Auton Robot (2007) 22:385–397
DOI 10.1007/s10514-006-9015-6

Socially Distributed Perception: GRACE plays social tag at AAAI 2005

Marek P. Michalowski · Selma Šabanović · Carl DiSalvo · Didac Busquets · Laura M. Hiatt · Nik A. Melchior · Reid Simmons

Received: 9 March 2006 / Revised: 4 October 2006 / Accepted: 31 October 2006 / Published online: 30 December 2006

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. *Socially distributed perception* is defined as a robot's ability to augment its limited sensory capacities through social interaction. We describe the task of social tag and its implementation on the robot GRACE for the AAAI 2005 Mobile Robot Competition & Exhibition. We then discuss our observations and analyses of GRACE's performance as a situated interaction with conference participants. Our results suggest we were successful in promoting a form of social interaction that allowed people to help the robot achieve its goal. Furthermore, we found that different social uses of the physical space had an effect on the nature of the interaction. Finally, we discuss the implications of this design approach for effective and compelling human-robot interaction, considering its relationship to concepts such as dependency, mixed initiative, and socially distributed cognition.

Keywords Social robotics · Human-robot interaction · Social situatedness · Socially distributed cognition · Mixed initiative · Observational analysis · Multidisciplinary design

M. P. Michalowski (✉) · C. DiSalvo · D. Busquets · L. M. Hiatt · N. A. Melchior · R. Simmons
Carnegie Mellon University,
Pittsburgh, PA 15213, USA

S. Šabanović
Rensselaer Polytechnic Institute,
Troy, NY 12180, USA

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, visiting town to attend a conference, makes plans to meet a former colleague for lunch. She has written him an 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, so he asks the hotel staff if they have seen a woman wearing a blue coat. The staff point him in the direction they last saw her, but 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 finds her. In this situation, the researcher's sensory information about the world is insufficient to solve his problem, as he cannot visually locate his colleague. However, his ablility to interact with other people who might have seen her allows him 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" (Brooks et al., 1999).

At the AAAI 2005 Mobile Robot Competition & Exhibition, the robot GRACE (Graduate Robot Attending a ConferencE) (Fig. 1) faced a similar situation. 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 (Michalowski et al., 2005, 2006). Tasks involving searching and object localization are a common theme at the Exhibition, and machine perception is often the limiting factor in these robotic systems. As the conference is attended by many people and takes place in an environment well-suited to human perception, we wanted to create a search task in which a "socially situated" (Suchman,1988) robot could determine

Fig. 1 The robot GRACE

the whereabouts of the team member primarily though social interaction with humans in the environment. Direct sensing is used to navigate, to locate people, and to identify the hat when it is nearby. Rather than treating humans as obstacles, the robot is aware of the social environment and uses it as a resource in its search. 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. The robot's performance at the Exhibition was specifically tailored to the conference setting and the population of artificial intelligence researchers who may, or may not, regularly work with embodied robots.

We introduce the term *socially distributed perception* to describe a robot's ability to augment its own perception through social interactions with people. Such a robot is perceptually situated within a social, as well as physical, environment and can treat both the social and the physical characteristics of that environment as affordances[1] (Gibson, 1979). In order for the robot to take advantage of these social affordances, it must be able to engage in social interaction with people. In our task, designing for social interaction involves ensuring that the robot is capable of appropriately requesting help and accepting it when offered.

In this article, we review research from a variety of domains in designing robots that can be helped by people, as one of the assumptions in our work is that humans will be willing to help robots accomplish their tasks. We propose the new approach of socially distributed perception as the distribution of perceptual demands by a situated robot across a group of people. The concept of "social situatedness" motivates our design of a task framework that, despite being simple, can support a varied range of social interactions when placed in the context of a complex social environment. We describe GRACE's design for the conference and discuss social tag as a task framework that makes use of socially distributed perception.

Our work introduces social science and design theory and analyses as complementary to engineering practices. The interaction between roboticists, social scientists, and designers was integral to the development and evaluation of this project. As human-robot social interaction was our focus, we conducted an observational analysis of the robot in action. Results from this analysis suggest GRACE was successful in obtaining help with her task, and we describe how interactions between GRACE and conference participants were shaped by the social and physical environments. In particular, we found that the social use of a space often had as significant an impact on interaction as did the physical nature of that space. Finally, we discuss socially distributed perception as a novel principle for designing appropriate, compelling, and beneficial social interactions between humans and robots.

2 Background

2.1 Humans helping robots

Within the existing robotics literature, relevant work that relies on humans providing spatial directions is found in the area of deictic gestures (pointing) (Sidner et al., 2005) and tele-robotics (Nourbakhsh et al., 2005). 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 or to lower human expectations of a robot's capabilities to more realistic levels (Breazeal (Ferrell) and Scassellati, 2000; Kozima, 2002; Okada et al., 2000). Non-traditional domains are also relevant and insightful to shaping this form of human-robot interaction: artists have produced several examples of robots that rely upon human presence and interaction for their operation (White, 1987; Goldberg, 2000; Penny, 1997; Rinaldo, 2000).

Recently, research has been done in allowing robots to request human assistance in cases where the robot's capabilities are insufficient for task completion. This research can be

[1] We use the term "affordances" to refer to the relational and emergent properties of interaction between an actor (in this case, the robot) and its physical and social environment. The social environment of the conference presents certain latent "action possibilities" to the robot, which are dependent on its capabilities but may exist independently of its recognition of those possibilities.

organized into "sliding autonomy," "mixed initiative," and deictic input.

Sliding autonomy refers to a form of human-robot collaboration in which control can be seamlessly transferred between autonomous robot operation and human teleoperation in order to complete complex tasks such as remote construction (Heger et al., 2005). Sliding autonomy typically involves robots working remotely from their human operators, therefore precluding true social interactions. Our work additionally differs in that GRACE is always autonomous and is never controlled by a human being. When the transfer of control can be initiated by either party, this is known as mixed initiative.

Mixed-initiative interaction is alternatively used to describe the design of interfaces or software agents that are expected to aid in the execution of information-processing tasks (Hearst, 1999). Such 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 allow either party to initiate a helpful *social* interaction.

Another approach is to combine deictic human input with sensors embedded and spatially distributed throughout the environment (Kawamura et al., 1995). 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 or dynamic. While distributed sensors may be considered affordances in a robot's physical environment, GRACE considers conference participants to be affordances in her social environment. She is able to use people's sensory abilities as sources of information because she has been given the interactive "tools" necessary to participate in the social environment.

2.2 Social situatedness and distributed cognition

The concept of "social situatedness" that we use to frame GRACE's interactions with conference participants augments the concept of "situated robotics" (Mataric, 2002) by taking social interaction into account. "Situated robotics" is founded on the notion that the complexity of the physical environment in which an embodied robot operates influences the robot's emergent behavior. "Social situatedness" expands the sources of this influence by considering social actors to be integral parts of the robot's environment. The notion of a "socially situated agent" (Dautenhahn et al., 2002) implies both social and physical interaction with the environment in order to acquire information about the social and physical domains. In the case of socially situated robotics, the organization of situated action is emergent from the interaction both among actors and between actors and their social and physical environments. The more the robot is "embedded" (Dautenhahn et al., 2002) in the environment, the more it will be able to respond to the social influences in its environment and the more complex its interactions with others can be. As we discuss in our observations, a robot that is not appropriately embedded ("structurally coupled") in the environment will not be robust to changes.

In "social tag," GRACE is a socially situated actor in an environment that is populated by humans and is therefore social as well as physical (Suchman, 1988). Conference attendees participated in the search task along with GRACE, sometimes in groups, looking for the person in the pink hat. This resembles a system of "socially distributed cognition," which occurs when group activities are focused on a task and distributed over a range of media, artifacts, and people (Hutchins, 1995). Interaction between machines and people in socially situated and distributed cognition 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 meaningful. In a distributed cognitive system, people can be seen as *partners* in GRACE's search for the pink hat rather than mere *users*. 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; this makes it seem purposeful to interaction partners (Turkle,1997; Suchman, 1988).

We consider "perception" to be a component of "cognition" that is simple enough for a robot to distribute socially, yet compelling enough to enable basic information exchange and to provide interesting behaviors for observation and research. Drawing on, and extending, the concepts of social situatedness and distributed cognition, our work aims to create a socially situated robot that performs a task in the context of a group of people by making use of the perceptual abilities distributed across the members of this group. 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. Such a system can be seen as an important first step toward the goal of creating a robot that can participate, rather than just be used, in a socially distributed cognitive system, which would be useful for mixed-initiative human-robot collaboration.

3 GRACE

GRACE (Fig. 1) 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 (Simmons et al., 2003). Our work builds on GRACE's ability to interact socially and applies it to a new task.

GRACE is an RWI B21 mobile robot. For this task, it 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 (Simmons et al., 2003).

4 Social tag

The task of social tag—finding the person with the pink hat—has several important attributes. The task is relatively simple and can be executed reliably, which allows us to focus on designing and observing human-robot interaction. Formulating the task as a fun and commonly-understood game increases people's willingness to interact with the robot. 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 robotic uses.

Many tasks in robotics research use distinctively colored objects, such as pink hats, to simplify the machine perception problem. Despite this simplification, in large crowded rooms the task of finding a hat is still extremely difficult for a robot. While the state of computer vision continues to improve, 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 whose appearance would be distinctive and thus easily remembered and recognized by conference participants. Indeed, there are cases in which no amount of local sensory ability would enable successful task completion; if the hat were in a different room, even a human would need to enlist the help of other humans who are moving through the environment, communicating with each other, and collectively remembering the locations of prominent objects. Accordingly, while GRACE depends on her own sense of vision to achieve the goal (i.e. recognizing the pink hat), her primary mode of perception is asking people for help in a socially acceptable manner. Without any social interaction, it is highly unlikely that GRACE would have ever found the pink hat, except by accident.

4.1 Interaction and interface design

GRACE's interface design attempts to support the goal of allowing humans to help robots by making the interaction simple and compelling. 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 (Fig. 1). Touchscreens have been found to be a simple and intuitive medium for transferring information to socially interactive mobile robots (Pineau et al., 2003; Michaud et al., 2005). 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 several full-screen interfaces that both convey the state of the robot and prompt for 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, in the form of appropriate sound clips played during different phases of the task, is another important component of GRACE's repertoire. In addition to adding to the playful character of GRACE, it calls attention to the robot and functions as another mode of expressing its current state to participants.

4.2 Task phases

The task has five main phases: *identification* of approachable humans; *approach* toward a human; *asking* for directions; *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 create deviations from this script by volunteering assistance without being asked. Our form of mixed-initiative interaction allows GRACE to handle such deviations.

The robot operates autonomously under the control of a simple state machine (Fig. 2), where states correspond with the phases of the task. We now describe these phases and discuss significant implementation details.

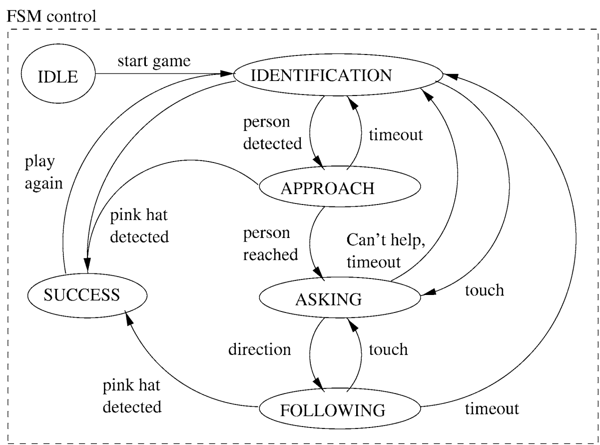

Fig. 2 The finite state machine that comprises GRACE's control task for Social Tag

4.2.1 Identification

When the game starts, 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, left). 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 (since people can engage the robot). Meanwhile, GRACE plays music and periodically says phrases such as "Where is the person in the pink hat?" Her animated face frequently changes direction, giving the appearance of actively looking around while wandering. If the screen is touched, GRACE transitions to the ASKING state.

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 color models. People can be located more reliably by correlating data between these two sensors (Michalowski and Simmons, 2006). This is done by registering the locations of faces and laser obstacles with each other in a robot-centric coordinate frame and labeling a laser obstacle as a person if there is a face located above it. When a person is detected, the robot enters the APPROACH state.

Fig. 3 The *wandering* and *directions* interfaces

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 and GRACE enters the SUCCESS state. 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.

4.2.2 Approach

When a person is detected and GRACE enters the APPROACH state, she 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 of the appropriate way to start an interaction and were fine-tuned through testing and modification prior to the conference. When the approach is complete, if the person is still there (approximately four feet away), GRACE enters the ASKING state. If the person is no longer there, GRACE transitions back to the IDENTIFICATION 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. The *directions* image depicts GRACE in the center and eight arrows pointing outwards from the robot (Fig. 3, right). The arrows are foreshortened to provide directional perspective from the participant's point of view.

Participants are then expected to 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.

4.2.4 Following

Once in the FOLLOWING state, 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 travelling about 3 m, 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.

4.2.5 Success

When GRACE finally locates the pink hat, she enters the SUCCESS state and 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, the task success is attributable to everyone and it is important to inform them when the person in the pink hat is found. The dance provides entertaining visual and auditory feedback to participants.

The person in the pink hat was typically stationary during a game, so success (or failure) in finding the hat was due to GRACE's social interactions rather than to movement of the hat.

5 Observations and results

Analyzing human-robot interaction as a situated activity performed in the context of particular concrete circumstances is an effective and appropriate approach to the evaluation of robots meant to function in non-laboratory settings (Sabanovic et al., 2006). Focusing our attention on human-robot interactions in dynamic, open environments enables us to evaluate and revise our designs so as to construct socially responsible and responsive robots. In the case of robots for which human-robot social interaction plays a focal role (Fong et al., 2002), observation and fine-grained behavioral analysis of the robot interacting in the environment can be used to analyze

- how humans react to, and interact with, the robot;
- how humans interact with each other while interacting with the robot;
- which aspects of the robot's and human's actions lead to breakdowns in the interaction; and
- how the robot succeeds or fails to engage humans in interaction.

These analyses provide detailed quantitative and qualitative data that can be used to improve socially situated robot interactions through iterative design processes. Although the attendees of the Exhibition might not be considered naive users, we believe that our findings are broad enough to have relevance to general human-robot interaction.

GRACE's ability to find the person in the pink hat rests 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). The robotics community conventionally applies quantitative measures such as speed or turns to task completion to judge a robot's performance and compare among robot designs and architectures. However, the contextual nature of social interaction precludes reliance on quantitative measures of human-robot interaction alone, as this would lead to the systematic exclusion of evaluations of phenomena that are not easily amenable to quantification. Due to the socially situated and interactive nature of the robot's task and design, we evaluated GRACE's performance by observing and analyzing three aspects of interaction:

- the robot's success in using social interaction to obtain perceptual assistance;
- the influence of the spatial and social nature of the environment on the human-robot interaction; and
- the environment's effects on the nature of social interactions between GRACE and conference participants and participants with each other.

GRACE operated for approximately 15 hours over the course of three days at the conference and played over 100 games. Video recordings and live observations of GRACE's performance were made during this time. Onsite observation allowed us to form a general impression of GRACE's interactions with conference participants that would not not have been available from the logs alone. For instance, some people became so invested in the task that they followed the robot around and helped it repeatedly. We also observed a challenge that was exacerbated by the high density of conference attendees—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, as in Thrun et al. (2000)). This would have been an appropriate context-dependent behavior: when a human helps the robot by giving a direction, the

Fig. 4 Sample frame from video of GRACE

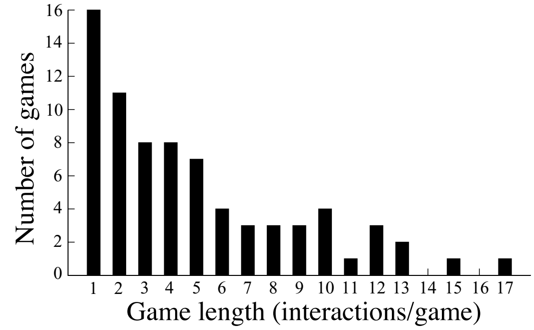

Fig. 5 The length of games in terms of number of interactions

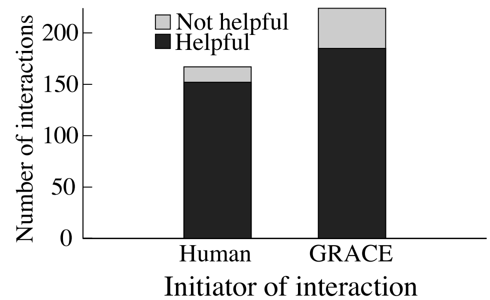

Fig. 6 The proportion of interactions in which help was provided or withheld

robot should resist immediate human interference until it is necessary once again.

The interactions between the robot and voluntary participants were recorded by a handheld camera. Using behavioral analysis software,[2] we coded and analyzed 3.6 hours of video (Fig. 4) taken over the three days of the conference. Manual codes were created and temporally applied to videotaped behaviors such as speech, spatial movement, gesture, and gaze as performed by GRACE, conference participants who interacted with her, and those who were in close proximity to GRACE (<1 m) but did not interact with her directly. The data was subjected to statistical analysis to describe the frequency of various types of events, while lag-sequential analyses[3] were used to determine the incidence of certain events being preceded or followed by others. Logs obtained from the robot were also used to measure the robot's success in getting help from conference participants. We focus our discussion on understanding situated interaction between people and robots; revealing factors that surpassed or challenged the initial design assumptions about social interaction; suggesting changes in the robot's design; and relating findings to more general applications in social robotics.

5.1 Success in obtaining help

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 had at least four interactions before locating the hat (which was typically located at least 10 m from the robot's starting location). 53% of GRACE's 75 logged games met this criterion (Fig. 5). 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 when the robot asked for help versus when a human offered to help.

The robot initiated 57% of the 420 logged interactions. When the robot initiated an interaction, humans provided help 83% of the time. When humans initiated the interaction, they provided help 91% of the time (i.e. 9% began an interaction but declined to help). 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. As expected, humans were more willing to participate in interactions they initiated; what is surprising is that they were almost as likely to help GRACE when she was the initiator (Fig. 6).

5.2 Social spaces

The original design assumptions did not take into account potential changes in the social and physical environments. Observational analysis focusing on the socially situated nature of the robot's task and design, however, revealed notable influences of the spatial and social characteristics of the setting on human-robot interaction.

We identified three spaces in the conference venue that varied in their spatial configuration and social use: the

---

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

[3] Lag-sequential analysis refers to the systematic observation of interaction and behavior sequences. It is used to detect commonly occurring sequences of events and to study the dynamics of interactive behavior from moment to moment. See Bakeman and Gottman (2006) for more information.

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 where they examined poster displays (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 activity in the hallway activity continued throughout the conference, and the banquet concluded the conference. 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 social spaces. This may be partially explained by the relaxed social atmosphere and the crowded conditions at the reception, where people frequently came into close contact with each other and with GRACE and were encouraged to meet and chat with people and robots alike.

5.2.1 *Interaction group size*

Dyadic (one-on-one) interaction is implicit in GRACE's design. As explained in Section 4, GRACE was designed to wander around the conference venue until she identified a person or a participant initiated a touch-screen interaction. The robot could identify, query, and receive directions from one person at a time. When the robot identifies one person as interacting, others are ignored until the ongoing interaction is terminated. Contrary to design assumptions, our observations show that it is equally likely for the robot to be approached by groups of people as by solitary individuals. A total of 171 interactions (dyadic or group) were recorded. Contrary to expectation, 53% of interactions involved more than one person gathering around GRACE and participating in the task either by taking turns giving her directions, by helping each other understand the task, or by 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 social spaces were quantitatively different (Fig. 7). 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%).

5.2.2 *Types of group interaction*

While interacting with GRACE, people often helped each other by discussing the robot's reactions, pointing out the salient aspects of the robot and its environment, and taking turns giving directions to, and getting responses from, the robot. Interactions between conference participants differed

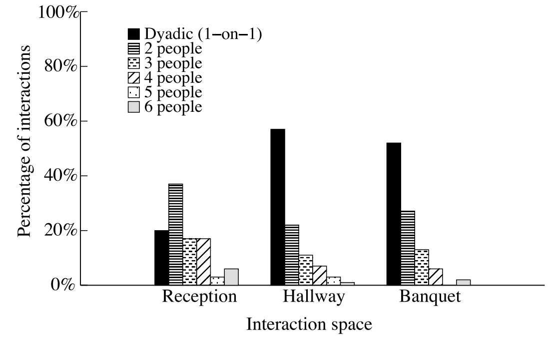

Fig. 7 The size of groups with which GRACE interacted

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

5.2.3 *Behavior of participants toward GRACE*

Two thousand separate instances of human behaviors that involved GRACE were coded (Fig. 8). 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 people who otherwise engaged with the robot (by moving, turning, or standing toward, in front of, or next to GRACE) was almost equal to the number of people

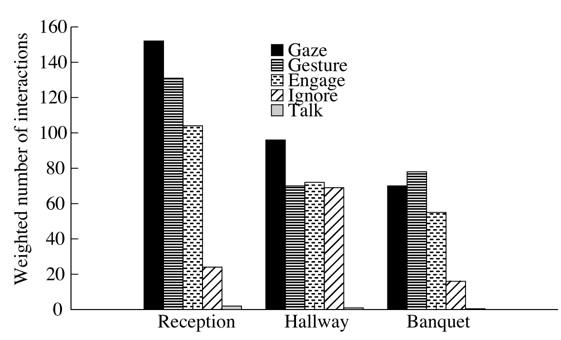

Fig. 8 The types of GRACE-oriented behaviors exhibited by conference participants

who passed her without interacting in the hallway. On the other hand, people passed GRACE without interacting much less 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.

5.2.4 *Participants' responses to GRACE's movements*

We also looked at the effect of GRACE's actions on how conference participants interacted with the robot. Lag sequential analyses (which are used to quantify the co-occurrence or different events within a small time window, e.g. 5 seconds) were used to study how participants reacted to GRACE after her attempts to engage (move toward or speak to a person) and disengage (move or turn away) from an interaction and when she was wandering through the conference space.

There were a number of interesting differences between the interaction spaces with respect to participants' responses to GRACE's actions (Fig. 9). The reception had the highest overall rate of participants' engagement with GRACE. Furthermore, when GRACE made an attempt to start an interaction, people were more likely to respond positively (by turning to the robot, looking at her, or touching her screen) in the reception and banquet. In the hallway, on the other hand, people were equally likely to engage as they were to disengage from the interaction in response to GRACE's interactive behavior. This can be attributed to the transitory nature of the hallway; the other two situations emphasized mostly stationary interaction. Moreover, people at the reception were more likely to attempt to continue interacting with the robot (through movement, gaze, or via touch-screen) even after GRACE made a disengaging action.

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

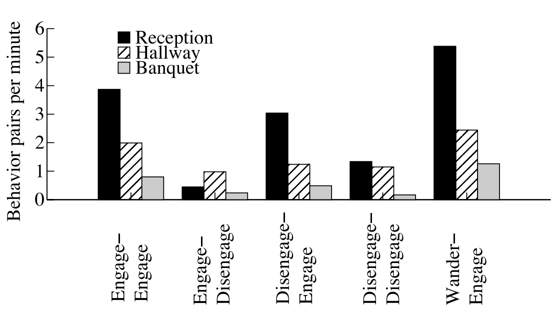

Fig. 9 A lag sequential analysis of GRACE's actions and human actions after 5 sec

6 Discussion

Our project was a multidisciplinary collaboration in which the design, development, and observation of the system 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. 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 useful and important tasks.

6.1 Multidisciplinary perspectives on social robots

Making robots socially interactive depends on designing devices that fit into an ongoing flow of human coordination, which raises novel technical, social, and design issues. Socially interactive robots are exemplary "boundary objects" (Star and Griesemer, 1989)—artifacts imagined, perceived, and interpreted differently by various disciplinary communities, yet still able to function within a shared intellectual space and provide a common focus for inquiry and action. They not only attract the interest of engineers, fascinated by the technical challenges they pose, but also intrigue social scientists who study the relationships and distinctions between humans and non-humans and the interplay between technological, scientific, and social factors. Robots furthermore epitomize the challenge of contemporary design: to appropriately and compellingly give form to heterogeneous and living networks of technologies, people, and artifacts. Socially interactive robots are meant to have a direct and personal impact on humans, and their construction calls for a systematic and comprehensive application of knowledge related to design, social, and technical factors.

The design of socially interactive robots such as GRACE is therefore a driving force for work at and across disciplinary boundaries—questioning and traversing the divisions between technical, design and social issues, as well as combining methods from different fields and encouraging interdisciplinary communication and collaboration. Our team, consisting of roboticists, a designer, and a social scientist, faced just these challenges. The unique multidisciplinary nature of this group shaped the iterative 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, modified, and evaluated. Differences in jargon and modes of knowledge-production made cross-disciplinary communication difficult at times, but our focus on a concrete object (i.e. the robot) and its behaviors helped to establish common ground.

6.2 Dependency as a design criterion

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, our dependence on machines is, at times, ethically or practically problematic.

However, it should be noted that dependency is not necessarily a unidirectional relationship. Since humans are better at different types of tasks than machines, humans might be dependent on a robot to execute a task or some parts of a task, while 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 each group is dependent on the other to ensure success in their common mission. It will be important that humans and robots alike have a mutual understanding of their skills, their ability to ask each other for help, and their interdependent relationship with each other.

It is also 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). Time and timing, then, are central qualities of interdependency that need to be taken into account in the design of human-robot interactions. For example, GRACE was unable to determine when a human was already actively engaged in conversation with others, so she met with annoyance on the part of some conference attendees. Beyond such annoyance, we believe that for a built-in dependency to be sustainable it should also be fault-tolerant and compelling. In our scenario we found that the use of humor (Goetz et al., 2003) and music were effective ways to facilitate and promote interactions with GRACE.

The idea of interdependency presents a new approach for designing interactive systems. Interdependency can be, at the same time, a catalyst for interaction, a quality of the interaction, and a means for achieving a goal. The limits of interdependency, however, are unclear. Identifying when dependency is counterproductive or inappropriate, as well as assessing the appropriate level of autonomy with respect to the social and technical conditions (e.g. when sensors fail), warrants future research. It is also necessary to identify what motivates people to interact with robots and to respond to their dependency. For example, music and humor seemed to promote interaction with GRACE, but it also appeared that people simply enjoyed participating in the robot's success and continued to monitor its progress after interacting. For dependency to become interdependency and collaboration, mixed-initiative interaction seems to be a critical component.

6.3 Situating mixed initiative in social interaction

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 believe 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. Social interaction, by definition, uses mixed initiative. 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 specific social situations in order to use mixed initiative and turn-taking most effectively. For example, the degree and nature of initiative exhibited by people was influenced by the social interaction space (Fig. 8). For humans and robots to benefit from each others' asymmetrical capabilities, they must have a reciprocal understanding of each others' (probably different) social affordances as well as of the affordances of the social situation in which they are involved.

6.4 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.

In addition to the simple dependency that GRACE exhibited, a socially distributed cognitive system might necessitate greater interdependence between the robot(s) and the humans involved. 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 use the robot's interface—and only then would the person transmit this knowledge to the robot.

An important aspect of both the socially situated and distributed cognition frameworks is that they place 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 social system in which she participated. 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 activities.

7 Conclusion

We have introduced *social tag* as a way of developing and demonstrating *socially distributed perception*, a robot's ability to augment its sensory capabilities through social interaction. We have discussed the structure of the task and described the design of GRACE's control system that allowed her to perform this task at AAAI 2005. Our observations 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 social and physical 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 off-load challenging sensory demands and thereby improve performance.

Acknowledgments This project was partially supported by National Science Foundation projects #IIS-0329014, #IIS-0121426, and #SES-0522630.

Thanks to Greg Armstrong and Brennan Sellner for technical contributions, to Sara Kiesler and Jodi Forlizzi for helpful comments, and to Linnda Caporael for analytical resources.

References

Bakeman, R. and Gottman, J.M. 2006. *Observing Interaction: An Introduction to Sequential Analysis*, 2nd ed. Cambridge University Press.

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

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

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

Fong, T., Nourbakhsh, I., and Dautenhahn, K. 2002. A survey of socially interactive robots: Concepts, design and applications: Technical Report CMU-RI-TR-02-29, Carnegie Mellon University Robotics Institute, Pittsburgh, PA.

Gibson, J.J. 1979. *The Ecological Approach to Visual Perception*. Lawrence Erlbaum Associates, New Jersey

Goetz, J., Kiesler, S., and Powers, A. 2003. Matching robot appearance and behavior to tasks to improve human-robot cooperation. In *Proceedings of the 12th IEEE Workshop on Robot and Human Interactive Communication (RO-MAN 2003)*, Millbrae, CA.

Goldberg, K. (ed.) 2000. *The Robot in the Garden: Telerobotics and Telepistemology in the Age of the Internet*. MIT Press.

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

Heger, F.W., Hiatt, L.M., Sellner, B., Simmons, R., and Singh, S. 2005. 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.

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

Kawamura, K., Pack, R.T., and Iskarous, M. 1995. Design philosophy for service robots. In *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, Vancouver, BC, vol. 4, pp. 3736–3741.

Kozima, H. 2002. Infanoid: A babybot that explores the social environment. In K. Dautenhahn, A.H. Bond, L. Canamero, and B. Edmonds (eds.), *Socially Intelligent Agents: Creating Relationships with Computers and Robots*, Kluwer Academic Publishers, Amsterdam, Chap. 19, pp. 157–164.

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

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

Michalowski, M.P., DiSalvo, C., Sabanovic, S., Busquets, D., Hiatt, L. M., Melchior, N. A., and Simmons, R. 2006. Socially distributed perception. In *Proceedings of the 1st Annual Conference on Human-Robot Interaction (HRI 2006)*. Salt Lake City, Utah.

Michalowski, M.P. and Simmons, R. 2006. Multimodal person tracking and attention classification. In *Proceedings of the 1st Annual Conference on Human-Robot Interaction (HRI 2006)*. Salt Lake City, Utah.

Michaud, F., et al. 2005. Modularity and integration in the design of a socially interactive robot. In *Proceedings of the IEEE International Workshop on Robot and Human Interactive Communication*, Nashville, TN, pp. 172–177.
Noldus Information Technology: The Observer. URL: http://www.noldus.com/products/observer.
Nourbakhsh, I., Hamner, E., Porter, E., Dunlavey, B., Ayoob, E. M., Hsiu, T., Lotter, M., and Shelly, S. 2005. The design of a highly reliable robot for unmediated museum interaction. In *Proceedings of ICRA*, Barcelona, Spain.
Okada, M., Sakamoto, S., and Suzuki, N. 2000. Muu: Artificial creatures as an embodied interface. In *ACM SIGGRAPH Conference Abstracts and Applications*, New Orleans, p. 91.
Penny, S. 1997. Embodied cultural agents: at the intersection of robotics, cognitive science, and interactive art. In *AAAI Socially Intelligent Agents Symposium*.
Pineau, J., Montemerlo, M., Pollack, M., Roy, N., and Thrun, S. 2003. Towards robotic assistants in nursing homes: challenges and results. *Robotics and Autonomous Systems*, 42:271–281.
Rinaldo, K. 2000. Emergent Systems: Exploring the confluence and coevolution of organic and technological cultures, http://accad.osu.edu/ ∼ rinaldo.
Sabanovic, S., Michalowski, M.P., and Simmons, R. 2006. Robots in the wild: Observing human-robot social interaction outside the lab. In *Proceedings of the 9th International Workshop on Advanced Motion Control (AMC 2006)*. Istanbul.
Sidner, C.L., Lee, C., Kidd, C., Lesh, N., and Rich, C. 2005. Explorations in engagement for humans and robots. *Artificial Intelligence*.
Simmons, R., et al. 2003. GRACE: An autonomous robot for the AAAI robot challenge. *AAAI Magazine*, 24(2):51–72.
Star, S.L. and Griesemer, J.R. 1989. Institutional ecology, 'translations' and boundary objects: Amateurs and professionals in berkeley's museum of vertebrate zoology, 1907-39. *Social Studies of Science*, 19(3):387–420.
Suchman, L. 1988. *Plans and Situated Actions: The Problem of Human/Machine Communication*. Cambridge University Press, Cambridge.
Thrun, S., Schulte, J., and Rosenberg, C. 2000. Interaction with mobile robots in public places. *IEEE Intelligent Systems*, pp. 7–11.
Turkle, S. 1997. *Life on the Screen: Identity in the Age of the Internet*. Touchstone, New York.
White, N. 1987. The Helpless Robot, http://www.year01.com/helpless.

Marek P. Michalowski is a Ph.D. student at Carnegie Mellon University's Robotics Institute. He received a B.A. in Computer Science & Psychology (2002) and a M.S. in Computer Science (2003) at Yale University, where his work focused on social attention in autism and perception in human-robot interaction. At Carnegie Mellon, he has worked on the perception and generation of attentive behaviors by a receptionist robot, and his current research focuses on interaction synchrony in robots designed for playrooms.

Selma Šabanović is a doctoral candidate at the Department of Science and Technology Studies at Rensselaer Polytechnic Institute. During her doctoral research, she was a visiting researcher at the Intelligent Systems Institute at the National Institute of Advanced Industrial Science and Technology (AIST) in Tsukuba, Japan and at the Robotics Institute at Carnegie Mellon University (CMU) in Pittsburgh, USA. Her current research lies at the intersection of social robotics, human-robot interaction, and science studies. She studies the impact of cultural factors on robot design and applies theories and methods from the social sciences to the design and evaluation of social robots.

Carl DiSalvo received a Ph.D. in Design from Carnegie Mellon University in 2006. As graduate student he worked on The Project on People and Robots in the Human-Computer Interaction Institute. He is currently a post-doctoral fellow in the Center for the Arts in Society and Studio for Creative Inquiry at Carnegie Mellon University.

D'dac Busquets is a Research Scientist at the University of Girona (Spain). He received the B.Sc. in Computer Science (1999) and the Ph.D. in Computer Science (2003) from the Technical University of Catalonia (UPC). His Ph.D. thesis was carried out at the Artificial Intelligence Research Institute (IIIA-CSIC, Barcelona). It was focused on qualitative navigation for robots and multiagent control architectures. He was a visiting postdoc at Carnegie Mellon University (2004–2006) as a Fulbright scholar, where he worked on auction mechanisms for coordinating teams of robots. He is now also working on the application of market mechanisms to task allocation and multiagent coordination. He has several papers in international scientific journals, conferences and workshops.

Laura Hiatt is a third year Ph.D. student in the School of Computer Science at Carnegie Mellon University. She came to Carnegie Mellon after graduation from Stanford University with a B.S. in Symbolic Systems in 2004. Her current research focuses on execution-time evaluation and modification of multi-agent plans.

Nik Melchior is a third-year Ph.D. student in the Robotics Institute at Carnegie Mellon University. He earned his Master's degree in computer science at Washington University in St. Louis. His present research interests include programming by demonstration and rough terrain path planning.

Reid Gordon Simmons is a Research Professor in Robotics and Computer Science at Carnegie Mellon University. He earned his B.A. degree in 1979 in Computer Science from SUNY at Buffalo, and his M.S. and Ph.D. degrees from MIT in 1983 and 1988, respectively, in the field of Artificial Intelligence. His thesis work focused on the combination of associational and causal reasoning for planning and interpretation tasks. The research analyzed the relationships between different aspects of expertise and developed a domain-independent theory of debugging faulty plans. Since coming to Carnegie Mellon in 1988, Dr. Simmons' research has focused on developing self-reliant robots that can autonomously operate over extended periods of time in unknown, unstructured environments. This work involves issues of robot control architectures that combine deliberative and reactive control, probabilistic planning and reasoning, monitoring and fault detection, and robust indoor and outdoor navigation. More recently, Dr. Simmons has focused on the areas of human-robot social interaction (including Grace and the roboceptionist project), coordination of multiple heterogeneous robots, and formal verification of autonomous systems. Over the years, he has been involved in the development of over a dozen autonomous robots.