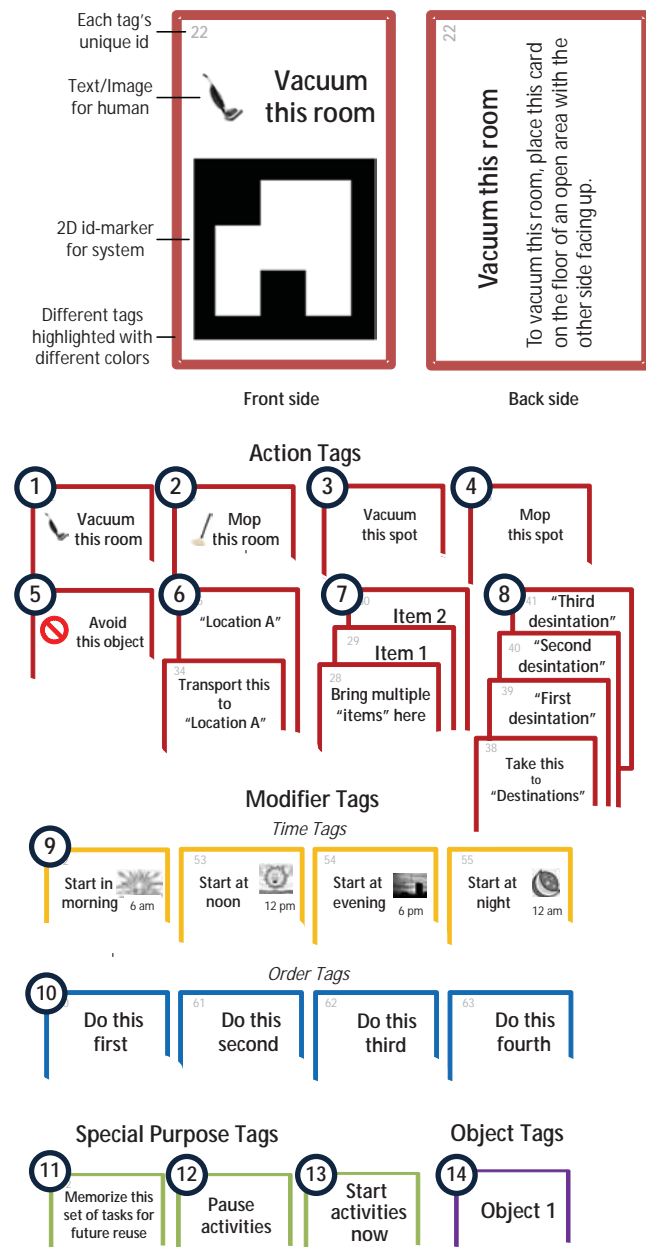


Fig. 4: Top of this figure shows front and backside of tag and highlights its different regions. Remaining figures show all tags for current prototype. Tags are first classified into 4 main types, and further divided into 14 functional sets within types.

“Vacuuming” and “mopping” actions are straightforward as indicated by their names. Two types of vacuuming and mopping are supported. *Vacuum/mop this room* will accomplish roomwise cleaning, while *vacuum/mop this spot* will instruct the robot to clean a spot specified by a user either as a separate event or combined with room-wise

cleaning to remove dirtier spots (Fig. 4, *set 1-4*). Delivery, where objects are moved by robots from the original place to required locations, can be classified into three types: 1) from one initial location to a destination (*one-to-one delivery*), 2) from one initial location to a number of destinations in sequence (*one-to-many delivery*), and 3) from a number of locations to one destination (*many-to-one delivery*).

Each of these delivery types follows examples from the real world. For example, “bring a chair from the bedroom to the living room” is an example of one-to-one delivery, “take the trash bin to the dumpster (location 1) in the morning and bring it back (location 2) in the afternoon” is an example of one-to many-delivery, and “collect all toys and put them in the toy bin” is an example of many-to-one delivery.

Since all three types of delivery involve more than one card, it is important for us to allow users to understand the three types of delivery behaviors.

One-to-one delivery is relatively straight-forward: the card associated with the starting position is labeled with “Transport this to ‘Location X’” where X is any alphabetical letter and the card for destination is labeled “Location X” where letter X matches the send card it is paired with (Fig. 4, *set 6*).

One-to-many delivery is more difficult, since it is not easy to intuitively express this to users. After several design iterations, we chose the following labels. The starting card reads “Take this to ‘Destinations’”, while each of the destination cards was labeled “First destination”, “Second destination”, “Third destination” and so on (Fig. 4, *set 8*). It uses the metaphor of riding a bus, where it will stop at different bus stops on the way.

Many-to-one delivery uses a different metaphor, but is easy to understand. It uses a “Bring multiple ‘items’ here” card to represent the designation, and “Item 1”, “Item 2”, “Item 3” and following cards to represent the items that need to be brought to the destination (Fig. 4, *set 7*).

Avoid tag: It often occurs that there are certain areas in the household one want robots to avoid, such as locations with fragile or expensive objects. By placing an “avoid this object” tag beside an object, the system will prevent any robots from moving close to it. In terms of functionality, it is similar to the “virtual wall” included in the Roomba package where it prevents Roomba from crossing certain points (Fig. 4, *set 5*).

Object tags

Object tags are arguably unnecessary, since action tags can implicitly indicate the location of objects and the system in theory can intelligently infer these objects by analyzing the visual images. However, current vision-based recognition systems are not perfect: automatically recognizing all objects in the environment is difficult, especially if these objects can be moved and more than one object can have

similar shapes and color. To allow unique identification and tracking of these objects, we choose to explicitly tag them with object tags (Fig. 4, *set 14*). These object tags are pre-registered and are printed on sticky paper so they can be easily applied to an object of interest. Note that this process only needs to be done once.

Modifier tags

Modifier tags define the context of action, specifying when, where, and how the action is performed. The current implementation supports three types of modifier tags: *time tags*, *location tags*, and *order tags*. Each of them is briefly described below.

Time tags • (Fig. 4, *set 9*) Time tags define when actions happen. If no time tags are specified, actions will start at a default time specified by the system. When interviewing our users, we found out that most of them typically had only a few important timeslots, and people found general terms such as “morning”, “noon”, “evening”, and “night” were more intuitive and easier to think about than precise terms such as 1:23 pm.

However, users often have different preferences, and to different people, their exact interpretation of the time concept such as “morning” or “evening” is different. To allow users to specify personalized time slots, we implemented a simple time-tag creation interface (see the accompanying video) where users could enter a convenient time and name and the system assigned a machine readable id to it. Users could print this out according to their needs.

Location tags • The location of an action is typically implicitly implied by the placement of the action tag in the physical world. However, certain actions (such as delivery) involve more than one location. Additional location tags are needed to specify target locations. Location tags are typically coupled with action tags and do not exist independently.

Order tag • (Fig. 4, *set 10*) Order tags help to specify the sequence of actions when several tasks are involved. For example, one can use order tags to ask the system to do vacuuming before mopping. We can use “Do this first”, “Do this second”, “Do this third” and so on to label the order tags.

Special purpose tags

In addition to the tags mentioned above, several special purpose tags are designed, including “memorize this set of tasks for future reuse” tag, “start activities now” tag, and “pause activities” tag.

“Memorize this set of tasks for future reuse” tag • (Fig. 4, *set 11*) So far, all the tags described have concerned simple tasks. However, it is important to allow users to group simple tasks into a higher-level compound task, and reuse it later. The challenge, however, is to come up with an intuitive design so that ordinary users (as opposed to computer scientists) can understand and use this in real life. Many design alternatives have been proposed, but many of

them did not pass our user study. However, one particular design passed our user test, which is described here.

In this approach, users no longer have to “compose” a compound task. They simply plan their tasks as usual. If they like their planning and want to use it later, they simply apply this card to tell the system to save the entire set of actions into a preset. To help the user to reuse this compound task, the system will print out a new tag with all the task descriptions printed on the back of the card, and this card can then be used as a handle to represent all the tasks the user wants to perform in a group. Note that this process can be done hierarchically, and a compound task that is made of other compound tasks is also possible.

Due to assistance from object tags, the system can uniquely remember objects involved in a compound task even if their location has changed. This provides some intelligence and robustness to the compound tasks. For example, if a compound task involves moving several objects to multiple locations, even if the initial object locations have been changed, the system can still find these objects and deliver them to the required locations specified by the compound task.

“Start activities now” tag • (Fig. 4, *set 13*) Although our system was designed for asynchronous interaction, this tag provides some real-time interactive capabilities. As soon as the system detects this tag, even if the image in the camera has not been stabilized, the system will start the activities immediately. This can be useful when immediate action is preferred.

“Pause activities” tag • (Fig. 4, *set 12*) We encouraged users to use robots while they were away from home. We did not preclude them from using them while at home. This was especially useful if some rooms were unoccupied and could be left for robots to do their work. In such cases, human hosts may occasionally want to interrupt robots’ activities if they want to use these rooms again. The “Pause activities” card was designed to serve this purpose. This card, together with the “Start tasks now” tag, once recognized by the system is executed immediately, regardless of whether any human motions are detected. Note that a “Resume activities” tag is not necessary since it can be implicitly indicated by removing the “Pause activities” card.

In summary, these are the basic cards supported by our system. They enable users to carry out various simple tasks in different configurations as well as compound tasks that are composed of simple or compound tasks. What we presented here obviously does not cover all the possible actions and configurations for performing domestic services; however, the overall design and the classification system provide a framework that can be extended in the future. In the next section, we will describe the overall system architecture as well as implementation details.

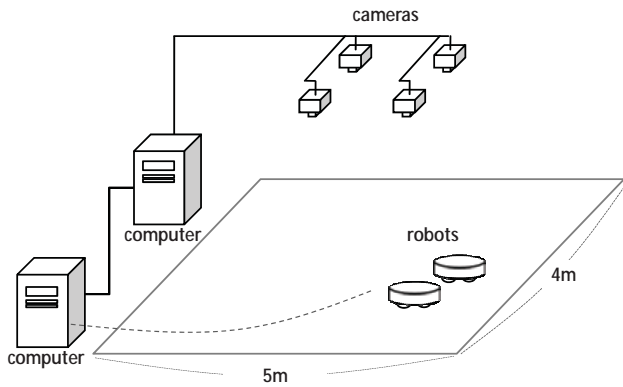


Fig. 5: Overall system architecture.

SYSTEM ARCHITECTURE AND IMPLEMENTATION

Overall system setup

The entire system hardware consists of sensors, computers, speakers, and robots (Fig. 5). The sensors are Logitech QuickCam® Pro for notebooks installed on the ceiling 2.5 meters above the floor. In the current implementation, they are connected to computers using extended USB cables. In the future, wireless connection is certainly preferred. Each camera covered an area of (2.5 x 2 m); with a combination of four cameras, a total area of 20 m² (5 x 4 m) was covered. Each camera had 960x780 resolution, and is capable of detecting markers with the size of 4.5 x 4.5 cm. The cameras are set up so that images slightly overlapped to allow image calibration and combination using the markers. The current implementation uses two computers connected via a TCP/IP network. The first has wireless receptors and is responsible for communicating with robots via Bluetooth.

Types of robots

All robots are products of iRobot Corporation. Of these, the vacuum robot (Roomba) and mopping robot (Scooba) are commercial robots (Fig. 6, top right) while the others are the modified use iRobot Create (Fig. 6, top left). According to their functionality, they can be roughly divided into working robots (robots that do actual housework) and administrative robots (robots that support other robots).

Working robots

Working robots include Roomba (vacuuming), Scooba (mopping), and delivery robots (iRobot Create), which do the vacuuming and mopping, and deliver objects. We turned off default automatic behavior (e.g., vacuuming while moving randomly) and directly controlled them using low-level control commands such as move forward and spin left. Vacuuming and mopping are relatively simple. The system simply moves the robot to the target location, turns on the vacuum or mop, and then moves it so that it covers the target area while avoiding obstacles. Delivery is a bit more involved. The system first moves the robot behind the target object and then moves the robot toward the target location. The system continuously tracks the position of the

robot and target object and adjusts its direction of movement when necessary.



Fig. 6: At top are photos of various robots from iRobot used in project. At bottom are card pickup robot and mobile printer robot modified using iRobot Create.

Administrative robots

The administrative robots include a printer robot and a card pickup robot. The printer robot is made of an iRobot Create robot with an HP Officejet H470B Mobile Printer attached to its back (Fig. 6, bottom left). The printer can be remotely controlled to print out card-size notes to be dropped on the ground by the iRobot Create. The printer robot has two uses as described earlier: 1) it could print out compound task tags as requested by the user and 2) it could also print out error messages to be dropped at the location where the error occurred.

The card pickup robot is also modified using an iRobot Create. In the current implementation, it is simply installed with a number of strong magnets in its extension and it picks up cards (with thin metal attached) using magnetic force as it moves over these cards (Fig. 6, bottom right). Due to this restriction, if users want to automatically pick up the cards, they can only be placed on the ground.

Working together

The system works in the following way. Using these tags is straightforward as described earlier in the usage scenarios. While a user is placing the tags, if none of the special tags that required immediate action were found (such as “Start activities now” or “Pause activities”), the system waits until the image stabilizes before instructing robots. This is similar to the design of the Collaborage system [16].

The recognition algorithm works as follows. The system needs to first group the cards into various tasks, and then prepare a plan for the robots to execute them. As previously described, a task may be defined by a number of cards representing the verb, object, and modifiers. Of these, the central piece is the “verb”, which is the action tag. The

system will first identify all the actions tags, and associate other cards with each of the action cards based on the distance to them.

Once all the tasks are defined and remembered, the system then brings out the card-pickup robot to collect all the cards, (except the “avoid object” card since the purpose of this card is to prevent any robots from getting close to this object, and collecting it using a robot defeats this purpose). After that, the system then brings out the appropriate robots to perform various tasks, such as vacuuming, mopping, and delivery.

If any errors occur during these, the system will remember the location (indicated by the robot, which is also equipped with traceable object tags), and once all the activities have finished, the system will bring out the printer robot to print an error message at the exact location where the error occurred. Finally, if it is a compound task, the system will ask the printer robot to print out a new tag and drop it at the same location as the “Memorize this set of tasks for future reuse” tag is dropped.

USER STUDIES

Early stage exploratory evaluation

Early in our design stage, in addition to analyzing results from the National Survey of Families and Households [2], and various survey results on domestic robots [7, 8, 25], we also conducted semi-structured interviews with six participants (one male and five females with ages ranging from 21 to 64, with a mean of 35.8. Three of them were married and one had children). These interviews taught us about their current living conditions and habits, and how they wanted to use our proposed system. A preliminary design for the paper tags were also presented to them for feedback. Interviews (each lasting about two hours) were conducted on separate days over the course of two weeks, during which we incorporated their feedback to iteratively refine the design of the paper tags.

This early exploratory evaluation helped us in many ways. First, it provided us with a better understanding of what people need and how our system can fit into their lives. Our interviewees helped us to confirm many of our initial ideas while clarifying some of our misperceptions.

All interviewees liked the asynchronous model of completing housework. Many of them spent a significant amount of time outside the home. They thought having their housework done while they were away was a great idea. We learned that most of them did both opportunistic cleaning as well as scheduled cleaning, consistent with the previous survey results [8]. Scheduled cleaning typically occurs on the weekend, while opportunistic cleaning can happen at any time.

We were originally more interested in spot-oriented cleaning, but our interviewees taught us that most people only undertake room-based or house-based vacuuming or mopping. Isolated spot-oriented cleaning occurs less and is

often dealt with in real time while our system is designed for asynchronous interaction; however, during the cleaning process, certain areas are often dirtier and require more cleaning, so it would be useful to ask robots to perform additional cleaning at these spots.

Second, the feedback from interviewees helped us greatly improve the design of our paper tags. Most of the tags (such as vacuuming, one-to-one delivery, and finish by morning tags) were conceptually simple and took only two to three iterations before an intuitive design was found and approved by all users. However, one type of tag, the compound-task tags, was much more challenging before we came up with an intuitive solution that could easily be understood by users.

These early explorations provided us with valuable feedback on the design of the *Magic Cards* interface. To find out the usability issues of the entire system, we also performed an evaluation with the working prototype, which is described below.

Evaluation of the working prototype

Participants

As soon as the system is implemented, 8 more participants (all females with ages ranging from 21 to 52, with a mean of 31.5) were invited to participate in the user study. Three of them are single; five are married; four are housewives; and three have children. Participants received \$20 per hour for taking part in our study.

Setup and procedure

The study involves three steps. 1) *Questionnaire* 2) *Paper tag evaluation* 3) *Using paper tags to complete tasks in a simulated home environment*. Hardware and software setup is exactly as described earlier. *Questionnaire* collects background information such as demographic information, living condition and habits, housework duties and types of work they want robots to help. *Paper tag evaluation* aims to further test the intuitiveness of the supported tags. Users were presented with 13 sets of tags (except object tags) after a brief introduction of the general purpose of the project. They were then asked to write down the meaning and usage of these tags based on the text and image presented on the front side of a tag, text from the backside was not provided. Once they finished, they proceeded to step 3 and were asked to use these tags to complete a few tasks in the simulated home (Fig. 1). Instructions to tasks were given in everyday language (e.g., “moving this box there”) with words directly from tag names purposely avoided. They started with simple tasks followed by more advanced tasks. Simple tasks included: vacuum, mop, and various types of delivery. Advanced tasks included combinations of simple tasks in different time and order. To save time, time based interaction were accelerated using a much short interval. After going through the planned tasks, users were asked to suggest their own tasks using the provided tags. All three steps, including break time, lasted about two hours for each participant.

Results and observations

Overall, users did very well in understanding the meaning of tags in step 2. 6 out of 8 users correctly interpreted the meaning of all 13 groups of cards. Two users misinterpreted the meaning for tag *set 8* (*one-to-many* delivery), and one of them also did not understand *set 11* (compound task). Their misinterpretations were not far from the correct meaning. One user interpreted *set 8* as distributing multiple mails to multiple persons, while the other person thought it was used to deliver an object to one of the destinations, instead all destinations in sequence. The person who did not understand the meaning of compound task tag simply wrote that she cannot imagine a scenario of how to use it.

All participants also did a great job in planning the tasks using paper tags in the simulated home environment. Interestingly, users found advanced tasks not more difficult to plan than simple tasks. All of them mentioned that it was straightforward to use them.

Users created some interesting compound tasks with the tags. For example, one subject suggested an interesting cleaning routine by combining delivery tags, vacuum-spot tags, and order tags. She first placed a delivery tag to move the toy bin to the side, a vacuum-spot tag to clean the area previously covered by the toy bin, and another delivery tag to move the toy bin back to the original location.

In the post experiment interview, users' response for the paper tag interface was very positive. All users highly applauded the simplicity of our paper-tag based interface. Many participants suggested that our interface is simple enough for elderly to use. This is further indicated by positive feedback received from the oldest participant in this study: a 52-year old housewife, self described as "terrible at any kind of instructions". She highly praised the simplicity and intuitiveness of the paper-tag-based interface which allows her to plan tasks without hitting a single button. She was self-described as hesitant to touch any buttons in fear of any undesirable consequences, but much more comfortable to handle and place paper cards because they seemed safe and harmless. She envisioned that our system would be useful for elderly and disabled to remain independence and perform housework by themselves.

While being aware of the limitations of the current system, users were surprised about the versatility of the system. By combining several tags, many potential useful tasks could be planned easily. Many expressed high interest in purchasing the system if it becomes available.

Compared with the positive responses, perhaps we, as researchers, were more interested in any problems and concerns found from the user study since these will lead to future improvements. One concern with the paper-tag interface was possible misuse by small children, as pointed out by the two users with small children. To them, cards cannot be left on the ground unattended. Cards also need to be big enough so their children cannot easily swallow them, and need to be made of harmless material such as soft

plastic to avoid any safety hazard. It is often essential to immediately collect these cards after planning, and as an extra precaution, they wanted to ensure all cards were collected before leaving home. In addition, they preferred to have the robots work while their children were not at home.

The use of multiple robots to complete different tasks is often a concern for users living in small apartments or houses. They prefer an all-encompassing robot that is capable of doing all tasks instead of using multiple robots. The use of many cameras in the environment also raises concerns for users who rent their abodes since mounting cameras to the ceiling may not be permitted by their landlords, and for aesthetics, they prefer cameras to be invisible. Instead of mounting cameras within the environment, future systems may consider using non-vision-based sensors or using robots equipped with a vision system that is capable of detecting tags in the environment.

During the course of the study, we experienced some surprising errors. Roomba may get stuck in an enclosed area or accidentally bump into another robot and cause the robot to reset itself. Communication with robots is via a Bluetooth connection and this may break down unexpectedly if battery power is running low. Considering that all these errors happened in a controlled environment, errors are unavoidable in more complex real-life situations. Currently, these errors are handled by more experienced technical people. How to implement an automatic error-recovery system (possibly by creating specialized rescue robots) can be an interesting and promising research topic for the future.

LIMITATIONS

As an exploratory project, this nevertheless had a number of limitations. The current system uses id-markers and computer vision to track these tags, which makes it vulnerable to occlusion problems as well as illumination conditions. The subjects need to be careful where they place the paper tags and when the tasks can be done (e.g., the current system does not work at night without lights). In the future, paper tags may be replaced by more advanced technologies such as Ubisense [3].

The robots used in this project are also primitive. These robots can only move on the ground level with a smooth surface. They have limited power and cannot support the delivery of large and heavy objects. Many advanced robots that are capable of walking, climbing the stairs, or imitating various human behaviors are being developed or already exist in laboratories or factories. These can be used to enhance our system to complete more interesting and complex tasks in the future.

While encouraged by the overall positive feedback, the raised concerns reconciled the earlier point that homes are not simply designed to accommodate service robot technologies; they need to be 'artfully integrate' with structures and practices of the home. While our user studies gathered initial evidence, further investigation in real home

environment for a prolonged period of time is necessary in order to fully understand the impact of such systems to people's lives.

CONCLUSION

In conclusion, we presented an implicit robot control interface using paper tags (cards) to systematically manage the robots to complete housework in a sensor augmented home environment. An initial prototype system that supported a limited number of housework tasks on smooth grounded surfaces was developed to explore this concept. Our user study has shown that our paper tag based interface is particularly simple to use, and provide users with flexibility in planning and controlling their housework tasks in a simulated home environment. The overall positive response from our users has indicated great potential for our paper-tag-based interface approach to design, and it arguably has a better chance of surviving over time because the design is based on more stable daily housework tasks.

Although our system can be used by everyone, particular interests have been raised about the potential benefit it can bring to the elderly and disabled. Instead of relying on outside assists, systems like ours can help them to regain independence by engaging [21] in housework themselves. The simplicity of the paper-tag-based interface may help domestic service robot technologies to be used by people previously difficult to reach. We hope our exploratory work can somehow accelerate the attainment of our dreams to complete housework tasks "magically" in the near future.

ACKNOWLEDGMENTS

We thank Minghui Sun for programming and helping make the video, Yoshiki Takeoka, Sorahiko Nukatani, and Tak Miyake for helping make the video, Zoey Yu for drawing figures and helping make the video, Yotam Gingold and other JST ERATO project staff and researchers for providing suggestions and help, and anonymous reviewers for their valuable and constructive feedback.

REFERENCES

1. ARToolKit. <http://www.hitl.washington.edu/artoolkit/>
2. The National Survey of Families and Household. <http://www.ssc.wisc.edu/nsfh/>.
3. Ubisense. <http://www.ubisense.net>.
4. Few DA, Bruemmer DJ, and Walton MC. Improved human-robot teaming through facilitated initiative. *Proc. RO-MAN '06*, (2006), 171-176.
5. Fiala M. ARTag, a fiducial marker system using digital techniques. *Proc. CVPR '05*; (2005), 2.
6. Fong T and Thorpe C. Vehicle teleoperation interfaces. *Autonomous robots* 2004; 11(1): 9-18.
7. Forlizzi J. How robotic products become social products: an ethnographic study of cleaning in the home. *Proc. HRI '07*, ACM Press (2007), 129-136.
8. Forlizzi J and DiSalvo C. Service robots in the domestic environment: a study of the roomba vacuum in the home. *Proc. of HRI '06*, ACM Press (2006), 258-265.
9. Grimm J and Grimm W. The elves and the shoemaker. *Fairy tales*.
10. Heiner JM, Hudson SE, and Tanaka K, Linking and messaging from real paper in the paper PDA, *Proc. UIST '99*, ACM Press, (1999), 179-186.
11. Hersch J and Stratton LS. Housework and wages. *The Journal of Human Resources* 2002; 37(1): 217-229.
12. Ishii H and Ullmer B, Tangible bits: towards seamless interfaces between people, bits and atoms, *Proc. CHI '97*, ACM Press, (1997), 234-241.
13. Johnson W, et al. Bridging the paper and electronic worlds: the paper user interface. *Proc. CHI '93*, ACM Press, (1993), 507-512.
14. Kidd CD, et al. The aware home: A living laboratory for ubiquitous computing research. *Proc. CoBuild '99*, (1999).
15. Lange BM, Jones MA, and Meyers JL, Insight lab: an immersive team environment linking paper, displays, and data, *Proc. CHI '98*, ACM Press, (1998), 550-557.
16. Moran TP, et al., Design and technology for Collaborage: collaborative collages of information on physical walls, *Proc. UIST '99*, ACM Press, (1999), 197-206.
17. Olson JT. Role conflict between housework and child care. *Work and Occupations* 1979; 6(4): 430-456.
18. Patel SN, Truong KN, and Abowd GD. PowerLine positioning: A practical sub-room-level indoor location system for domestic use. *Proc. Ubicomp '06*, (2006), 441-458.
19. Rekimoto J and Ayatsuka Y, CyberCode: designing augmented reality environments with visual tags, *Proc. DARE '00*, ACM Press, (2000), 1-10.
20. Robinson JA and Robertson C. The LivePaper system: augmenting paper on an enhanced tabletop. *Computers & Graphics* 2001; 25(5): 731-743.
21. Rogers Y. Moving on from Weiser's vision of calm computing: Engaging UbiComp experience. *Proc. Ubicomp '06*, (2006), 404-421.
22. Saund E. Image mosaicing and a diagrammatic user interface for an office whiteboard scanner. *Proc. CoBuild '99*, (1999).
23. Siio, I., Mima, Y. IconStickers: Converting computer icons into real paper icons. *Proc. HCI International '99*, (1999), 271-275.
24. Stafford-Fraser Q and Robinson P, BrightBoard: A video-augmented environment, *Proc. CHI '96*, ACM Press, (1996), 134-141.
25. Sung J-Y, et al. Housewives or technophiles?: understanding domestic robot owners. *Proc. HRI '08*, ACM Press, (2008), 129-136.
26. Tao Q, Oyster girl, *Shou shen hou ji*.
27. Taylor AS and Swan L, Artful systems in the home, *Proc. CHI '05*, ACM Press, (2005), 641-650.
28. Weiser M. The computer for the 21st century. *Scientific American* 1991; 265(3): 94-104.
29. Wellner P. Interacting with paper on the digital desk. *Communications of the ACM* 1993; 36(7): 87-96.