Socially Distributed Perception

Marek Michalowski, Carl DiSalvo, Didac Busquets, Laura M. Hiatt, Nicholas Melchior, Reid Simmons
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, PA 15213

Selma Sabanovic
Rensselaer Polytechnic Institute
110 8th Street
Troy, NY 12180

ABSTRACT

This paper presents a robot search task (*social tag*) that uses social interaction, in the form of asking for help, as an integral component of task completion. We define *socially distributed perception* as a robot's ability to augment its limited sensory capacities through social interaction. We review previous work in robots that are helped by humans, social situatedness, and socially distributed cognition. We discuss the task of social tag and its implementation on GRACE for the AAAI 2005 Robot Exhibition. We then present observations and results that suggest that we were successful in promoting a form of social interaction that allowed people to help the robot achieve its goal. Finally, we discuss the implications of this design approach for effective and compelling human-robot interaction.

Categories and Subject Descriptors

I.2.9 [**Artificial Intelligence**]: Robotics—*Operator interfaces, Sensors*; K.4.0 [**Computers and Society**]: General

General Terms

Design, Human Factors

Keywords

Human-robot interaction, Social robotics, Mixed initiative

1. INTRODUCTION

We often take routine activities for granted without realizing the complexity and amount of social interaction necessary to achieve the simplest of goals. Consider this example: A researcher travels to a new city to attend a conference. He makes plans to meet a former colleague for lunch. She tells him by email that she will be in the hotel lobby at noon wearing a blue coat. Arriving several minutes late, the researcher cannot find his colleague. He asks the hotel staff if they have seen a woman wearing a blue coat. The staff respond that indeed they have and point him in the direction they last saw her. But after looking around, the researcher still cannot find his colleague. So he continues to approach and ask strangers if they have seen a woman in a blue coat until, after following several sets of directions, he locates his colleague. In this typical situation, the researcher's sensory information about the world is insufficient to solve his problem, as he cannot visually locate his colleague. However, he can use other abilities, such as being able to interact with other people who might have seen her, to supplement his own spatial and temporal knowledge. This is an example of how, in everyday life, "social interaction allows humans to exploit other humans for assistance, teaching and knowledge" [2].

Figure 1: The robot GRACE.

A similar scenario was played out by the robot GRACE (Graduate Robot Attending a ConferencE) (Fig. 1) at the AAAI 2005 Robot Exhibition. GRACE played a game of *social tag* in which the task was to locate and rendezvous with a team member who was wearing a pink hat [12]. In this game, our purpose was not to create an object localization task (such as a scavenger hunt); rather, we wanted to create a task that would require a "socially situated" robot [20] that could enlist the help of humans through frequent social interactions. We designed the game of social tag so that the robot's information about the whereabouts of the team member came not primarily through the modalities of vision or acoustics, but rather through social interactions

Human Robot Interaction 2006, Salt Lake City, Utah, USA

with strangers in the environment. The task explores issues in human-robot interaction that involve shared space, intuitive interface design, and negotiation of an environment filled with "strangers" – individuals who have not been explicitly trained to interact with the robot.

Most machine perception systems are still severely limited compared to the sensory abilities of humans. Our project was motivated by the idea that robots built to operate in human environments (which are usually well-suited for human perception) would benefit from an ability to off-load sensory demands to human partners. It is up to a robot's designers to ensure that a robot is capable of *requesting* this help and *accepting* it when given. We introduce the term *socially distributed perception* to describe a robot's ability to augment its own perception through social interactions with people (and possibly other robots) in the environment. Such a robot is perceptually embedded within a social as well as a physical environment and can treat both the social and the physical characteristics of that environment as affordances [5]. In order for the robot to take advantage of these social affordances, it must be able to engage in social interaction with people.

In section 2 we review research in designing robots that can be helped by people. We then discuss two bodies of work, social situatedness and socially distributed cognition, that can be used to frame such research. It is from a union of these two bodies of work that we propose socially distributed perception as a new approach. In section 3 we review the robotic platform on which our work was implemented. In section 4 we introduce and discuss social tag as a task framework that makes use of socially distributed perception. We maintain that this framework is appropriate beyond our general game task and applicable to a wide range of activities for robots used in everyday dynamic environments. We then describe the system we developed to perform this task at AAAI 2005. In section 5 we present observations and results that suggest that we were successful in this respect, and describe how interactions between GRACE and conference participants were shaped by different social and physical environments. In section 6 we discuss socially distributed perception as a novel design principle that suggests new ways of designing appropriate and compelling social interactions between humans and robots.

Our team consisted of roboticists, designers, and social scientists. The unique multidisciplinary nature of this group shaped the development of the task and the accompanying robotic system. The importance of design considerations, the principles of social interaction, and the feasibility of implementing robotic behaviors were taken into account as the robot was designed, tested, and modified. Finally, the observation and analysis of the resulting system was performed with these different perspectives in mind.

2. BACKGROUND

2.1 Humans helping robots

Because of the broad and emergent nature of the field of human-robot interaction, related work appears in a variety of domains and disciplines. Within the existing robotics literature, similar work that relies on humans providing spatial directions is found in the area of deictic gestures (pointing) [18] and telerobotics [16]. A number of robots have been designed to appear "young" in order to evoke in humans the desire to act in the capacity of an assistive caregiver [1] [10].

Recently, research has been done in allowing robots to request human assistance in cases where the robot's capabilities are insufficient for task completion. "Sliding autonomy" refers to a form of human-robot collaboration in which control is transferred between autonomous robot operation and human teleoperation in order to complete a task such as remote construction [7]. When the transfer of control can be initiated by either party, this is known as "mixed initiative." Our work differs in that GRACE is always autonomous and is never controlled by a human being. Rather, she is dependent on humans for help; this is discussed in Section 6. Additionally, sliding autonomy typically involves robots working remotely from their human operators, therefore precluding true social interactions. "Mixed-initiative interaction" is also used to describe the design of interfaces or software agents that are expected to aid in the execution of information-processing tasks [6]. This work is similar to sliding autonomy in that a human and a machine can each contribute what they are best suited to do at the appropriate time. In our work, we consider a form of mixed-initiative interaction in which the machine is being helped by humans rather than vice versa, and the goal is to initiate a helpful social interaction. This interaction can be initiated either by the robot or by humans in the environment.

Another approach is to combine deictic human input with sensors embedded and spatially distributed throughout the environment [9]. While such a solution works for relatively stable environments, this approach is not sufficient to support tasks that might be assigned to robots in scenarios where the environment is either unknown prior to encountering it or is dynamically changing. Our project, on the other hand, considers people to be sources of information similar to distributed sensors. This can be understood in terms of a framework such as socially situated action and perception [20], within which the environment of objects, artifacts and other actors are used by the robot as resources that make intelligent action possible or as affordances available to the robot in its environment [5].

The theme of "humans helping robots" can be theoretically framed by the methodologies of social situatedness and socially distributed cognition. Our notion of socially distributed perception draws from these two bodies of work.

2.2 Social Situatedness and Distributed Cognition

"Situated robotics" [11] is based on the notion that complex environments have a strong influence on the nature and complexity of the behavior of embodied robots that are embedded in them. In social tag, GRACE is socially situated, or (inter)actively embedded, in an environment that is populated by humans, which is therefore social as well as physical [20]. The notion of a "socially situated agent" [4] implies that a robot interacts both socially and physically with the environment in order to acquire information about the social and physical domains. We relax this definition slightly: while GRACE was designed to locate people, participate in the social domain, and use people's sociability as a resource in accomplishing her task, the robot uses social interaction to gain information only about a physical aspect of the environment.

In socially distributed cognition, group activities and work are focused on a task and distributed over a range of media, artifacts and people [8]. In social tag, people participated in the search task along with GRACE, looking for and consulting each other about the whereabouts of the person in the pink hat. People can therefore be seen as partners in GRACE's search for the pink hat rather than mere "users." GRACE and her human partners constitute a system of socially distributed perception.

Situated actions are taken in the context of particular, concrete circumstances [20], and distributed cognition occurs in the context of socially, culturally and historically specific task structures. Interaction between machines and people in both frameworks implies mutual intelligibility and understanding. In social tag, participants are engaged through the familiar narrative of a search for a friend, which makes the simple engineering task take on a meaningful significance. The robot itself can be seen as a "social object" because it is reacting to people and has a goal that is understandable to them, which makes it seem purposeful to users [22] [20].

Drawing on and extending the methodologies of social situatedness and distributed cognition, our work aims to create a socially situated robot that performs a task in the context of a large social group by making use of the perceptual abilities distributed across the members of this group. Rather than attempting to create a system that uses socially distributed cognition to solve a complex task, we consider "perception" to be a component of "cognition" that is simple enough to distribute socially yet compelling enough to enable basic information exchange and to provide interesting behaviors for observation and research. Our hope was that an otherwise minimal social presence and engagement with humans would minimize the opportunity for system error or inappropriate human attributions of the robot's capabilities. The extent of GRACE's social interaction was therefore limited to a mixed-initiative form of requesting and accepting assistance (i.e., either the robot or a human could initiate an interaction that advanced the robot's progress). We implemented a system for the robot GRACE to use socially distributed perception to perform the task of social tag.

3. GRACE

GRACE grew out of a multi-institution collaboration to design a robot capable of performing the AAAI Robot Challenge, which involves autonomously registering for the annual AAAI conference, navigating through the conference area, interacting with people, and delivering a talk. GRACE performed most of these tasks at AAAI 2002 [19] [3].

GRACE is a RWI B21 mobile robot. For this task, she was equipped with a SICK laser scanner, a Canon VC-C4 PTZ camera, an LCD monitor with an animated face, and an ELO 1224 LCD touchscreen. The robot has two computers on board: one for controlling mobility, sensing with the laser, avoiding obstacles, and handling the touchscreen interface, and the other for vision, control of the face and voice, and general task control. The software architecture follows the design used in previous years for our entries in the Robot Challenge. Independent processes communicate via IPC message passing (http://www.cs.cmu.edu/∼IPC). No remote communications were used except for startup and shutdown.

4. SOCIAL TAG

The task of social tag – finding the person with the pink hat – was conceived as having a number of important benefits. The task is relatively simple and can be executed reliably, which allows us to focus on questions of human-robot interaction. Formulating the task as a fun and commonly-understood game also increases the chances that people will be willing to interact with the robot. While these properties were important for designing and observing human-robot interaction, the interaction's ultimate purpose is to locate the pink hat. The generality of object localization is another important property of our selected task, as it involves search, planning, navigation, etc., and can be applied to a wide range of useful robotic tasks.

In typical robot tasks that involve detecting a visual target, distinctively colored objects such as pink hats are used to simplify the machine perception problem. However, even with this simplification, in large crowded rooms the task of finding a hat is still extremely difficult. While research in computer vision and speech recognition continues to improve machine perception, GRACE relies on the assistance of humans with fully developed senses of sight and hearing and the ability to communicate. In this scenario, the pink hat was intended to be as much for the benefit of other people as for GRACE herself. The team member was to be a highly visible individual who would be easily recognized and remembered by conference participants so that it would be easy for them to help a wandering robot. Accordingly, while GRACE depends on her own sense of vision to achieve the goal (i.e. recognizing the pink hat), her primary mode of gathering information is asking people for help in an intuitive and socially acceptable manner. We are more interested in this journey, and in the social interactions that occur along the way, than in the completion of the task itself.

The task has five main phases: *identification* of approachable humans; *approach* toward a human with whom the robot would like to interact; *asking* for directions to the person with the pink hat; *following* those directions until a pink hat is found visually or more help is required; and demonstrating *success* when the hat has been found. As in typical human interactions, these phases may be regarded as a "script" that suggests an appropriate plan of action. However, humans may deviate from this script by volunteering assistance without being asked. Our form of mixed-initiative interaction allows GRACE to handle such deviations in a context-dependent manner.

4.1 Interaction and interface design

An LCD touchscreen mediated the interactions between GRACE and conference attendees. The touchscreen is mounted on the front of GRACE, below the screen that displays her face, at approximately chest level with an average-sized adult. Touchscreens have been found to be a simple and intuitive medium for transferring information to socially interactive mobile robots [17] [13]. This mode of interaction was selected over modes that involve natural language (either spoken or typed) for reasons of reliability and simplicity.

The touchscreen interaction consists of a number of full-screen interfaces that both convey the state of the robot and prompt certain types of input. The general approach to designing the screens was to make them as bold and simple as possible (so as not to burden participants) while still being aesthetically appealing. To this end, we limited the amount of text on the screen and supplemented the screens with spoken information (via a text-to-speech system with lip synchronization on the animated face). In order that GRACE's verbalizations do not become repetitive or boring, she chooses randomly from a library of appropriate phrases at each point in the task. Music is another important component of GRACE's repertoire. Appropriate sound clips were played throughout the phases of the task. In addition to the music adding to the playful character of GRACE, it also functions as another mode of (auditory) expression to participants, emphasizing the current state of the robot.

4.2 Task phases

The robot is controlled by a simple state machine (Fig. 2), where states correspond with the phases of the task as mentioned above. We now describe these phases and discuss significant implementation details.

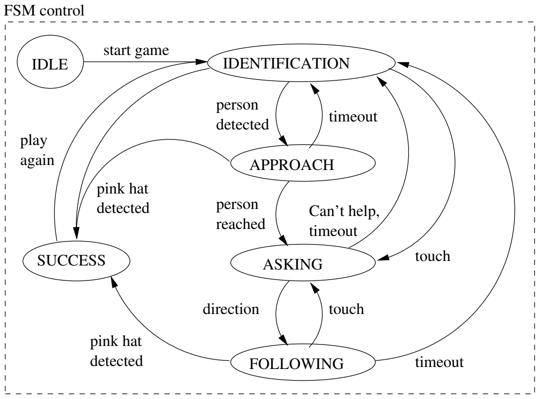

Figure 2: The finite state machine that comprises GRACE's control task for Social Tag.

4.2.1 *Identification*

When the game is started, GRACE is in the `IDENTIFICATION` state. She is looking for the pink hat or for a person that might help her find it. The touchscreen displays the *wandering* image (Fig. 3), which depicts a pink hat, question marks, and the phrase "Touch Me. I am looking for a pink hat." If the screen is touched, GRACE transitions to the `ASKING` state. The *wandering* screen serves three purposes: it informs people about what GRACE is doing; it provides an opportunity to interact with the robot; and it mitigates the limitations of the vision system by not relying on it to find people. Meanwhile, GRACE plays music and periodically says phrases such as "Where is the person in the pink hat?" or "I need to find that person in the pink hat!" Her animated face frequently changes direction, giving the appearance of actively looking around while wandering.

GRACE is equipped with a laser scanner near human knee-height and a camera near human face-height. The laser scanner clusters short adjacent range readings, labels those that appear to be human beings, and tracks those objects over time using a Kalman filter. The camera locates faces using appearance-based frontal face detectors and tracks them using skin color models. Data from these two sensors are combined to determine more reliably where people are located. This is done by registering the locations of camera-located faces and laser-located obstacles with each other in a robot-centric coordinate frame and labeling a laser obstacle as a person according to whether there is a face located above it. When a person is detected, the robot enters the `APPROACH` state.

A color model was obtained for the pink hat during a calibration procedure in the particular lighting environment of the conference center. If the number of pixels in the camera image that match this color model exceed a certain threshold, the hat is considered found; GRACE enters the `SUCCESS` state and a *Gotcha!* screen is displayed, asking to be touched to start the game again. If neither the pink hat nor a person is found after some time, the robot moves randomly and looks again.

There were many challenges involving sensing. Poor lighting, mirrors, and bright sunlight through windows made it difficult to obtain a pink-color model that was effective yet discriminating; pink or red shirts occasionally registered as the hat. Additionally, computational constraints limited the resolution of our camera images and therefore the distance at which frontal faces would be detected. However, these challenges were fully anticipated in our formulation of the task and in our development of the robot; in fact, these challenges were the very motivation for this endeavor.

Figure 3: The *wandering* and *directions* interfaces.

4.2.2 *Approach*

When a person is detected, GRACE first says, "I think I've found someone to ask," and begins to move toward the person to initiate an interaction. During this approach, the robot tries to observe societal norms such as speed, direction of approach, greeting (e.g. saying, "Excuse me!"), and personal space. These were designed from common sense understanding and fine-tuned through testing. When the approach is complete, if the person is still there (approximately four feet away), GRACE enters the `ASKING` state.

4.2.3 *Asking*

In the `ASKING` state, GRACE displays a "Can you help me?" screen, with buttons "Yes" and "No." Meanwhile, she says "I am looking for the person in the pink hat. Can you help me?" If the person presses "No," GRACE thanks the person, displays a *Thank you* screen, and returns to the `IDENTIFICATION` state. If the person presses "Yes," GRACE asks the person to point her in the direction of the pink hat. To facilitate this, the *directions* image is displayed on the touchscreen (Fig. 3). This image depicts GRACE in the center and eight arrows pointing outwards from the robot along with the line "Which way is the person in the pink hat?" The arrows are foreshortened to provide directional perspective from the participant's point of view.

Participants then touch the arrow that most accurately points in the direction of the person in the pink hat. For example, pressing on the top-most arrow causes GRACE to turn 180° and travel away from the participant. Once an arrow is pressed, GRACE briefly turns her animated face in the indicated direction (to help convey that the information was properly received), displays the *Thank you* screen, verbally thanks them, and enters the `FOLLOWING` state.

There were some instances of participants being confused about the orientation of the arrows in the interface relative to the robot in space; perhaps the interface was not completely self-evident. In particular, some participants provided the opposite direction, believing they had to flip their frame of reference.

4.2.4 Following

GRACE turns and follows the suggested direction, avoiding obstacles found by the laser scanner. At the same time, GRACE looks for the pink hat. If it is not found after traversing a minimum distance of 3m, GRACE returns to the `IDENTIFICATION` state and looks for another person to approach. During this phase of the task, the *wandering* image is displayed on the touchscreen. If the image is pressed at any point in this phase, GRACE stops and enters the `ASKING` state. Whenever the robot needs to wait for human input, there is a timeout in case no input is given.

4.2.5 Success

When GRACE finally locates the person in the pink hat, she displays a *Gotcha!* screen. This may happen at any point during the task, except when she is interacting with someone in the `ASKING` state. GRACE performs a dance (spinning two times in a circle) and plays the Aerosmith song "Pink." Given the distributed social nature of the task, it is not enough for the robot simply to finish; since she has distributed her perceptual burdens among others, it is important to inform them when the person in the pink hat is found. The success is not just the robot's but is attributable to everyone who interacted with the robot. Thus the dance provides entertaining visual and auditory feedback to participants.

5. OBSERVATIONS AND RESULTS

GRACE's ability to find the person in the pink hat rested more on her capabilities as a socially interactive system (finding and communicating with people) than as a mobile perceptual robot (directly searching for the hat). Due to the socially situated and interactive nature of the robot's task and design, its success was not judged primarily by metrics such as time or turns to completion. Rather, we evaluated GRACE's performance by observing and analyzing three aspects of interaction: the influence of the spatial and social nature of the environment on the human-robot interaction; the nature of social interactions between GRACE and conference participants and participants with each other; and the robot's success in using social interaction to obtain perceptual assistance.

GRACE operated for approximately 15 hours over the course of three days at the conference. Video recordings and live observations of GRACE's performance were made during this time. On-site observation allowed us to form a general impression of GRACE's interactions with conference participants. For example, a challenge that was exacerbated by the high density of conference attendees was that participants would often stand in GRACE's way after giving her directions (despite GRACE's visible but unsuccessful attempts to find a path to her new goal). In retrospect, it would have been beneficial to implement a set of increasingly aggressive behaviors (e.g. facially, verbally, and in terms of motion). This would have been an appropriate context-dependent behavior: when a human helps the robot by giving a direction, the robot should resist immediate human interference until it is necessary once again. The museum tour-guide robot Minerva successfully employed such methods after her designers encountered similar difficulties [21]. On the other hand, some people ignored the robot entirely, while others became so invested in the task that they followed the robot around and helped it repeatedly.

Video recordings were subjected to a more fine-grained analysis and provided more detail. We coded and analyzed 3.6 hours of video using Noldus Observer software [14]. Manual codes were created and temporally applied to videotaped behaviors such as speech, spatial movement, gesture, and gaze as performed by GRACE, conference participants that interacted with her, and those that were in close proximity to GRACE (< 1m) but did not interact with her. Statistics gathered on the data in different spaces (described below) were weighted to compensate for differences in the length of video available for each space.

5.1 Social spaces

For analytical purposes, and in accordance with our emphasis on socially and physically situated interaction, we categorized the data according to the social spaces in which the interactions between GRACE and participants occurred. There are three categories corresponding to areas in the conference venue that varied in their spatial configuration and social use: the *reception*, a social event held in a large hall in which people were contained and crowded (31.25 minutes recorded); the *hallway*, a place through which people walked on their way to the various conference presentations and in which they examined displayed posters (104 min.); and the *banquet*, a social event during which the hallway was furnished with food tables for the occasion (82.4 min.). Sequentially, the reception occurred the first evening of the conference, the hallway activity occurred throughout the conference, and the banquet occurred last. A comparison of interactions within the three categories shows the effect of both social and spatial factors on human-robot, as well as interpersonal, interaction. Observation and analysis of the video showed that there were salient differences in the way people were affected by and interacted with GRACE in these three spatial and social environments.

5.1.1 Interaction group size

As explained in section 4, GRACE was designed to wander around the conference venue until a person was identified by the robot or initiated an interaction by touching the screen. GRACE's design encouraged dyadic (one-on-one) face-to-face interaction, as the robot could identify, query, and receive directions from one person at a time. A total of 171 touchscreen interaction sequences were observed in the video. Overall, 47% of them involved GRACE interacting with a single individual. Contrary to expectation, 53% of interactions involved more than one person gathering around GRACE and participating in the task by either taking turns giving her directions, helping each other understand the task, or locating the person in the pink hat before pointing GRACE in the right direction. The incidences of dyadic and multiple-person interactions in the various socio-spatial locations (hallway, reception, and banquet) were quantitatively different (Fig. 4). While the banquet and hallway both had very similar distributions of interactions, in the reception hall there were actually more interactions with GRACE involving two people (37%) than with one person (20%). This may be partially accounted for by the reception being spatially contained and crowded, with people coming into close contact with each other and with GRACE quite frequently, as well as by the social atmosphere of the event, which encouraged meeting and chatting with new people and robots.

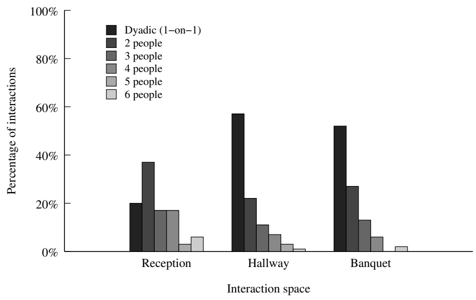

Figure 4: The size of groups with which GRACE interacted.

5.1.2 *Types of group interaction*

Interactions between conference participants differed qualitatively depending on the social and spatial location where they occurred. We recorded 331 interactions between conference participants while they were interacting with GRACE. Conversation between participants was the most common form (44%), followed by interaction through gaze (33%), spatial movement (walking, standing, turning) (17%), and gesture (touch, waving, pointing) (6%). The distribution over these different types of interaction is similar across the three spaces, but the reception saw a higher frequency of interpersonal interaction (3.3 per minute) than the hallway (1.7 per minute) and the banquet (0.6 per minute).

5.1.3 *Behavior of conference participants toward GRACE*

2000 separate instances of GRACE-oriented behavior were coded (Fig. 5). The most common behavior was gaze, which does not necessarily involve further physical interaction or aiding GRACE in her quest. This was closely followed by gesture, (i.e. using the touch-screen and interacting closely with GRACE). It is interesting to note that the number of "engaging" movements (such as moving, turning, and standing toward, in front of, or next to GRACE) was almost equal to the number of "disengaging" movements (walking, turning, or standing away from GRACE) in the hallway. Disengaging movements were significantly lower than engaging movements in both the banquet and reception situations. There were even a few cases in which participants talked to GRACE, although she did not have the capability to understand natural speech.

5.1.4 *Participants' responses to GRACE's movements*

Next, we looked at the effect of GRACE's actions on how conference participants interacted with the robot. A lag sequential analysis (a technique which measures the number

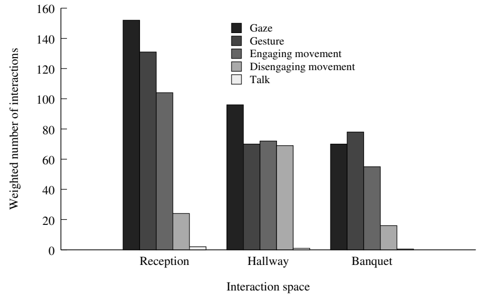

Figure 5: The types of GRACE-oriented behaviors exhibited by conference participants.

of times certain events are preceded or followed by other events) was performed to see how participants reacted to GRACE's engaging, disengaging, and "random" wandering movements within 5 seconds of their occurence.

There were a number of interesting differences between the interaction spaces with respect to participants' responses to GRACE's actions (Fig 6). For example, the reception had the highest overall rate of participants' engagement with GRACE. Furthermore, when GRACE made an engaging movement (turning toward a participant), people were more likely to make an engaging movement than a disengaging movement in the reception and banquet. On the other hand, in the hallway, people were equally likely to engage or disengage in response to an engaging movement by GRACE. This is possibly due to the transitionary nature of the hallway; the other two situations, although they were in different spaces, had a similar social purpose. Another notable phenomenon was that in the reception, after GRACE made a *dis*engaging action, people were more likely to re-engage the robot, either through movement, gaze, or using the touchscreen, than in the other two spaces. It is possible that a time effect (familiarity) may be partly responsible for the differences between the interaction spaces, but the similarities between the reception and the banquet suggest that the social activity and the physical space are both important, in different ways, to the quality of the interactions that occur.

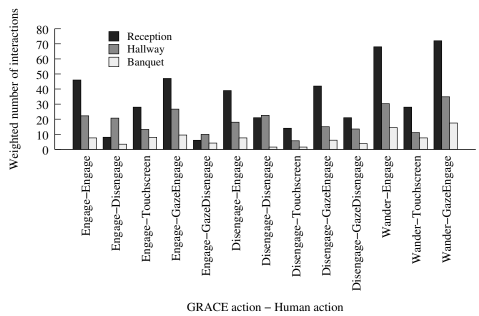

Figure 6: A lag sequential analysis of GRACE's actions and human actions after 5 sec.

It is possible that a time effect (familiarity) may be partly responsible for the gross differences between the interaction spaces, but the results of the video analysis (such as similarities between the reception and banquet) suggest that the physical and social environment in which the robot is situated have a significant effect on the resulting human-robot interaction. While GRACE was reasonably successful at getting help from conference participants, it is apparent that the way she was designed made her more effective in instigating interaction in certain social and physical environments than others. This suggests that the social and spatial situation in which the robot will be placed during the interaction should be seriously considered when designing social robots.

5.2 Success in obtaining help

Finally, we looked at whether the robot was able to use interaction effectively in order to seek and obtain human help in achieving its goals. We logged the robot's internal state for 7.2 hours of her operation. As we were interested in promoting and observing human-robot interaction, we considered a "good" game to be one in which GRACE interacted with at least four people before locating the hat, which was typically located far (e.g. 10m) from the robot's starting location. 53% of GRACE's 75 logged completed games met this criterion (Fig. 7). 86% of the 391 logged touchscreen interactions were "helpful;" that is, a human interactor said that they could help and provided a direction. We did not record whether these directions were accurate, although it would have been interesting to compare people's honesty and usefulness when the robot asked for help versus when a human offered to help.

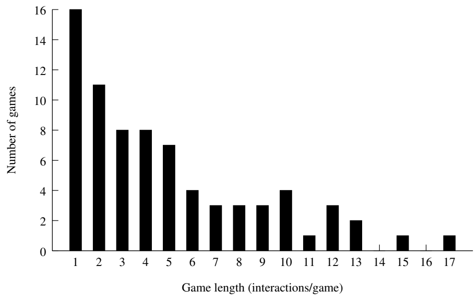

Figure 7: The length of games in terms of number of interactions.

The robot initiated interactions 57% of the time, so humans initiated 43% of interactions (Fig. 8). When the robot initiated a logged interaction, humans provided help 83% of the time. When humans initiated the interaction, they provided help 91% of the time. We consider it a success that the robot initiated requests for help about as frequently as humans initiated offers of help, since this demonstrates that a mixed-initiative approach was appropriate. It is also a significant result that humans were more willing to help the robot if they initiated the interaction than if the robot initiated it; this result should be investigated further.

6. DISCUSSION

Our design, development, and observation of this system has raised a number of interesting questions that should be explored further. GRACE's reliance on human perception to supplement her own introduces a dependency that has a number of positive and negative implications. In this section, we discuss the possible use of mixed-initiative interaction to transform this sort of dependent relationship into the type of interdependent relationship that is necessary for robots to help humans perform truly useful and important tasks.

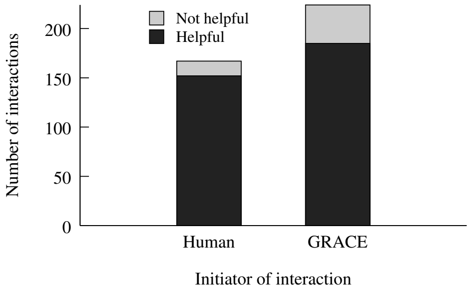

Figure 8: The proportion of helpful interactions to unhelpful interactions.

6.1 Dependency

In allowing a robot to transfer a part of its perceptual burden to humans in the environment, we create a dependent relationship: the robot depends on people to help it. This is interesting because we typically expect the opposite – that most robots will be built to help and serve people. In general, for both practical and ethical reasons, people tend to want to limit their dependency on our machines. This raises the question of whether machines that are dependent on humans may be more appropriate and/or engaging in certain situations if they instill a sense of comfort that humans are still a necessary part of the loop.

However, it should be noted that dependency is not necessarily a unidirectional relationship. Humans might be dependent on a robot to execute a task or some parts of a task, and the robot might be dependent on humans for assistance at certain points. Indeed, it may be more useful to think of interdependency: doctors and nurses in a hospital have different skill sets and therefore different job descriptions, and each group is dependent on the other to ensure successful execution of their common mission.

It is important in well-designed systems that built-in dependencies are sustainable, i.e. that a robot's requests for assistance do not become bothersome. What qualifies as bothersome behavior is of course task-specific. However, it is reasonable to say that a socially situated robot should attempt to determine when it can ask for assistance and when it should wait (e.g. when a human is busy). For GRACE, it was difficult to determine when a human was actively engaged in a conversation with others, so she met with annoyance on the part of some conference attendees. Barring this, we believe that for a built-in dependency to be sustainable it should be fault-tolerant [15] and compelling.

The idea of dependency presents a new approach for designing interactive systems. Dependency can be, at the same time, a catalyst for interaction, a quality of the interaction, and a means for achieving a goal. Its limits, however, are unclear. When is dependency counterproductive or inappropriate? How much autonomy to push for: as much as possible, with interaction occurring only when necessary, or is a minimum degree of interaction necessary to serve other purposes? For dependency to become interdependency, mixed-initiative interaction seems to be a critical component.

6.2 Social Mixed Initiative

Social interaction, by definition, uses mixed initiative. A great deal of psychology research has been concerned with turn-taking in human beings. However, mixed-initiative interaction has been somewhat narrowly defined in the human-computer interaction literature: it often assumes a rather well-defined task that is mutually understood by the human and the system and it does not necessarily take into account the context of the social situation. We belive that mixed-initiative interaction should be considered more broadly; that is, it should account for the possibility of the natural social initiative (or interruptions) that may occur between physical, mobile agents. In our work, the robot could ask for help or a human could offer it; this paradigm is not commonly employed and should be explored further.

Moreover, as we have found, it is necessary to anticipate the qualities of the specific social situation in order to use mixed initiative and turn-taking most effectively. For humans and robots to benefit from each others' asymmetrical capabilities, they must have a reciprocal understanding of each others' (possibly asymmetrical) social affordances. In this way a positive feedback loop can be established: "behavior affords behavior" [5].

6.3 Developing socially distributed cognition

The socially distributed perception that GRACE exhibited can be seen as a preliminary step in designing robots for participation in socially distributed cognition. Distributed cognition is centered around a task, such as GRACE's search for the pink hat, but it is also informed by a cultural and historical context that specifies the interaction and makes it possible. Distributed cognition also requires a high degree of coordination between the robot and the people who interact with it, similar to the mixed initiative in GRACE's performance. However, unlike the simple dependency that GRACE exhibited, a socially distributed cognitive system would necessitate more interdependence between the robot(s) and the humans involved and would be an instructive framework for human-robot collaboration.

In our observations, we saw this kind of collaboration among the participants who explained to each other what the robot was doing and then helped each other through the process of giving it assistance. The robot, by asking people for directions, not only managed to tap their resources to accomplish its goal, but also instigated a distributed chain of searching. For example, some people did not see the person with the pink hat, but others would point the person out to them, or indicate where the hat was last seen, or demonstrate how to work the robot controls – and only then would the person transmit this knowledge to the robot.

An important aspect of both the socially situated and the socially distributed cognition frameworks is that it places intelligence within the whole system, rather than at just the individual level. Accordingly, even though GRACE was neither very smart nor very perceptive, she was able to find the pink hat by relying on the intelligence of the system in which she participated through her social abilities. Likewise, when designing a collaborative human-robot system, we can rely on the complexity of the social and physical environment to imbue even a simple design with complex and useful resulting actions.

7. CONCLUSION

We have introduced *social tag* as a way of developing and demonstrating *socially distributed perception*, or a robot's ability to augment its sensory capabilities through social interaction. We have discussed the structure of the task, and we have described the design of GRACE's control system that allowed her to perform this task at AAAI 2005. Our observations and results suggest that the robot was able to successfully use social interaction to request and accept assistance from conference participants with the aim of finding the person in the pink hat, and that the specific social situation had a significant effect on the nature of this interaction.

The game of social tag is a useful construct for testing a robot's ability to use people as a resource for enhancing its perceptual understanding of the environment. Furthermore, this ability is a potentially important design principle for mobile robots operating in human environments. Much of the work in human-robot social interaction is targeted toward a particular application, e.g. in the areas of entertainment, service robotics, or psychological research. We propose that *any* robot intended to perform tasks in human environments should be able not only to safely negotiate such environments but also to actively engage the environment and its inhabitants in order to improve performance.

8. ACKNOWLEDGMENTS

This project was partially supported by National Science Foundation ITR projects #IIS-032914, #IIS-0121426, and #SES-0522630.

Thanks to Greg Armstrong and Brennan Sellner for technical contributions and to Sara Kiesler, Rachel Gockley, and Kristen Stubbs for helpful comments.

9. REFERENCES

[1] C. Breazeal (Ferrell) and B. Scassellati. Infant-like social interactions between a robot and a human caretaker. *Adaptive Behavior*, 8, 2000.

[2] R. Brooks, C. Breazeal, M. Marjanovic, B. Scassellati, and M. Williamson. The Cog Project: Building a humanoid robot. In C. L. Nehaniv, editor, *Computation for Metaphors, Analogy and Agents*, volume 1562 of *Springer Lecture Notes in Artificial Intelligence*. Springer-Verlag, 1999.

[3] A. Bruce, I. Nourbakhsh, and R. Simmons. The role of expressiveness and attention in human-robot interaction. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA '02)*, May 2002.

[4] K. Dautenhahn, B. Ogden, and T. Quick. From embodied to socially embedded agents - implications for interaction-aware robots. *Cognitive Systems Research*, 3:397–428, 2002.

[5] J. J. Gibson. *The Ecological Approach to Visual Perception*. Lawrence Erlbaum Associates, NJ, 1979.

[6] M. A. Hearst. Trends & controversies: Mixed-initiative interaction. *IEEE Intelligent Systems*, 14(5):14–23, 1999.

[7] F. W. Heger, L. M. Hiatt, B. Sellner, R. Simmons, and S. Singh. Results in sliding autonomy for multi-robot spatial assembly. In *Proceedings of the 8th International Symposium on Artificial Intelligence,*

Robotics and Automation in Space, Munich, September 2005.

[8] E. Hutchins. *Cognition in the Wild*. MIT Press, 1995.

[9] K. Kawamura, R. T. Pack, and M. Iskarous. Design philosophy for service robots. In *IEEE International Conference on Systems, Man and Cybernetics*, volume 4, pages 3736–3741, Vancouver, BC, 1995.

[10] H. Kozima. Infanoid: A babybot that explores the social environment. In K. Dautenhahn, A. H. Bond, L. Canamero, and B. Edmonds, editors, *Socially Intelligent Agents: Creating Relationships with Computers and Robots*, pages 157–164. Kluwer Academic Publishers, Amsterdam, 2002.

[11] M. J. Mataric. Situated robotics. In *Encyclopedia of Cognitive Science*. Nature Publishing Group, Macmillan Reference Ltd., November 2002.

[12] M. Michalowski, D. Busquets, C. DiSalvo, L. Hiatt, N. Melchior, R. Simmons, and S. Sabanovic. Social tag: Finding the person with the pink hat (technical report). In *Proceedings of the National Conference on Artificial Intelligence (AAAI '05) Mobile Robot Competition Workshop*, July 2005.

[13] F. Michaud et al. Modularity and integration in the design of a socially interactive robot. In *Proceedings of the IEEE International Workshop on Robot and Human Interactive Communication*, pages 172–177, Nashville TN, 2005.

[14] Noldus Information Technology: The Observer. http://www.noldus.com/products/observer.

[15] D. A. Norman. *The Design of Everyday Things*. MIT Press, 1998.

[16] I. Nourbakhsh, E. Hamner, E. Porter, B. Dunlavey, E. M. Ayoob, T. Hsiu, M. Lotter, and S. Shelly. The design of a highly reliable robot for unmediated museum interaction. In *Proceedings of ICRA*, Barcelona, Spain, 2005.

[17] J. Pineau, M. Montemerlo, M. Pollack, N. Roy, and S. Thrun. Towards robotic assistants in nursing homes: challenges and results. *Robotics and Autonomous Systems*, 42:271–281, 2003.

[18] C. L. Sidner, C. Lee, C. Kidd, N. Lesh, and C. Rich. Explorations in engagement for humans and robots. *Artificial Intelligence*, May 2005.

[19] R. Simmons et al. GRACE: An autonomous robot for the AAAI Robot Challenge. *AAAI Magazine*, 24(2):51–72, 2003.

[20] L. Suchman. *Plans and Situated Actions: The Problem of Human/Machine Communication*. Cambridge University Press, Cambridge, UK, 1988.

[21] S. Thrun, J. Schulte, and C. Rosenberg. Interaction with mobile robots in public places. *IEEE Intelligent Systems*, pages 7–11, July/August 2000.

[22] S. Turkle. *Life on the Screen: Identity in the Age of the Internet*. Touchstone, New York, 1997.