CHI Workshop

Shaping Human-Robot Interaction

Understanding the Social Aspects of Intelligent Robotic Products

April 25th, 2004, Vienna

http://www.bartneck.de/workshop/chi2004

Imprint

Abstract

Intelligent robotic products have begun to integrate into our daily lives. While robotic technology has achieved a certain amount of maturity, the social impact of these products is largely unknown. Little information is available on how intelligent robotic products need to be designed to fulfill their social role in our society. This workshop focuses on three aspects of human-robot interaction: (1) technical implementation of intelligent robotic products, (2) form, function and behavior, and (3) human behavior and expectations as a means to understanding the social aspects of interacting with these products.

Organizers

Jodi Forlizzi
HCII and School of Design
Carnegie Mellon University
Pittsburgh, PA 15221 USA
forlizzi@cs.cmu.edu

Christoph Bartneck
Department of Industrial Design
Eindhoven University of Technology
Den Dolech 2, 5600 MB Eindhoven
The Netherlands
christoph@bartneck.de

Table of contents

Bringing Robots To Life:

Applying Principles Of Animation To Robots

A.J.N. van Breemen
Philips Research
Prof. Holstlaan 4 (WDC 1.34)
5656 AA Eindhoven, The Netherlands
albert.van.breemen@philips.com
Tel: +31 40 27 43440

ABSTRACT
The acceptance of user-interface robots as "social friend" depends among others on the ability of the user to understand the robot's behavior − to understand what it is doing and thinking. Body gestures are a natural channel to communicate this. Traditionally, the control of robotic body parts is carried out by feedback control loops. This results, however, in rather machine-like behavior that does not reveal much about what the robot is doing or thinking. We argue that in order to bring robots to life − such that they show behavior that can be naturally understood and anticipated − principles known from the field of character animation should be applied. In this article we will discuss this idea and present results obtained from experiments with our user-interface robot "iCat".

Author Keywords
Robot, user-interface robot, character animation, Ambient Intelligence.

ACM Classification Keywords
H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

INTRODUCTION

In the consumer electronics market a new genre of robots is introduced, known by names such as: **Personal Robots**, **Service Robots** or **User-Interface Robots** [7]. The application domain of these devices varies between messaging, multimedia applications (e.g. taking photos, playing melodies, telling stories), web-services and gaming. Some examples of these robots are AIBO ERS-7 [10], PaPeRo [8], ER1 Robot [6], and RoboScout [11].

Users create a special relationship with their user-interface robot and consider them as being a "social friend". Toys such as the "Tamagotchi" and "Furby" already introduced this aspect a few years ago. Owners of these toys are responsible for taking care of their toys − feeding them, playing with them and talking to them. From the robot's perspective, a good social relationship opens possibilities to build rich user profiles that can be applied to personalize the robot's tasks. From the user's perspective, personalization adds value to the robot application. However, in order to preserve the social relation it is important that the user understands the robot's behavior and is able to anticipate it [2].

Animated User Interface Characters have been applied for similar purposes as user interface robots and also here the principles of animation have proven their value [5].

Robot behavior manifests itself, besides sound and lights, through moving body parts. The movements of these parts reveal a lot about what the robot is doing and thinking. Traditionally, feedback loops control the movement of the body parts. For instance, an object tracking behavior is created by a feedback loop between the estimated object position from a camera and the servos of the robot's head [4]. This results in *machine*-like behavior that − in contrast to *life*-like behavior − can not be naturally interpreted.

Early day's animators dealt with a similar problem: creating the illusion of life of characters on paper, such that the audience understands what the character is doing and thinking [12]. Principles of animation were discovered that are the basis for creating the illusion of life. We consider user-interface robots to have the same problem of early day's animations: they miss the illusion of life, which is so important to be understandable for users. In this article we propose to apply the principles of animation to user-interface robots and show several results of this idea.

The remainder of this article is organized as follows. In the next section we present user-interface robots and discuss the problem of understanding and anticipating their behavior. Section 3 presents principles of animation and summarizes the twelve main principles. Section 4 explains how these animation principles can be applied to robots and present results obtained during experimentation with our user-interface robot "iCat". Section 5 summarizes our work.

USER-INTERFACE ROBOTS

Instead of being a universal house-hold robot like "Rosie" from "The Jetsons", user-interface robots currently entering the consumer market have specialized application domains. At Philips Research we are currently researching the usage of user-interface robots for home automation. The user-interface robot's main purpose is to provide an interface to devices in an Ambient Intelligence Home environment such as the "HomeLab" [1].

During previous research we build Lino (see figure 1), a mobile user-interface robot with a special head to create facial expression [3]. Lino is able to recognize spoken commands, to autonomously navigate in the home and to recognize objects by using vision. Our latest user-interface robot is the "interactive Cat", or just "iCat" (see figure 1). In contrast to Lino, iCat is smaller and lacks mobility, so that we can solely focus on the robot-human interaction. The iCat has 13 servos to move different parts of the head in order to create facial expressions, a stereo microphone to determine the direction of sound and a build-in webcam. Furthermore, iCat can be connected to a home network to control devices (e.g. VCR, TV) and to use the Internet.

Figure 1. Our user-interface robots Lino and iCat.

It's iCat's task to recognize users, build profiles of them and handle user requests. The profiles are used to personalize different kind of ambient functions performed by the robot. For instance, different light and sound conditions are used for one user asking iCat to create a 'relaxing atmosphere' and another user requesting the same. In order to learn rich user-profiles, a good social relationship between the iCat and the user is required, because both should understand each other and be willing to spend time in teaching each other things about themselves. As argued in the introduction, this relationship only lasts when the user is able to understand what iCat is doing and thinking.

From preliminary experimentation we observed that facial expressions can reveal what iCat is thinking. In one particular experiment a user had to tell iCat that "Madonna has released a new album". iCat reacted by showing a surprised face and said "I didn't know that". From this reaction, the user anticipated that iCat did like Madonna very much − something that is hard to infer when the sentence "I didn't know that" was spoken without any emotion. Experiments like this one let us hypothesize that the complete robot's reaction, including the movements of all its body parts, need to be properly choreographed in order to communicate the right message to the user.

PRINCIPLES OF ANIMATION

More is needed to create naturally understandable behavior then only having a robot react to stimuli from the environment. People need to be able to explain the behavior of the robot in terms of intentional, mental and emotional states [2]. Thomas and Johnson [12] mention in *Illusion of Life* that change of expression can reveal the thought process of a character: "It is the change of shape that shows the character is thinking. It is the thinking that gives the illusion of life. It is the life that gives the meaning to the expression." (p471).

Over the years animators have discovered several fundamental principles to animate characters. The next sections present twelve principles of animation mentioned by [12].

Squash and stretch

Squash and stretch is one of the most important principles of animation. Moving objects with a fixed shape appear to be very rigid (e.g. a chair or table). The movement emphasizes the rigidity. However, moving objects that change their shape appear to be flexible. Things made of living flesh always change their shape while moving (e.g. a moving arm with swelling biceps). Living objects are therefore always animated using the squash and stretch principle.

Anticipation

A viewer watching a character will not understand its actions unless these actions are preceded by a clearly planned sequence of other actions. This sequence of preceding actions is known as anticipation. For instance, before throwing a ball the character first swings its arm backwards.

Staging

Staging is the presentation of any idea so that it is completely and unmistakable clear. Actions, personality, expressions and moods can all be staged. If a character is sad you should not play happy music. If you want to stage an action, be sure only one action is seen. If too much happens the user might become confused.

Straight Ahead Action and Pose-to-Pose

There are two basic methods to create animations. The Straight Ahead Action method works from the first drawing to the next until the final drawing is reached. The Pose to Pose method starts with planning the key drawings and then starts to draw the in-betweens. Whereas the first method results often in more spontaneous animations, the second methods results in clearer animations with more strength.

Follow Through and Overlapping Action

Whenever the action of a character is suddenly stopped they appear to be stiff and unnatural. For instance, it is unnatural for a character that throws a ball to stop moving its arm when the ball is just released. Therefore, just like anticipation is used to precede an action, a sequence of actions is needed to end a major action.

A Follow Through is a sequence of actions that follows the major action. For instance, when throwing a ball, the character first moves backwards (anticipation), then throws and releases the ball (major action) and finally its arm overshoots and keeps moving along the direction of movement and finally returns to the end position (follow through).

An Overlapping Action is an action that is caused by the major action. For instance, a dog with long ears that moves and stops (major action) will have ears that keeps moving for a little while (Overlapping Action).

Slow In and Slow Out

The Slow In and Slow Out principle states that the movement of objects seems to become more natural whenever the in-betweens are close to the “extremes” or key drawings. Only one in-between should be drawn between two key drawings. Slow In and Slow Out causes objects to move nonlinear.

Arcs

Movements of living organisms always occur in arcs. Movements in straight lines are very mechanical and seldom performed by characters or parts of characters.

Secondary Action

A Secondary Action is a supplementary action that supports the main action. Secondary Actions make scenes richer and more natural. For instance, a sad person that wipes a tear as he turns away.

Timing

Timing is essential to animation. By changing the number of inbetweens the meaning of a movement can be changed. For instance, a fast eye blink makes the character alert and awake, whereas a slow eye blink makes the character tired.

Exaggeration

Exaggeration is used to accent actions, personalities and expressions in order to make them more convincing. It creates a bigger contact with the audience. So, if a person is sad, make him sadder. If an object is bouncing, make it really bounce.

Solid Drawing

The basics of solid, three-dimensional drawing are weight, depth and balance. Drawings that are very symmetrical (e.g. arms and legs having the same pose) appear to be very stiff and “wooden”. A character becomes more natural when each part of its body varies in some way from the corresponding opposite part.

Appeal

Appeal means anything that a person likes to see, a quality of charm, pleasing design, simplicity, communication and magnetism. Drawings lack appeal when they are too complex or hard to read.

ANIMATING ROBOTS

The traditional way to control a robot’s movements is to use feedback control loops that let the robot react to stimuli from the environment. However, this results in very machine-like behavior − constant velocities and moving in straight lines. More life-like behavior is created by designing pre-programmed motions based on the principles mentioned in the previous section. These motions could be stored in a library and played during the right situation.

In practice, a mixture between pre-programmed motions and feedback loops is required. The pre-programmed motions are designed to make the robot better understandable, whereas feedback loops let the robot react to stimuli from the environment. This mixture has implications for the software architecture of the robot. In the world of games specialized animation layers that combine pre-programmed motions and feedback loops are already known for virtual characters [9]. We applied these ideas to our iCat and developed a set of motions to animate iCat. The motions were carefully designed using the principle of animation mentioned in the previous section. Below, we will discuss two of them: turning iCat’s head to the left and letting iCat fall asleep.

Example 1: Turning to other side

There are many ways in which a character can turn its head, and there are also many reasons why a character would do it. For instance, a character might turn its head because it has been hit, or the character’s attention was caught suddenly, or the character wants to look severe at somebody in order for that person to be silent. In the first case the character will move the head very fast. In the second case the character will move its head quickly and will have its eyes and mouth wide open (probably with fast eye blinks). In the last case the character will move its head slowly, look angry and will have very few eye blinks.

Figure 2 (top) shows iCat turning to the left in a feedback loop-like manner: the robot moves its head with constant velocity. This case simulates for example, the robot tracking an object. The movement is unnatural, especially when we focus on the eyes of iCat. They just look into infinity, something we don't expect from living things. The iCat behaves zombie-like.

Figure 2 (bottom) shows an animated "turn to the left" movement. First, the eyes of iCat move to the left, as if iCat sees something at its left side. This action is used to prepare the user for the major action: turning to the left. An eye blink is added as secondary action to make the scene more natural. Also, all movements (head and eyelids) are performed using slow in and slow outs.

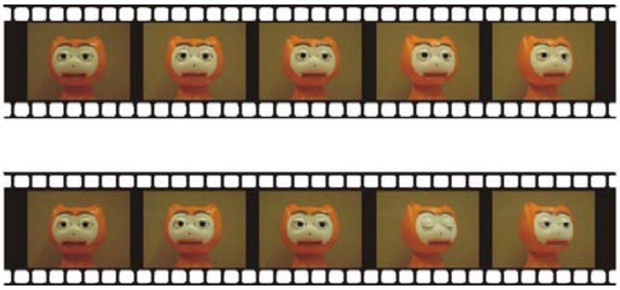

Figure 2. Robot animation example "turning to the left". Top: linear motion, Bottom: applying principles of animation.

Example 2: Falling asleep

Although there is no physiological reason why a user-interface robot would fall to sleep, every living being has a rhythm of being awake and being asleep. Suddenly falling asleep by closing the eyes and lowering the head is unnatural. Therefore, we use again some principles of animation in order to make the scene more natural. Figure 3 shows the movements of the iCat falling to sleep. The scene starts with an anticipating action, namely iCat first yawns. This should prepare the user for the major action: falling to sleep. Again, the falling to sleep movement is carried out by using slow in and slow outs.

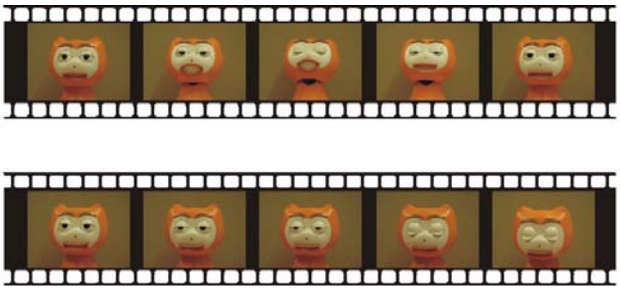

Figure 3. Robot animation example "falling to sleep".

As mentioned in the introduction, besides moving body parts, robot behavior manifests itself also by sound − e.g. music, noises and speech − and light. Robot movements alone are not enough for presenting a particular action. In addition to animated movements, spoken words, previous actions, background sounds and lights all contribute to the staging of an action.

SUMMARY

In this article we argued that the social relationship between a user and a user-interface robot depends on the degree a user is able to understand the robot's behavior and whether it can anticipate it. The way in which a robot moves reveals a lot about what the robot is doing and thinking. Therefore, robot movement needs to be carefully controlled.

Traditionally, feedback loops control the robot's movements. This results however in machine-like behavior that does not tell a user what the robot is doing and why it is doing it. We proposed to apply principles of traditional animation to make the robot's behavior better understandable. Examples from our user-interface robot iCat were given to illustrate this idea.

Using a preprogrammed set of carefully designed animation motions requires a change in the software architecture of the robot. A general architecture that merges pre-programmed motions and feedback control loops is needed.

REFERENCES

1. Aarts, E., Harwig, R., & Schuurmans, M. (2001). Ambient Intelligence. In P.Denning (Ed.), The Invisible Future, pp. 235-250, New York: McGraw Hill.
2. Breazeal, C. (1999), "Robot in Society: Friend or Appliance?". In Agents99 workshop on emotion-based agent architectures, Seattle, WA. 18-26.
3. Breemen, A.J.N. van, Crucq, K., Krose, B.J.A., Nuttin, M., Porta, J.M. and Demeester, E. (2003). "A User-Interface Robot for Ambient Intelligent Environments". In Proc. of ASER 2003, Bardolino, Italy, pp.176-182.
4. D'Azzo, J.J. and Houpis, C.H. (1995). Linear Control System Analysis and Design, McGraw-Hill, Singapore, 4th edition.
5. Diederiks, E. (2003). "Buddies in a Box: Animated Characters in Consumer Electronics" In Proceedings of IUI 03< Miami, USA, pp. 34-38.
6. Evolution Robotics (2004), ER1 Personal Robot System, http://www.evolution.com .
7. Krose, B.J.A., Porta, J.M., Breemen, A.J.N. van, Crucq, K., Nuttin, M. and Demeester, E. (2003). "Lino, the User-Interface Robot", First European Symposium on Ambient Intelligence (EUSAI 2003), November 3-4, Veldhoven, The Netherlands, pp. 264-274.
8. NEC (2004), Personal Robot PaPeRo, http://www.incx.nec.co.jp/robot/PaPeRo/english/p_index.html .
9. Rabin, S. Ed. (2002), AI Game Programming Wisdom, Charles River Media.
10. Sony (2004), AIBO ERS-7, http://www.aibo.com .
11. Sharper Image Design (2004), RoboScout, http://www.roboscout.com .
12. Thomas, F. and Johnson, O. (1981), The Illusion of Life – Disney Animation, Walt Disney Productions.

Effect of Shared-attention on Human-Robot Communication

Junji Yamato
NTT Communication Science Laboratories
3-1 Morinosato-Wakamiya, Atsugi,
Kanagawa 243-0198 JAPAN
yamato@eye.brl.ntt.co.jp

Kazuhiko Shinozawa, Futoshi Naya
ATR Intelligent Robotics and Communication Laboratories
2-2-2 Hikaridai Keihanna Science City
Kyoto 619-0299, JAPAN
shino@atr.co.jp

ABSTRACT
In our pursuit of ways to quantitatively evaluate communication between humans and robots, we recently focused on the effect of shared attention on human decision-making. We used a head robot that can make facial expressions and has human face tracking capability, and designed the interaction so that the robot often looked at the same objects the subjects looked at. Subjects were asked to select a color name from two candidates while the robot made a recommendation. The ratio that subjects selected the recommended candidate was used as a measure of the robot's persuasive communication ability. We found that the matching ratio was correlated with the period time that shared attention (SA) was established for a group of submissive subjects. There was a significant difference in the matching ratio between a high-SA group and a low-SA group. This suggests that establishing shared attention is important for human-robot communication.

Author Keywords
Human-robot interaction, communication robot, shared attention, Social agent, interaction design

ACM Classification Keywords
H5.2. Information interfaces and presentation (e.g., HCI): User interfaces.

INTRODUCTION
The recent growth in computer power has enabled the development of embodied conversational agents. Also being developed are personal robots or pet robots that can serve as communication partners. A computer can be given a personality by using minimal superficial cues [5]. So, how about the robots, and what are the differences and similarities between agents and robots? We have experimentally investigated these questions using agent and robots in a simple color-naming task. We have found that both an agent and a robot can make influences user decision-making and that the foot-in-the door technique also works for agent [12,13]. Another finding is that spatial consistency is important; an agent had more influence when color samples were displayed on a screen and robot had more influence when actual color plates were used [7,8]. While a robot had less influence when color samples were displayed on a screen. So, became clear that a crucial condition is that the robot has to share the same space with users in order to perform as a good persuader. In this paper, we focused on why this is so.

Our hypothesis is this. This condition is good for establishing shared attention of human and robot. When a robot, color sample plates, a button box, and a user are all in the same physical space, not separated in a screen and physical space, shared-attention to the color plate or button box are established naturally through the interaction. On the other hand, when color sample and selection buttons were displayed on screen, it is difficult for user to feel that robot is actually putting attention on these objects while the robot talking about the objects. Shared-attention is considered a key factor in the theory of mind [1], and some robots that are designed for communication have some features to achieve this [3,6]. However, a little quantitative evaluation has been made for shared-attention between humans and robots really matters in human-robot communication. We conducted experiments to measure the effect of shared-attention in human-robot communication by measuring the effect on human decision-making in the color name selection task. The effect was quantitatively measured and statistically significant for a specific type of subject.

EXPERIMENT
The experiment was designed to quantitatively measure the influence of robot's recommendation on human decision-making. We used the color name selection task. Subjects looked at the color sample plate, and were asked to select

the color name from two candidates. All colors in the task were ambiguous and some names were not so familiar to ordinary people. For example, carmine or vermilion were displayed as candidates for a bright red color plate, although all words were actually Japanese. The answer was not obvious and most subjects had no prior reference. Before a subject made decision on any given question, a robot in front of the subject made a recommendation supporting one candidate. Subjects could accept the recommendation or not based on their own preference. The experimental set ups are shown in Figure 1.

The ratio that subjects selected the same candidate the robot recommended (matching ratio) was the measure of the robot's persuading capability or the measure of the goodness of the communication between a subject and the robot. The robot recommended the same candidates for all subjects. Which one it would recommend was determined beforehand based on a preliminary experiment. We chose candidates so that expected average matching ratio would become around 0.5 for no-recommendation condition.

Subjects were 28 people (14 male and 14 females), aged from 21 to 29. The average age was 24.0 and the standard deviation was 1.83. There was no instruction as to how they should handle the robot's recommendations. Each subject saw 30 color plates in total, and the order of presentation was the same for all subjects.

Besides performing the color name selection task, subjects took a personality-profiling test to categorize them according to certain personality traits for detailed analysis. We employed the TEG (Tokyo University Egogram)[10], which was developed by the Medical School of the University of Tokyo based on transactional analysis [9] using principal component analysis of a large-scale public survey. The TEG consists of 60 questions and measures five personality factors: CP (critical parent), NP(nurturing parent), A(adult), FC(free child), and AC(adapted child). Each factor is scaled from 0 to 20.

Achievement of shared-attention was evaluated from recorded video after the experiments. The gaze direction of the robot was classified into four classes: subject, color plate, button box, and other. Figure 2 shows the robot for each class of gaze. The gaze of the subjects was classified as robot face, robot (other), color plate, button box, or other. When both looked at the color plate or button box simultaneously, shared-attention was considered to be achieved. Our original expectation was that the effect of robot's recommendation, measured by the matching ratio, would correlate with the period of time shared-attention was established.

Robot

We used a head robot with a human face tracking feature [11]. The robot was build by the MIT Artificial Intelligence Laboratory as a new version of Kismet [2]. It has two eyes with pan and tilt control, two eyebrows with up and down

Figure 1 Experimental set ups.

control, eyelids, a mouth with expressive lips (two degrees of freedom (DoF) for each), and two fan-like ears for expression of its emotional state. It also has neck movements with three DoF. The robot has a video camera in each eye, and another one in the center of its face.

The vision system of the robot can extract and track the skin color region in the captured image. We used this feature to establish eye contact with the subject. The speech of the robot was generated by the "Fluet" Japanese speech synthesizer [4] developed by NTT.

RESULTS

The matching ratio was 0.57. This was higher than a reference group in a no-recommendation condition. However, there was no statistically significant difference.

There was no correlation between the matching ratio and the amount of shared-attention time (SA time) for all subjects. This implies shared-attention did not influence on human-decision making in general.

However, an interesting interaction was found. A high-AC group of subjects whose AC factor of TEG test was more than 11 (75 percentile in standard distribution) demonstrated a strong correlation between the matching ratio and SA time (Speaman's $r=0.51$, $p=0.051$. see Figure 3). On the other hand, for a low-AC group, the correlation was negative, but it was not statistically significant.

Based on this finding, the matching rations of the high-SA group and low-SA group among high-AC subjects were compared. As shown in Figure 4, the matching ratio for the high-SA group was higher. The difference was statistically significant by t-test ($p<0.05$)

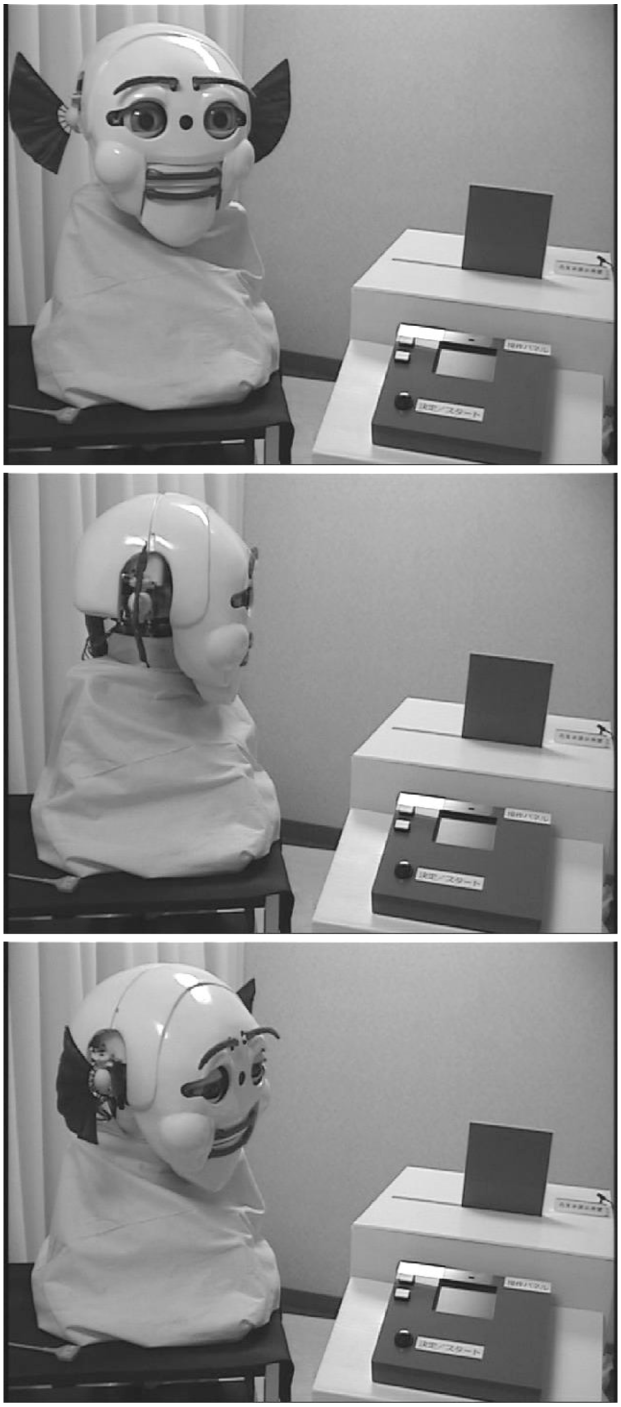

Figure 2. Robot's gaze.

(From the top, subject, color plate, and button box)

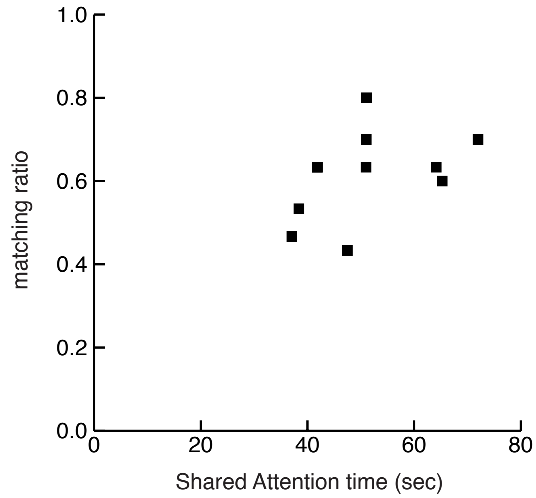

Figure 3. Correlation of SA-time and matching ratio.

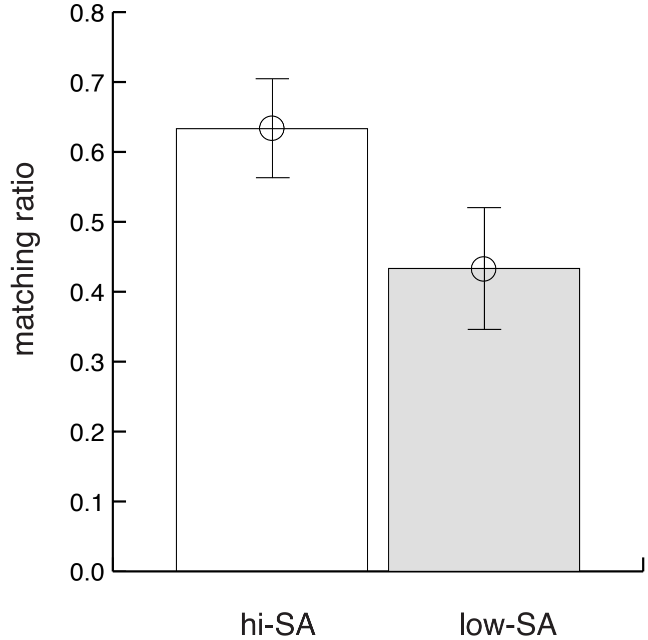

Figure 4. Matching ratios of high-SA and low-SA groups.

Discussion

This result shows that, as in human-human communication shared-attention is important for human-robot communication. Shared-attention had an influence on user's behavior, but not always a positive one. The kind of influence depended on the personality type of the user. Because high-AC people are considered to be submissive or obedient, this result might sound trivial: the high-AC group was obedient, so of course it chose the recommended one. But it was not so simple. The average matching ratios of high-AC group and low-AC group were not so different; 0.59 and 0.54. Furthermore, the average SA time of the high-AC group was less than that of the low-AC group; 50.2 sec versus 53.4 sec. Distributions of the two groups mostly overlapped. But one has strong positive correlation, and the other negative. This means that the behavior of the high-AC group was influenced by how much time they established shared-attention with the robot. This indicates the effect of shared-attention in communication

Another point is that even a pseudo shared-attention worked well in this experiment. The robot can track a subject's face to establish eye contact using the skin-color detection feature, but it cannot detect subject's eye direction. Therefore, based on preliminary experiments, we designed the robot's behavior so that it looked at where subjects tend to look. So the shared-attention in these experiments was rather stochastically achieved. This suggests that a robot can pretend to follow a user's gaze using an information source other than actual eye direction in real time, which is still difficult to robustly measure from video images. This is important in designing effective interactive robot actions, especially when the robot has limited performance in terms of image acquisition or image processing.

CONCLUSION

We measured the effect of shared-attention on human-robot communication and found that the amount of time that shared-attention was achieved has a positive correlation with the strength of the effect on human decision-making. Although the results are limited to a specific type of subjects (a high-AC factor (adapted child) group), they could be useful in designing more persuasive communication robots.

REFERENCES

1. Baron-Cohen, S. (1995). *MIND BLINDNESS.* Cambridge, MA: The MIT Press.
2. Breazeal, C. (2002), *Designing Sociable Robots* Cambridge, MA, MIT Press.
3. Kozima, H., Nakagawa, C., and Yano, H. Emergence of Imitation Mediated by Objects, *Proc. The Second International Workshop on Epigenetic Robotics (2002).*
4. Mizuno, O. & Nakajima, S. (1998) Synthetic Speech/Sound Control Language: MSCL. 3rd ESCA/COCOSDA *Proc. International Workshop on Speech Synthesis,* pp 21-26.
5. Reeves, B., & Nass, C. (1996). *The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places.* Stanford, CA: CSLI Publications.
6. Scassellati, B. Theory of Mind for a Humanoid Robot. *Proc. the first IEEE/RSJ International Conference on Humanoid Robotics,* (2000)
7. Shinozawa, K., Naya, F., Yamato, J., & Kogure, K. (2001). Quantitative evaluation of effect of embodied conversational agents on user. *Proc. HCI International ' 01.*
8. Shinozawa, K., Naya, F., Yamato, J., & Kogure, K. (2001). Difference in Effect of Recommendation made by Robot and Screen Agent. *Proc. SI2002.* SICE (In Japanese)
9. Stewart, I., and Joines, V. (1987) *TA Today: A New Introduction to Transactional Analysis.* Lifespace Publishing, Nottingham, England
10. TEG Research Group in the University of Tokyo Medical school, *TEG: Tokyo University Egogram (New Ver.) ,* Kaneko Shobou(2000).
11. Yamato, J., Shinozawa, K., Brooks, R., and Naya, F. Human-Robot Dynamic Social Interaction. *NTT Technical Review* 1, 6(2003), 37-43.
12. Yamato, J., Shinozawa, K., Naya, F. and Kogure, K. Effects of Conversational Agent and Robot on User Decision. *The IJCAI-01 Workshop on Autonomy, Delegation, and Control: Interacting with Autonomous Agents*
13. Yamato, J., Shinozawa, K., Naya, F. and Kogure, K. Evaluation of Communication with Robot and Agent: Are robots better social actors than agents? *Proc. INTERACT'01* IOS Press (2001), pp.690-691

Applying CSCW and HCI Techniques to Human-Robot Interaction

Jill L. Drury
The MITRE Corporation
Mail Stop K320
202 Burlington Road
Bedford, MA 01730
jldrury@mitre.org

Jean Scholtz
National Institute of Standards and Technology
100 Bureau Drive, MS 8940
Gaithersburg, MD 20899
jean.scholtz@nist.gov

Holly A. Yanco
Computer Science Department
Univ. of Massachusetts Lowell
One University Avenue
Lowell, MA 01854
holly@cs.uml.edu

ABSTRACT
This paper describes our approach for human-robot interaction (HRI) research and summarizes our progress to date. We have concentrated on HRI in urban search and rescue (USAR) because it is an example of a safety-critical application. We analyzed the performance of robotic teams at two USAR robotics competitions using adaptations of techniques from the human-computer interaction (HCI) field and determined that problems experienced by the operators or robots could be traced to a lack of awareness on the part of the operator of the robots' status, location, or immediate surroundings. To aid analysis, we developed a taxonomy of HRI-related characteristics, evaluation guidelines, a coding scheme that categorizes HRI activities, and a fine-grained definition of HRI awareness based on awareness research from computer-supported cooperative work (CSCW). As a result, we are beginning to determine design guidelines for HRI that are being used in developing next-generation robots at the University of Massachusetts Lowell.

Author Keywords
Human-robot interaction (HRI), Computer-Supported Cooperative Work (CSCW), awareness, urban search and rescue (USAR), human-computer interaction (HCI) evaluation techniques.

ACM Classification Keywords
H.5.2 User Interfaces, Evaluation/Methodology

INTRODUCTION
Much progress has been made in robotics in the last decade. For example, roboticists have worked hard to improve communications between humans and robots (and also between robots), the variety and fidelity of sensors on-board the robot, the ability of the robot to traverse rough terrain, and the level of autonomy that robots are able to achieve. By comparison, relatively little progress has been made in optimizing the partnership between people and robots through improved techniques for human-robot interaction. To address this gap, our research partnership includes robotics, HCI, and CSCW expertise.

We chose to focus on USAR robots because they are a prime example of a class of safety-critical situations: situations in which a run-time error or failure could result in death, injury, loss of property, or environmental harm [Leveson 1986]. Safety-critical situations, which are usually also time-critical, provide one of the bigger challenges for robot designers due to the vital importance that robots perform exactly as intended and support humans in efficient and error-free operations.

The rest of this paper describes our methodology, analysis frameworks, results, and future work.

METHODOLOGY
There are few opportunities to study USAR operations in real disaster situations. Thus, we have used a strategy based on usability tests and robotics competitions.

We have arranged for typical users of USAR robotics to perform rescue tasks in a mock-up of a disaster situation, taking place in NIST-developed test arenas that simulate a building with various levels of destruction [Jacoff et al, 2000; Jacoff et al, 2001]. Consonant with traditional HCI usability testing, we ask participants to "think aloud" [Ericsson and Simon 1980] as they perform rescue tasks, enabling us to identify those portions of the interface that hinder participants or impede efficiency.

However, to date, most of our opportunities to study HRI have come in conjunction with USAR robotics competitions. These opportunities differed from traditional usability testing in two significant ways. First, the robot developers operated the robots (rather than members of the rescue professions). We viewed performance, therefore, as an upper bound: if the robot developers had problems with a part of the interface, it is likely that any other user would also have difficulties. Second, we were restricted to being silent observers who could not ask the operators to do anything differently during the competition than they would have already done. To a least partially make up for a lack of “thinking aloud,” our observer performed a quick debriefing of the operators via a short post-run interview to obtain the operators’ assessment of their (and the robots’) performance. In addition, we were given the scoring materials from the competition judges that indicated where victims were found and penalties that were assessed.

HRI Taxonomy

To better understand the different types of HRI, we developed a taxonomy to characterize robotic interaction [Yanco and Drury 2002]. Besides determining the classification categories, we defined values to describe each classification. The list of classification categories and their description is contained in Table 1 (values are omitted here due to space limitations).

ANALYSIS FRAMEWORKS

We feel that some of our more important contributions to HRI are our analysis frameworks, since they may help other researchers, robotic designers and evaluators to better understand when and how HRI can be improved.

We have used three different mechanisms to structure our analyses: a detailed definition of HRI awareness, a coding scheme for HRI activities, and Scholtz’ [2002] evaluation guidelines. Each of these mechanisms is discussed below. All three led us to focus on “critical incidents,” which we defined as cases in which the robot, USAR victims, or environment sustained actual or potential damage or harm.

Classification	Description
Autonomy	% time a robot performs a task on its own
Amount of intervention	% time that a human operator must control a robot
Human-robot ratio	The ratio of operators to robots.
Level of shared interaction	Various combinations of whether the humans and robots act independently or as part of team(s).
Composition of robot teams	Whether teams of multiple robots are homogeneous or heterogeneous.
Available sensors	A list of sensor types available on the robot platform.
Sensor fusion	A list of functions mapping the sensor data to the fused output.
Criticality	The importance of getting the task done correctly in terms of its negative effects should problems occur.
Time	Whether the humans and robots work together in the same time (synchronously) or different times (asynchronously).
Space	Whether the humans and robots work together in the same place (collocated) or in different places (non-collocated).

Table 1: Taxonomy for Human-Robot Interaction [Yanco and Drury 2002]

HRI Awareness

Much research has been performed in the CSCW community to characterize awareness. While there are many definitions of awareness in the CSCW literature (see Drury, Scholtz, and Yanco [2003] for a summary), we started with the definition in Drury [2001], the informal version of which is: awareness in a multi-user computing system is a participants’ understanding of the presence, identities, and activities of another participant. There are two differences between CSCW and robotic systems that affect how awareness can be understood, however. The first difference is the fact that CSCW addresses multiple humans working together, whereas HRI can involve single or multiple humans working with single or multiple robots. The second is that human participants will bring at least a minimum level of free will and cognitive ability to the collaboration that cannot be brought by the robotic participants. Thus the HRI awareness framework must account for all combinations of single and multiple humans and robots, and must accommodate the non-symmetrical nature of the human-robot collaboration. The simplest case of HRI occurs when one human works with one robot.

> *HRI awareness (base case)*: Given one human and one robot working on a task together, HRI awareness is the understanding that the human has of the location, activities, status, and surroundings of the robot; and the knowledge that the robot has of the human’s commands necessary to direct its activities and the constraints under which it must operate.

Note that greater or lesser amounts of HRI awareness are needed depending upon the level of autonomy that the robot achieves, so the expectations of awareness need to be tailored for the anticipated level of autonomy. The HRI awareness base case can be generalized to cover multiple humans and robots coordinating in real time on a task.

> *HRI awareness (general case):* Given n humans and m robots working together on a synchronous task, the general case of HRI awareness consists of five components:

- <u>Human-robot</u>: the understanding that the humans have of the locations, identities, activities, status and surroundings of the robots. Further, the understanding of the certainty with which humans know the aforementioned information.
- <u>Human-human</u>: the understanding that the humans have of the locations, identities and activities of their fellow human collaborators.
- <u>Robot-human</u>: the knowledge that the robots have of the humans’ commands necessary to direct their activities and any human-delineated constraints that may require a modified course of action or command noncompliance.

- Robot-robot: the knowledge that the robots have of the commands given to them, if any, by other robots, the tactical plans of the other robots, and the robot-to-robot coordination necessary to dynamically reallocate tasks among robots if necessary.

- Humans' overall mission awareness: the humans' understanding of the overall goals of the joint human-robot activities and the moment-by-moment measurement of the progress obtained against the goals.

In human-robot awareness, "activities" refer to such phenomena as speed and direction of travel and progress towards executing commands. Status information includes battery power levels, the condition of sensors, etc.

Sufficient HRI awareness is needed to ensure smoothly functioning human-robot coordination on a shared task. When insufficient HRI awareness is provided, we say this is an HRI awareness violation:

> *HRI awareness violation:* HRI awareness information that should be provided is not provided.

There are five possible types of HRI awareness violations, corresponding to the five types of HRI awareness defined above. We discussed the results from a USAR competition in terms of types of awareness violations that occurred during critical incidents in Drury, Scholtz and Yanco [2003].

Coding Scheme

To help in analyzing videotapes of the robot competitions and usability test runs, we noted each critical incident and categorized it in terms of the type of HRI awareness violation that occurred (if one was present) and the type of task being attempted at the time of the incident.

Because all cases that we analyzed so far concerned a single operator and one or more robots that did not coordinate with each other, HRI awareness problems consisted solely of human-robot awareness violations. We anticipate that more of the HRI awareness framework will be employed when we analyze more diverse configurations.

We defined five types of tasks relating to critical incidents.

> Local navigation: An operator is navigating in constrained or tight situations, and encounters difficulty because of the constraints. An example of a local navigation problem is when the robot slips down a ramp or bumps a wall.

> Global navigation: An operator is navigating in all other situations. An example of a global navigation problem is when an operator does not have a clear understanding of the robot's position, potentially leading to driving the robot out of the arena unintentionally or covering areas already searched.

> Obstacle encounter: An operator is working to free the robot from an obstacle; the robot is hindered in moving towards a goal.

> Victim identification: An operator is attempting to characterize the state of a victim (e.g., conscious or not, warm or cold, speaking or silent, moving or not moving). An example of a problem occurring during victim identification is inaccurate interpretation of sensor data.

> Vehicle state: An operator is attempting to perform USAR tasks despite the fact that the robot is in a degraded state (e.g., it is not stable or upright or its sensors are impaired or broken).

We analyzed data from a USAR competition using this coding scheme; the results are summarized in Scholtz, Young, Drury, and Yanco [in submission].

Scholtz's Guidelines

Scholtz [2002] developed six evaluation guidelines for evaluating HRI. We treated these guidelines as heuristics to be tailored for USAR systems (Nielsen [1993] recommends tailoring heuristics to be appropriate to the systems being evaluated). After tailoring (including combining two of the guidelines into one heuristic), we evaluated the robotic systems in a major USAR competition against the following:

Is sufficient status and robot location information available so that the operator knows the robot is operating correctly and avoiding obstacles?

Is the information coming from the robots presented in a manner that minimizes operator memory load, including the amount of information fusion that needs to be performed in the operators' heads?

Are the means of interaction provided by the interface efficient and effective for the human and the robot (e.g., are shortcuts provided for the human)?

Does the interface support the operator directing the actions of more than one robot simultaneously?

Will the interface design allow for adding more sensors and more autonomy?

A discussion of how we tailored these heuristics, plus our results after applying the heuristics, is contained in Yanco, Drury, and Scholtz [to appear].

RESULTS TO DATE

We found that all critical incidents could be traced to HRI awareness violations. Thus, when we developed a preliminary set of guidelines for designing interfaces for HRI [Yanco, Drury, and Scholtz, to appear], we began with awareness and also included guidelines to address the other major problems we observed with HRI:

Enhance awareness. Provide a map of where the robot has been. (Operators using systems with maps were more successful in navigating the arena.) Also, provide more spatial information about the robot in the environment; operators must be aware of their robots' immediate surroundings to avoid bumping into obstacles or victims.

Lower cognitive load. Provide fused sensor information to avoid making the user fuse the data mentally.

Increase efficiency. Provide user interfaces that support multiple robots in a single display/window. In general, minimize the use of multiple windows. With additional sensor fusion, more information could be displayed in a single window, which is more efficient for users than having to switch between windows.

Provide help in choosing robot modality. Provide the operator assistance in determining the level of robotic autonomy that would be most appropriate for a given situation.

DISCUSSION AND FUTURE WORK

One of the primary goals of our further research is to expand and refine our set of design guidelines. We have taken the design guidelines developed so far and are in the process of applying them to new robots and interfaces being developed at the University of Massachusetts Lowell [Hestand and Yanco, in submission].

We found coding to be very difficult at times. Our first attempt at coding (not described in this paper) involved accounting for every second of human/robot activities; we found that the detailed data did not yield as many insights as hoped. In contrast, the scheme described in this paper concentrated on characterizing anomalous behavior, analogous to an HCI expert concentrating on users' problems operating interfaces during usability testing. We anticipate that the coding scheme will likely evolve further.

We plan to expand our use of HCI analytical and inspection evaluation techniques. For example, we anticipate performing a Goals, Operators, Methods, Selection rules (GOMS) analysis of several robotic systems.

Few of the robots studied so far include much autonomy. We plan to investigate HRI under varying levels of autonomy, especially via usability testing.

As we evaluate systems that include multiple humans and robots that communicate with each other, we plan to more fully exercise the HRI awareness framework and determine whether it should evolve.

We also plan to refine the taxonomy. By characterizing the robotic system in a useful way, we hope to be able to use the taxonomy to roughly predict the likely level of efficiency and cognitive load.

ACKNOWLEDGMENTS

The work was supported in part by NSF IIS-0308186, NIST 70NANB3H1116, and the DARPA MARS program.

REFERENCES

1. Drury, J. (2001). Extending Usability Inspection Evaluation Techniques for Synchronous Collaborative Computing Applications. Sc.D. Thesis, University of Massachusetts Lowell, Department of Computer Science.
2. Drury, J., J. Scholtz and H.A. Yanco (2003). Awareness in Human-Robot Interactions. In Proceedings of the IEEE Conference on Systems, Man and Cybernetics, Arlington, VA, October 2003.
3. Ericsson, K. A. and H. A. Simon (1980). "Verbal reports as data." *Psychological Review* **87**: 215 - 251.
4. Hestand, D. and H.A. Yanco. Layered Sensor Modalities for Improved Feature Detection. Submitted to ICRA-04.
5. Jacoff, A., E. Messina, and J. Evans (2001). "A reference test course for autonomous mobile robots." In *Proceedings of the SPIE-AeroSense Conference*, Orlando, FL, April 2001.
6. Jacoff, A., E. Messina, and J. Evans (2000). "A standard test course for urban search and rescue robots." In *Proceedings of the Performance Metrics for Intelligent Systems Workshop*, August 2000.
7. Leveson, N.G. (1986). "Software safety: why, what and how." *ACM Computing Surveys* **18**(2): 125 – 162, June 1986.
8. Nielsen, J. (1993). Usability Engineering. Chestnut Hill, MA, AP Professional Press.
9. Scholtz, J. (2002). "Evaluation methods for human-system performance of intelligent systems". In *Proceedings of the 2002 Performance Metrics for Intelligent Systems (PerMIS) Workshop*.
10. Scholtz, J., J. Young, J.L. Drury, and H.A. Yanco. Evaluation of Human-Robot Interaction Awareness in Search and Rescue. Submitted to 2004 IEEE International Conference on Robotics and Automation.
11. Yanco, H.A. and J.L. Drury (2002). "A taxonomy for human-robot interaction." In *Proceedings of the AAAI Fall Symposium on Human-Robot Interaction*, November 2002, Falmouth, MA. AAAI Technical Report FS-02-03, pp. 111 – 119.
12. Yanco, H.A., J.L. Drury and J. Scholtz. Beyond Usability Evaluation: Analysis of Human-Robot Interaction at a Major Robotics Competition. To appear in the Journal of Human-Computer Interaction.

Investigating socially interactive robots that give the right cues and make their presence felt

Position Paper for CHI-04 Workshop on Shaping Human-Robot Interaction

Helge Hüttenrauch, Kerstin Severinson Eklundh
Royal Institute of Technology
S-10044 Stockholm, Sweden
{hehu, kse}@nada.kth.se

ABSTRACT

In order to make socially interactive robots usable and enjoyable, it is mandatory to understand how their locomotion, body language, approach and positioning as well as (non-)verbal communicative cues affect users. The attributes of robots' behavior and the dynamics of the mutual adaptation between user and robot need to be identified and the different interaction-steps user-tested empirically. The design for socially interactive robot behaviors ought to be guided by findings in the human sciences about human-to-human social behavior. To initiate a discussion about the design and investigation into socially interactive robots' behavior our view upon a possible approach and possible interaction scenarios to test with is put forward.

Author Keywords

Human-Robot Interaction, embodied interaction, posture and positioning, service robots, user-testing

INTRODUCTION

During long-term user testing [8] and experimentations with the Cero service robot, designed to transport light objects in an office environment, we repeatedly observed how the robot's movement behavior, its turning, and spatial positioning in the environment was part of the unfolding interaction and reaction towards the robot: The ultrasound-sensor guided corridor-following behavior, which tries to keep the robot in the middle of a corridor to traverse, sometimes swiveled the robot slightly so that people encountering it at that moment, had trouble to decide on which side to pass. In another situation the robot was inhibited by a doorstep to enter into an office, thus the user was approaching to robot for the loading and unloading of it instead. When this doorstep had be removed by the facilities management, the robot was nevertheless made (by the interaction designers) to stop in front of the office and ask the user whether the door was open wide enough for entering (as this was not handled by robot on-board sensors). However, this behavior of first standing and asking for allowance to enter the room was also interpreted as a social function: If a co-worker tries to enter, the social norm is to first knock on the door, stick in the head and ask whether one could interrupt. The robot was interpreted as quasi mimicking this behavior.

The collected evidence showed three effects: (1) the Cero robot's movement behavior is interpreted and prompts re-actions (e.g. its moving, avoiding obstacles or approaching people), (2) its overall behavior and communication (e.g. speech interaction) is related to social norms of human-to-human behavior, communication, and interaction, and (3), handing-over objects requires fine detailed spatial management of both the robot and its user.

We want to study these effects more systematically to understand how the movement behavior, communicative cues, and negotiation of the spatial relationship might contribute to the design for enjoyable interaction with service robots. What patterns of body language would make it easier for robots to convey their intention? What re-actions by human users can robots expect to encounter in approaching people (or being approached by them) and then perhaps taking up contact with spoken language? What non-verbal communication cues can be put to work in handling the evolving interaction?

As this work has just start up in the framework of "Cognitive robot companion" (COGNIRON) [1], we will point out some attributes that will likely influence the mutual perception of movement, spatial relation, and communicative cues as backdrop. We will then sketch and reason in more detail our research approach strategy to prepare for some closing questions to the research community which we would like to discuss.

INFLUENCTAL FACTORS

The potential design space for socially interactive robots that react according to users expectations as far as spatial distance, orientation, approachability, and possibly, intent is concerned, seems infinite. A reduction according to some classification of involved attributes might be helpful, even if their implications can not be discussed here in detail.

Socially interactive robots come in many shapes, suit a wide variety of purposes, and are subject to many challenging research questions [6]. This points out that their expressive form [3] (or designed look and feel) is highly influential to the robot' exhibited and user perceived motions and communicative cues. For example is the naturalness in communication or life-likeness of imitating agents, e.g. in the form of humanoid robots, questioned in [2] and [12].

An alternative approach could be to look upon the task, purpose, or domain the robot is to be employed in: Robots as diverse as systems for house cleaning [10], therapy [14], education and/or entertainment [5] will with benefit have the different socially interactive behavior.

Another dimension in the design space are the indented users, their prior experience, expectations in the robot, and level of interaction to be achieved. Human-Robot interaction user taxonomies like [11] may be helpful in designing social robot behaviors that are targeted at certain user groups like operators, peers, or bystanders.

The embodied interaction [4] by the robot and the mobility of both the robot and its users are closely related factors: Spatial, interpersonal distances, orientation, and signaled approachability define how human-robot encounters will be conducted in space. The space might be further defined by attributes such as 'open', 'closed', 'public', 'private', or 'intelligent' [8] and have connotations of special places [13], like kitchen which define not only a certain type of room, but equally what kind of actions might take place there. Note however that the term 'space' might be too limited: The interaction situation and context will not only be defined by area the interaction encounter will happen in, but also the temporal patterns of it.

This temporal aspect becomes obvious with another involved domain: The investigation into socially interactive robots can be approached from a communicative theory perspective: Spoken language, non-verbal communication (or visual cues), dialogues, the use of body language and gestures happen in short-term time-frames, not in separation, but as multi-modal communication.

If robots, as a long-term goal, are to be perceived as socially competent the grounding of this perception is to be found in the human sciences – e.g. psychology, sociology, or anthropology – and their findings, e.g. for human-to-human interpersonal relationships. This strategy however assumes that the findings in human sciences can actually be transferred from a human-to-human context to a human-robot one. E.T. Hall's [7] categories of interpersonal distances might suit as an example. Hall showed that there are different spatial distances (i.e. *intimate, personal, social,* and *public* distance) at which humans feel comfortable to encounter and interact with on another, dependent upon the situational context, the intended interaction and communication.

The different identified interpersonal distances at which people feel comfortable with could be taken as a test-bed to investigate, if e.g. the personal or social distance zones identified in human-to-human encounters can be used in human-robot-interaction, too.

RESEARCH APPROACH

From the influencing factors just described the multi-dimensionality of the design and investigation space for socially interactive robots was argued. How is one than to conduct research into the field without drowning in the sea of possibilities? How can investigations lead to findings that can be generalized and not only describe individual cases?

We propose to start off by conducting explorative studies that try to have users experience socially interactive robots in prototypical robot-encounter situations. This collection of empirically observations is argued for at this state as 1) real data of user encounters with socially interactive robots is scarce and 2) can suit as a valuable guidance to stay focused on the human-robot-interaction, instead of the development of technological components (e.g. sensory systems) in isolation.

The most simplistic case to look at seems to be that one human is to encounter one robot where either the robot approaches the user, the user approaches the robot, or both one another simultaneously. The prototypical situation could incorporate specially designed for situations, like encountering and performing a special task, or following with a robot.

With such a scenario, interesting concepts, processes, dynamic mutual adaptations, treatment variables etc. could be tested for practically. One could observe, quantify, and analyze different measurements of interaction and social cues as given, e.g. in speech, by eye-gaze exchanges, different gestures, the use of body-language to name but a few. Furthermore, observations can be related to subjective measurements inquiring the *user experience* of the situational encounters with certain robot behaviors.

QUESTIONS FOR DISCUSSION

The field of research into socially interactive robots is challenging and untested. Accordingly, the motivation, background, and initial investigation strategy presented in this position statement might still have flaws that we would like to have pointed and worked out in cooperation with interested researchers and developers. We phrased our as questions, e.g.

What are the heuristics of robot-human body-postures and positioning, including communicative clues of socially interactive mobile agents?,

Will the perception of such agents and their behavior feel natural and provide a type of socially interactive robot-genre?,

What are the truly salient interaction patterns and situations of socially interactive human-robot-interaction?,

What are appropriate robot application for such robots – just entertainment, therapy, and educational ones?,

If the robot and the human capabilities in perceiving and producing social cues and interpersonal behaviors is unbalanced (the human being more competent), how can this imbalance be communicated to users to avoid disappointing unrealistic user expectations?, and finally

If findings in human-robot interaction are supposed to be limited to certain cultural settings – how can more universal results be achieved?.

To repeat the obvious once more: The format of this presentation does only allow for a incomplete treatment, its intention is to raise the issues and allow for a shared exchange of ideas and thoughts.

REFERENCES

1. COGNIRON, (2003). The Cognitive Robot Companion - Annex 1 Description of Work, 6th Framework FET Proactive initiative “Beyond Robotics”. FP6-IST-002020.
2. Dautenhahn, K. (2002). Design spaces and niche spaces of believable social robots. Robot and Human Interactive Communication, 2002. Proceedings. 11th IEEE International Workshop on , 25-27 Sept. 2002 pp:192 - 197
3. DiSalvo, C.; Gemperle, F.; Forlizzi, J.; Montgornery, E. (2003). The hug: an exploration of robotic form for intimate communication. Robot and Human Interactive Communication, 2003 (ROMAN 2003). The 12th IEEE International Workshop on, pp: 403 – 408
4. Dourish, Paul. (2001). Where the Action Is: the Foundations of Embodied Interaction. A Bradford Book, The MIT Press.
5. Druin, A., & Hendler, J. (Eds.), (2000). Robots for kids : exploring new technologies for learning. San Mateo, CA: Morgan Kaufmann.
6. Fong, T., Nourbakhsh, I., Dautenhahn, K. (2003). A Survey of Socially Interactive Robots, Robotics and Autonomous Systems, Vol. 42, (3-4), March 2003, pp.143-166. Elsevier Science.
7. Hall, E. T. (1966). The hidden dimension : man ’s use of space in public and private. Bodley Head, London.
8. Hashimoto, H. (2002). Intelligent space - how to make spaces intelligent by using DIND? Systems, Man and Cybernetics, 2002 IEEE International Conference on, Vol. 1, pp:14 – 19.
9. Hüttenrauch, H. and Severinson Eklundh, K. (2002). Fetch-and-carry with CERO: observations from a long-term user study with a service robot. In Proc.11th IEEE Intern. Workshop on *Robot and Human Interactive* Communication (RO-MAN 2002). pp.158–163.
10. RobotCleaner 3000, Kärcher product information, on-line: http://www.robocleaner.de/english/work1.html
11. Rogers, E., and Murphy, R. (2001), DARPA/NFS Study on Human-Robot Interaction, Invitation Workshop, Sept.29-30, 2001, San Luis Obispo, CA; see www.csc.calpoly.edu/~erogers/HRI/ for online Final Report (as of: 2003-12-19).
12. Schmidt, C.T. (2002). Socially interactive robots. Why our current beliefs about them still work. Robot and Human Interactive Communication, 2002. Proceedings. 11th IEEE International Workshop on , pp:560 – 564.
13. Svensson, M. (2003). Defining, Designing and Evaluating Social Navigation, *Doctoral thesis, SICS-P-333-SE*, Dept. of Computer and Systems Sciences, Stockholm University / Royal Institute of Technology.
14. Wada, K., Shibata, T., Saito, T., and Tanie, K. (2003). Effects of robot assisted activity to elderly people who stay at a health service facility for the aged. Intelligent Robots and Systems, 2003. (IROS 2003). Proceedings. IEEE/RSJ International Conference on, Vol. 3, pp:2847-2852.

Robots that Work in Collaboration with People

Guy Hoffman
MIT Media Lab
77 Massachusetts Ave, NE18-5th floor
guy@media.mit.edu

Cynthia Breazeal
MIT Media Lab
77 Massachusetts Ave, NE18-5th floor
cynthiab@media.mit.edu

ABSTRACT
Developing robots with social skills and understanding is a critical step towards enabling them to cooperate with people as capable partners, to communicate with people intuitively, and to learn quickly and effectively from natural human instruction. These abilities would enable many new and exciting applications for robots that require them to play a long-term, supportive, and helpful role in people's daily lives. This paper describes our work towards building sociable autonomous robots that can work in collaboration with people. Our approach puts an emphasis on task dialog and social communication under the theoretical framework of joint intention theory.

INTRODUCTION
Many of the most useful and new applications for autonomous robots require them to work alongside people as capable, cooperative, and socially savvy partners. For instance, robots are being developed to provide the elderly with assistance in their homes. Such a robot should be persuasive in ways that are sensitive to the person, like reminding them when to take medication, without being annoying or upsetting. In other applications, robots are being developed to serve as members of human-robot teams. NASA JSC's *Robonaut* is a great example [1]. This humanoid robot is envisioned to work shoulder-to-shoulder with astronauts assisting them in space station maintenance operations.

To provide a human teammate with the right assistance at the right time, a robot partner must not only recognize what the person is doing (i.e., his observable actions) but also understand the intentions or goals being enacted. This style of human-robot cooperation strongly motivates the development of robots that can infer and reason about the mental states of others within the context of the interaction they share.

For applications where robots interact with people as partners, it is important to distinguish **human-robot collaboration** from other forms of human-robot interaction (HRI). Namely, whereas interaction entails acting *on* someone or something else, collaboration is inherently working *with* others [2,3].

Much of the current work in human-robot interaction is thus aptly labeled given that the robot (or team of robots) is often viewed as an intelligent tool capable of some autonomy that a human operator commands (perhaps using speech or gesture as a natural interface) to perform a task [4,5]. This sort of master-slave arrangement does not capture the sense of partnership that we mean when we speak of working "jointly with" others as in the case of collaboration.

Human robot collaboration has been studied using autonomous vision-based robotic arms [6] and teleoperated humanoids, such as NASA JSC's *Robonaut*. In other teleoperation work, the notion of partnership has been considered in the form of *collaborative control* (e.g. [7]), allowing a robot to ask a human for help in resolving perceptual ambiguities. In this approach, the human is used by the robot as a source of information rather than as a peer. In our work, the robot is autonomous and works with the human as a peer and a member of a collocated team to accomplish a shared task.

True human-robot collaboration is thus a relatively unexplored kind of human-robot interaction. This paper describes how we apply our theoretical framework (based on joint intention theory) to enable an expressive humanoid robot, *Leonardo*, to work shoulder-to-shoulder with a human teammate towards accomplishing a joint task. To this end, *"Leo"* uses collaborative discourse, gesture, and accompanying social cues. Leo is shown in Figure 1(a).

JOINT INTENTION THEORY
What characteristics must a robot have to work effectively with its human collaborator? To answer this, we look to insights provided by *Joint Intention Theory* [8]. According to this framework, joint action is conceptualized as doing something together as a team where the teammates share the same goal and a common plan of execution. This *collaborative plan* does not reduce to the sum of the individual plans [3], but consists of an interplay of actions inspired and affected by a joint intention.

Several models have been proposed to explain how joint intention relates to individual intention. Searle [9] argues that collective intentions ("*We*-intentions") are not reducible to individual intentions of the agents involved ("*I*-intentions"), and that the individual acts exist solely in their role as part of the common goal. Bratman's analysis of Shared Cooperative Activity (SCA) [2] introduces the idea of meshing singular sub-plans into a joint activity. We generalize this concept to the idea of dynamically meshing sub-plans.

Bratman also defines certain prerequisites for an activity to be considered shared and cooperative; he stresses the importance of *mutual responsiveness, commitment to the joint activity* and *commitment to mutual support*. Cohen and his collaborators [8] support these guidelines and provide the notion of *joint stepwise execution*. Their theory also predicts that an efficient and robust collaboration scheme in a changing environment commands an open channel of *communication*. Sharing information through communication acts is critical given that each teammate often has only partial knowledge relevant to solving the problem, different capabilities, and possibly diverging beliefs about the state of the task.

Our work integrates these ideas to model and perform collaborative tasks.

MODELLING COLLABORATIVE TASKS

Humans are biased to use an intention-based psychology to interpret an agent's actions [10]. Moreover, it has repeatedly been shown that we interpret intentions and actions based on goals, not specific activities or motion trajectories (e.g. [11]). A goal-centric view is particularly crucial in a collaborative task setting, in which goals provide a common ground for communication and interaction.

All of this argues that goals and a commitment to their successful completion must be central to our intentional representation of tasks, especially if those should be performed in collaboration with others.

Intention and Task Representation

We represent tasks and their constituent actions in terms of *action tuples* [12] with the additional notion of goals. These goals play a central role both in the *precondition* that triggers the execution of a given action tuple, and in the *until-condition* that signals when the action tuple has successfully completed.

Our task representation currently distinguishes between two types of goals: (a) *state-change* goals that represent a change in the world, and (b) *just-do-it* goals that need to be executed regardless of their impact on the world. These two types of goals differ in both their evaluation as preconditions and in their evaluation as until-conditions. As part of a precondition, a state-change goal must be evaluated before doing the action to determine if the action is needed. As an until-condition, the robot shows commitment towards a state-change goal by executing the action, repeatedly if necessary, until the robot succeeds in bringing about the new state. This commitment is an important aspect of intentional behavior [2,8]. Conversely, a just-do-it goal will lead to an action regardless of the world state, and will only be performed once.

Tasks are represented as a hierarchical structure of actions and sub-tasks (recursively defined in the same fashion). Since tasks, sub-tasks, and actions are derived from the same *action tuple* data structure, a tree structure is naturally afforded. It should be noted that goals are also associated with the successful completion of an overall task or sub-task, separate from the goals of each of the task's constituents.

Intention and Decision-Making

When executing a task, goals as preconditions and until-conditions of actions or sub-tasks manage the flow of decision-making throughout the task execution process. Additionally, overall task goals are evaluated separately from their constituent action goals. This top-level evaluation approach is not only more efficient than having to poll each of the constituent action goals, but is also conceptually in line with a goal-oriented hierarchical architecture. For example, consider a task with two actions. The first action makes some change in the world (and has a state-change goal), and the second action reverses that change (also a state-change goal). The overall task goal has no net state change and becomes a just-do-it goal even though its constituent actions both have state-change goals.

Task manager

The *task manager module* maintains a collection of known task models and their associated names. Given this set of tasks, the robot listens for speech input that indicates a task-related request from the human partner. These can be in the form of: "Leo, do *task x*" or "Leo, let's do *task x*." These requests can also be made in the form of a question: "Leo, can you do *task x*?" In the case of a question, given Leonardo has no speech generating capabilities yet, the robot will answer by either nodding "yes" or shaking its head "no." If the robot does not recognize the name of the requested task, or if the robot does not know how to perform it, he looks puzzled or shrugs his shoulders "I don't know."

The task manager distinguishes between requests for autonomous task completion and invitations to task collaboration, and starts the appropriate execution module. If Leo is asked to do a known task on his own, then the task manager executes it autonomously by expanding the task's actions and sub-tasks onto a focus stack (in a similar way to [13]). The task manager proceeds to work through the actions on the stack popping them as they are done and, upon encountering a sub-task, pushing its constituent actions onto the stack. The robot thus progresses through the task tree until the task's goals are achieved.

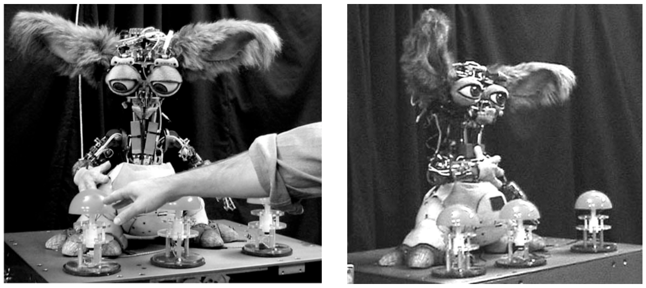

Figure 1: (a) Leonardo participating in a collaborative button-pressing task. (b) Leonardo negotiating his turn for an action he is able to perform.

The major contribution of this work, however, concerns the collaborative scenario: if a collaborative task execution is requested, the task manager starts the collaboration module to jointly execute a common plan.

PERFORMING TASKS WITH HUMANS

When collaborating with a human partner, many new considerations come into play. For instance, within a collaborative setting the task can (and should) be divided between the participants; the partner's actions need to be taken into account when deciding what to do next; mutual support must be provided in cases of one participant's inability to perform an action; and a clear channel of communication must be used to establish mutual beliefs and maintain common ground for intentions and actions.

Our implementation supports these considerations as Leonardo participates in a collaborative discourse while progressing towards achieving the joint goal. To do so, and to make the collaboration a natural human interaction, we have implemented a number of mechanisms that people use when they collaborate. In particular, we have focused on communication acts to support joint activity (utilizing gestures and facial expressions), dynamic meshing of sub-plans and turn taking.

Experimental Setup

In our experimental scenario there are three buttons in front of Leonardo. The buttons can be switched ON and OFF (which lights the button up). Occasionally, a button that is pressed does not light up, and in our tasks this is considered a failed attempt. We use tasks comprised of vision and speech recognition and simple manipulation skills. For instance, Leonardo can learn the names of each of the buttons and is able to point to and press the buttons.

To test our collaborative task execution implementation, we designed a set of tasks involving a number of sequenced steps, such as turning a set of buttons ON and then OFF, turning a button ON as a sub-task of turning all the buttons ON, turning single buttons ON and other tasks. This task set represents simple and complex hierarchies and contains tasks with both state-change and just-do-it goals.

Dynamic Meshing of Sub-plans

Leo's intention system is a joint-intention model that dynamically assigns tasks between the members of the collaboration team. Leo derives his *I*-intentions based on a dynamic meshing of sub-plans according to his own actions and abilities, the actions of the human partner, Leo's understanding of the common goal of the team, and his assessment of the current task state.

Leonardo is able to communicate with the human teammate about the commencement and completion of task steps within a turn-taking interaction. Specifically, the robot is able to recognize changes in the task environment, as well as successes and failures on both Leo's and his teammate's side. Moreover, Leonardo is able to communicate to the human teammate his inability to accomplish a task step crucial to the complete joint action.

Self Assessment and Mutual Support

At every stage of the interaction, either the human should do her part in the task or Leo should do his. Before attempting an element of the task, Leo negotiates who should complete it. To accomplish this, Leo has the ability to evaluate his own capabilities. In the context of the button task, Leonardo can assess whether he can reach each button or not. If he is able to complete the task element (e.g., press a particular button) he will offer to do so. Conversely, whenever he believes that he cannot do the action (e.g., because he cannot reach the button) he will ask the human for help.

Since Leonardo does not have speaking capabilities, he indicates his willingness to perform an action by pointing to himself, and adopting an alert posture and facial expression (Figure 1(b)). Similarly, when detecting an inability to perform an action assigned to him, Leo's expression displays helplessness, as he gestures toward the human in a request for her to perform the intended action. Leo also shifts gaze between the problematic button and his partner to direct her attention to what it is he needs help with.

Communication to Support Joint Activity

While usually conforming to this turn-taking approach, the robot can also keep track of simultaneous actions, in which the human performs an action while Leo is working on another part of the task. If this is the case, Leonardo will take the human's contribution into account and reevaluate the goal state of the current task focus. He then might decide to no longer keep this part of the task on his list of things to do. However, the robot needs to communicate this knowledge to the human to maintain mutual belief about the overall task state.

We have implemented a variety of gestures and other social cues to allow the robot communicate his internal state during collaboration – such as who the robot thinks is doing an action, or whether the robot believes the goal has been met. For instance, when the human partner unexpectedly changes the state of the world, Leo acknowledges this change by glancing briefly towards the area of change before redirecting his gaze to the human. This post-action glance lets the human know that the robot is aware of what she has done, even if it does not advance the task.

If the human's simultaneous action meets a task goal, such as turning the last button ON during the buttons-ON task, Leo will glance at the change and give a small confirming nod to the human. Similarly, Leo uses subtle nods when he thinks he completed a task or sub-task. For instance, Leo will give an acknowledgement nod to the human after completing the buttons-ON sub-task and before starting the buttons-OFF sub-task, in the case of the buttons-ON-then-OFF task.

These cues play a crucial role in establishing and maintaining mutual beliefs between the teammates on the progress of the shared plan.

RESULTS AND FUTURE WORK

In summary, during the trials for the collaborative button task, Leonardo displayed successful meshing of sub-plans based on the dynamic state changes as a result of his successes, failures, and the partner's actions. Leo's gestures and facial expressions provided a natural collaborative environment, informing the human partner of Leo's understanding of the task state and his attempts to take or relinquish his turn. Leo's requests for help displayed his understanding of his own limitations, and his use of gaze and posture served as natural cues for the human to take appropriate action in each case. See Appendix A for a transcript of a typical collaborative interaction.

As future work, we would like to improve the complexity of the task representation as well as the interaction and dialog. Leonardo can understand a few spoken requests of the human, but he does not speak himself. Although his gestures and facial expressions are designed to communicate his internal state, combining this with an ability to speak would give the robot more precision in the information that he can convey. We would also like to implement a richer set of conversational policies to support collaboration. This would be useful for negotiating the meshing of sub-plans during task execution to make this process more flexible and efficient. We continue to make improvements to Leonardo's task representation so that he can represent a larger class of collaborative tasks and more involved constraints between the tasks' action components.

CONCLUSION

Building sociable robots has profound implications for how we will be able to engage robots in the future – far beyond making them appealing, entertaining, or providing an easy interface to their operation. It is a critical competence that will allow robots to assist us as capable partners.

This paper presents an overview of our work to build sociable robots that work cooperatively with people using natural dialog, gesture, and social cues. We have presented how our ideas, informed by joint intention theory, can be applied to building and demonstrating robots that engage in self-assessment and provide mutual support, communicate to support joint activity, perform dynamic meshing of sub-plans, and negotiate task division via turn taking.

ACKNOWLEDGMENTS

The work presented in this paper is a result of the ongoing efforts of the graduate and undergraduate students of the MIT Media Lab Robotic Life Group and our collaborators. Leonardo's speech understanding abilities are developed in collaboration with Alan Schultz and his group at the Naval Research Lab. Stan Winston Studio provided the physical Leonardo robot. Leonardo's architecture is built on top of the *C5M* code base of the Synthetic Character Group at the MIT Media Lab. Geoff Beatty and Ryan Kavanaugh provided the virtual model and animations. This work is funded in part by a DARPA MARS grant and in part by the *Digital Life* and *Things that Think* consortia.

REFERENCES

1. Bluethmann, W., Ambrose, R., Diftler, M., Huber, E., Goza, M., Lovchik, C., and Magruder, D. Robonaut: A Robot Designed to Work with Humans in Space. *Autonomous Robots* 14 (2003), 179-207.
2. Bratman, M. Shared Cooperative Activity. *The Philosophical Review* 101,2 (1992), 327-341.
3. Grosz, B. Collaborative Systems: 1994 AAAI Presidential Address. *AI Magazine* 2,17 (1996), 67-85.
4. Jones, H., and Rock, S. Dialog-based Human-robot Interaction for Space Construction Teams. *Proc. IEEE Aerospace Conference* (2002).
5. Perzanowski, D., Schultz, A., Adams, W., Marsh, E., and Bugajska, M. Building a Multimodal Human-Robot Interface, *IEEE Intelligent Systems*, (2001), 16-20.
6. Kimura, H., Horiuchi, T., and Ikeuchi, K. "Task-Model Based Human Robot Cooperation Using Vision," *Proc IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'99)* (1999), 701-706.
7. Fong, T., Thorpe, C., and Baur, C. Collaboration, Dialogue, and Human-Robot Interaction. *Proc of the International Symposium of Robotics Research*, (2001).
8. Cohen, P., and Levesque, H. Teamwork. *NOÚS* 25, (1991), 487-512.
9. Searle, J. Collective Intentions and Actions. In Cohen, Morgan and Pollack (eds.) *Intentions in communication*, MIT Press (1990), Chapter 19.
10. Dennett, D. Three kinds of intentional psychology. In *The Intentional Stance*, MIT Press (1987), Chapter 3.
11. Baldwin, D., and Baird, M. Discerning intentions in dynamic human action. *Trends in Cognitive Sciences*, 5, 4, (2001), 171-178.
12. Burke, R., Isla, D., Downie, M., Ivanov, Y., and Blumberg, B. Creature Smarts: The art and architecture of a virtual brain. *Proc. of the Computer Game Developers Conference*, (2001).
13. Grosz, B. Collaborative systems, in *AI Magazine*, 17(2): 67-85, 1996.

#	Human	Leonardo	Notes
1	"Leo, let's do task BUTTONS"	Shrugs "I don't know"	Leo does not know this task.
2	"Let's do task BUTTON-ONE"	Looks at the buttons	Leo acknowledges that he understands the task, and visibly establishes mutual belief on the task's initial conditions.
3		Points to himself	He can do the first (and only) part of the task, and suggests doing so.
4	"OK, you go"	Presses button one, looking at it	Looking away from the partner while operating establishes turn taking boundaries.
5		Looks back at his partner	Gaze shift is used to signal end of turn
6		Nods shortly	Communicates the robot's perceived end of task
7	"Leo, let's do task BUTTON-ONE"	Looks at the buttons; points to himself	As in steps 2-3
8	"I'll go "	Looks at his partner	
9	Presses button one	Looks at button one	Acknowledges partner's action, creates mutual belief
10		Nods shortly	Communicates perceived end of task.
11	Moves button one out of Leo's reach		
12	"Let us do task BUTTON-ONE"	Looks at buttons	Leo acknowledges that he understands the task, and visibly establishes mutual belief on the task's initial conditions.
13		Looks at button one, then back at the human partner; extends his arms in "Help me" gesture.	Leo assesses his capabilities and consequently requests support.
14	Presses button one	Looks at button one; looks back at human; nods shortly.	Glance acknowledges partner's action; nod creates mutual belief as to the task's completion.
15	"Let us do task BUTTON-ONE-AND-TWO"	Looks at buttons	Leo acknowledges that he understands the task, and visibly establishes mutual belief on the task's initial conditions
16		Points to himself	He can do the first part of the task, and suggests doing so.
17	"OK, you go"	Presses button one, looking at it	
18	At the same time as 17, presses button two		
19		Looks at button two; looks back at the human; nods shortly	Acknowledges partner's simultaneous action; creates mutual belief as to the task's completion.

Table 1: Sample task collaboration on single-level tasks.

APPENDIX A – TASK COLLABORATION TRANSCRIPT

Table 1 shows a sample transcript describing a typical interaction task collaboration between Leonardo and a human teammate. We chose to display the following simple, non-hierarchical tasks for reasons of transcript brevity: *BUTTON-ONE* – Toggle button one, *BUTTON-ONE-AND-TWO* – Turn buttons one and two ON.

While these do not illustrate the Leonardo's full range of goal-oriented task representation capabilities, they offer a sense of the joint intention and communicative skills fundamental to the collaborative discourse stressed in this paper.

Robotic Pets in the Lives of Preschool Children

Peter H. Kahn, Jr.[1], Batya Friedman[2], Deanne R. Perez-Granados[3], and Nathan G. Freier[2]

[1]Department of Psychology
Box 351525
University of Washington
Seattle, WA 98195-1525
pkahn@u.washington.edu

[2]The Information School
Box 352840
University of Washington
Seattle, WA 98195-2840
{batya, nfreier}@u.washington.edu

[3]School of Education
485 Lasuen Mall
Stanford University
Stanford, CA 94305-3096
granados@stanford.edu

ABSTRACT

This study examined preschool children's reasoning about and behavioral interactions with one of the most advanced robotic pets currently on the retail market, Sony's robotic dog AIBO. Eighty children, equally divided between two age groups, 34-50 months and 58-74 months, participated in individual sessions that included play with and an interview about two artifacts: AIBO and a stuffed dog. Results showed similarities in children's reasoning about the two artifacts, but differences in their behavioral interactions. Discussion focuses on how robotic pets, as representative of an emerging technological genre in HCI, may be (a) blurring foundational ontological categories, and (b) impacting children's social and moral development. More broadly, results inform on our understanding of the human-robotic relationship.

Keywords

AIBO, children, companionship, ethics, human-robotic relationship, human values, moral development, robotic pets, social responses to technology, user conceptions, Value Sensitive Design, virtual pets.

INTRODUCTION

Animals have long been an important part of children's lives, offering comfort and companionship, and promoting the development of moral reciprocity and responsibility [5]. Yet in recent years there has been a movement to create technological substitutes for pets, such as the Tamagotchi, i-Cybie, Tekno, and Poo-Chi. In turn, researchers have begun to ask if technological pet counterparts, now or in the future, can provide children with similar developmental outcomes [6, 9].

Short paper submitted to CHI 2004.

January 10, 2004.

In this study, we investigated preschool children's reasoning about and behavioral interactions with one of the most sophisticated deployed personal robots on the market – Sony's robotic dog AIBO. This artifact, AIBO, represents the integration of two long-standing areas of research within the CHI community. The first area involves computer persona that exist on the desktop computer or through voice interfaces, including virtual embodied agents and social responses to computer technology [7]. The second area involves computational artifacts (without a persona) that link people to a physical world, including augmented reality, tangible computing, and telepresence [4]. By bringing both areas of research together – through the use of computation to embed interactive persona into physical artifacts – personal robots represent a new genre for human-computer interaction.

Building on Friedman, Kahn, & Hagman [3], and principles of Value Sensitive Design [1, 2] we sought data that would inform on how robotic pets (as representative of this emerging technological genre in HCI) may be (a) blurring foundational ontological categories, and (b) impacting children's social and moral development.

METHODS

Participants

Eighty children participated in this study, equally divided between two age groups, 34-50 months and 58-74 months. There were equal numbers of males and females in each age group.

Artifacts

Two main artifacts were used in this study: a robotic dog and a stuffed dog. The robotic dog was Sony's version 210 AIBO, at the time of data collection (2001-2002) the most advanced robotic animal on the retail market. The stuffed dog was roughly the same size as the robotic dog and made of a soft-plush fabric. Both the robotic and stuffed dog were black-hued in color.

Procedures and Measures

Each of the 80 children participated in an individual session lasting approximately 45 minutes. One part of the session involved an interactive period with AIBO, and another part an interactive period with the stuffed dog (which we called SHANTI). The presentation order of the two artifacts was counterbalanced.

With each artifact (AIBO or the stuffed dog), the child first engaged in a short (2-3 minute) unstructured introductory "play" period. Then the child was allowed to continue to play with the artifact while engaging in a semi-structured interview. In order to limit the total number of questions asked of any one child – to fit within the 45-minute session – children by sex and age were randomly divided into two groups. One group was asked 10 questions that pertained to each artifact's *biological properties* (e.g., "This is a dog biscuit. Do you think AIBO will eat this?"), and *mental states*, including intentionality (e.g., "This is a doggie toy. I'm going to put it here. Do you think AIBO will try to get the toy?") and emotion (e.g., "Can AIBO feel happy?"). The other group was asked 12 questions that pertained to each artifact's *social rapport*, including reciprocal friendship relations (e.g., "Can AIBO be your friend?" "Can you be a friend to AIBO?" "If you were sad, would you want to spend time with AIBO?"), and *moral standing* (e.g., "Do you think it's OK that I hit AIBO?" "Is it OK to leave AIBO alone for a week?"). Then every child was asked 5 questions about each artifact's potential *animacy* (e.g., "Is AIBO alive or not alive?" "Can AIBO die?"). The interviewer asked the questions in as relaxed a format as possible, with the child often engaged in playing with AIBO or the stuffed dog. We believed this method increased the ecological validity of the interviews.

Children's behaviors with both artifacts were video-recorded continuously during the interactive sessions, and then reviewed for coding.

Coding and Reliability

Building on coding categories from [3], a detailed reasoning and behavioral coding manual [technical report citation omitted for blind review, 2003] was developed from half of the data and then applied to the entire data set. 17.5% of the data was recoded by a second individual trained in the use of the coding manual. Intercoder reliability was assessed through testing Cohen's kappa at the .05 significance level. All tests were statistically significant. For evaluations, $k = .85$ ($Z = 18.02$), and for behavioral responses, $k = .76$ ($Z = 45.21$).

RESULTS

Averaging evaluations within question type, about a quarter of the children accorded animacy to both artifacts (AIBO 25%, stuffed dog 20%), about half the children accorded biological properties (AIBO 46%, stuffed dog 48%), and about two-thirds of the children accorded mental states (AIBO 66%, stuffed dog 64%), social rapport (AIBO 76%, stuffed dog, 82%), and moral standing (AIBO 63%, stuffed dog 67%).

Children's behavioral interactions with the artifacts were coded with the 6 overarching categories and 22 subcategories listed in Table 1. Intercoder reliability was established at the level of the 22 subcategories. The six overarching categories are exploration, apprehension, affection, mistreatment, endowing animation, and an attempt at reciprocity. Table 1 also provides a definition, example, and representative still image from the video data of each category; the video figure provides clips of each behavioral category. In total, 2,360 behavioral interactions were coded, 1,357 with AIBO (58%) and 1,003 with the stuffed dog (43%).

Statistical results (using the Wilcoxon signed-rank test) showed that children engaged in a comparable amount of affection with both AIBO (294 occurrences) and the stuffed dog (310 occurrences). But otherwise, across the five other overarching categories, children differed in their behavioral interactions with AIBO and the stuffed dog. Specifically, children more often engaged in exploratory behavior with AIBO (221 occurrences) than with the stuffed dog (150 occurrences) ($p = .013$); more often engaged in apprehensive behavior with AIBO (143 occurrences) than with the stuffed dog (1 occurrence) ($p = .000$); more often engaged in attempts at reciprocity with AIBO (683 occurrences) than with the stuffed dog (180) ($p = .000$); less often engaged in mistreatment of AIBO (39 occurrences) than with the stuffed dog (184 occurrences) ($p = .000$); and less often engaged in endowing animation with AIBO (20 occurrences) than with the stuffed dog (207 occurrences) ($p = .000$).

DISCUSSION

There was no difference in children's evaluations that pertained to AIBO and to the stuffed dog. One interpretation of these results is that the children engaged in imaginary play with AIBO in the same way and to the same degree that they engaged in imaginary play with the stuffed dog. Yet this interpretation is called into question by our behavioral results. Namely, children engaged more often in exploratory behavior, apprehensive behavior, and attempts at reciprocity with AIBO, and more often mistreated the stuffed dog and endowed it with animation. These behavioral results show that the children substantially distinguished between the two artifacts. The behavioral results also map well on to how one might expect children to respond to AIBO if they were treating AIBO *as if* it were a live dog. For example, children flinching away from AIBO immediately after AIBO initiated an action (e.g., standing, walking, or approaching the child) is evidence that the children believed that AIBO could be a threat. Our findings are of a piece with research in the field of human-computer interaction [7] which shows that with minimal social cues computers can pull for social responses.

Table 1. Coding Categories for Behavioral Interactions

Behavioral Category	Definition and Example	Still Image from Video
1. Exploration 1.1 Anatomy Check 1.2 Touch/Move Limbs 1.3 Demonstrate w/ Artifact 1.4 Feed	Reference to the child's visual or tactile exploration, manipulation, inspection, pointing, and feeding of the artifact. *E.g., child explains to the interviewer that AIBO is a boy while inspecting the hindquarters of AIBO.*	
2. Apprehension 2.1 Startle 2.2 Wariness	Reference to the child exhibiting a startle response, wariness, or other intentional movement away from the artifact. *E.g., child touches AIBO's head, AIBO begins moving, and child reacts with startle.*	
3. Affection 3.1 Non-exploratory Touch 3.2 Pet 3.3 Scratch 3.4 Kiss 3.5 Embrace 3.6 Verbal	Reference to the child engaging in petting, scratching, kissing, carrying, embracing, and one-way verbal greetings to the artifact. *E.g., child squeezes the stuffed dog in a big hug.*	
4. Mistreatment 4.1 Rough Handling 4.2 Thumping 4.3 Throwing	Reference to the child's behavior showing disregard for the artifact, including rough handling (e.g., hitting, squishing) and throwing. *E.g., child swings the stuffed dog overhead and then thumps it to the floor.*	
5. Endow Animation 5.1 Vocalize 5.2 Movement 5.3 Object-oriented Play 5.4 Feed	Reference to the child enlivening the artifact in order to perform a behavior or action with it, including making sounds and moving the artifact around. *E.g., child throws the doggie toy and says "Go get it!" Then child picks up the stuffed dog and begins to hop it toward the toy.*	
6. Attempt at Reciprocity 6.1 Motion 6.2 Verbal 6.3 Offering	Reference to the child's behavior not only responding to the artifact, but expecting the artifact to respond in kind based on the child's motioning behavior, verbal directive, or offering. *E.g., AIBO is searching for a ball. Child observes AIBO's behavior and puts the ball in front of AIBO and says, "Kick it!"*	

Our results also support the proposition that a new technological genre may be emerging that challenges traditional ontological categories (e.g., between animate and inanimate). This genre comprises artifacts that are *autonomous* (insofar as they initiate action), *adaptive* (act in response to their physical and social environment), *personified* (convey an animal or human persona), and *embodied* (the computation is embedded in the artifacts rather than just in desktop computers or peripherals). If we are correct, then it may be that the English language is not yet well equipped to characterize or talk about this genre. As an analogy, we do not normally present people with an orange object and ask "is this object red or yellow?" It is something of both, and we call it orange. Similarly, it may not be the best approach to keep asking people if this emerging technological genre is, for example, "alive" or "not alive" if from the person's experience of the subject-object interaction, the object is alive in some respects and not alive in other respects, and is experienced not simply as a combination of such qualities (in the way one can inspect a tossed salad and analytically distinguish, for example, between the green leaf lettuce and the red leaf lettuce) but as a novel entity. Thus the human-computer interaction question for the future may not be, "Do young children treat such new technologies as either X or Y?" (e.g., animate or inanimate, having agency or not, or being a social other or not) because the answer may not be one or the other. Rather, the question, or at least an initial question – and one that this current study has purchase on – is "What are some more fine-grained characteristics of children's reasoning and behavior in relation to this new technological genre?"

In the moral developmental literature [8], reciprocity is central to moral development, setting into motion concerns for the wellbeing of others and the construction of equality, fairness, and justice. Thus it was surprising to find that almost half of the children's behavioral interactions with AIBO involved an attempt at reciprocity (668 occurrences). While children may form certain types of moral relationships with robotic pets, it is our supposition that the nature of these relationships will be impoverished in several ways. First, what does it mean to morally care about an entity that (as the majority of the children recognized) is not alive? In this sense, a person can "care" very deeply about a car they have owned for decades, and cry when it is finally towed to the junkyard; but that would seem to us a derivative form of caring, supported only by the person's projection of animacy and personality onto the artifact, concepts which may first have to be developed in the company of sentient others. Second, to the extent interactions with the robot partially replace children's interactions with sentient others, and as long as the robot only partially replicates the entire repertoire of its sentient counterpart, then such interactions may impede young children's social and moral development.

Future studies could move in a number of important directions. One direction would be to conduct research that compared children's reasoning of and behavior with AIBO in comparison to a live dog (rather than a stuffed dog, as in the present study). Another direction would be to investigate differences in children's relationships with robotic humanoids compared to robotic animals. It is our intuition that because people do not expect full social responsiveness from animals, that children (and adults) will find human-animal robotic relationships more satisfying than human-humanoid robotic relationships, especially until the robotic technology is able to mimic more realistically human behavior.

ACKNOWLEDGMENTS

[Funding acknowledgments omitted for purposes of a blind review.]

REFERENCES

1. Friedman, B., and Kahn, P. H., Jr. Human agency and responsible computing: Implications for computer system design. *Journal of Systems Software*, *17* (1992), 7-14.
2. Friedman, B., Kahn, P. H., Jr., and Borning, A. *Value sensitive design: Theory and methods* (UW CSE Technical Report 02-12-01). Seattle: University of Washington, Department of Computer Science and Engineering (2002).
3. Friedman, B., Kahn, P. H., Jr., and Hagman, J. Hardware companions?: What online AIBO discussion forums reveal about the human-robotic relationship. *Proc. of CHI 2003*, ACM Press (2003), 273-280.
4. Ishii, H. and Ullmer, B. Tangible bits: Towards seamless interfaces between people, bits and atoms. *Proc. of CHI 97,* ACM Press (1997), 234-241.
5. Kahn, P. H., Jr. *The Human Relationship with Nature: Development and Culture*. MIT Press, Cambridge MA, 1999.
6. Perez-Granados, D. R. *Computerized toys as social partners? Exploring the role of technology in young children's play and social interactions*. Paper presented at the Annual Meeting of the AERA, New Orleans, LA (2002, April).
7. Reeves, B., and Nass, C. *The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places*. Cambridge University Press and CSLI Publications, New York NY and Stanford CA, 1996.
8. Turiel, E. Moral development, in N. Eisenberg (ed.), *Social, emotional, and personality development*. Vol. 3 of W. Damon (Ed.) *Handbook of Child Psychology*. 5th ed. Wiley, New York NY, 1998, 863-932.
9. Turkle, S. *Life on the Screen: Identity in the Age of the Internet*. Simon & Schuster, New York NY, 1997.

Towards A Design-Centered Framework for Social Human-Robot Interactions

Jodi Forlizzi
HCII and School of Design
Carnegie Mellon University
Pittsburgh, PA 15221 USA
forlizzi@cs.cmu.edu

ABSTRACT

Robotic technology has achieved a certain amount of maturity, and is being integrated more and more into consumer products. Although several intelligent products have begun to integrate into our daily lives, the social impact of these products is largely unknown. How should these products be designed to offer ready access to the technological innovations they offer? What interactions and behavioral cues are most appropriate? Finally, is it possible for this class of intelligent products to play a social role in our society and culture?

Our group, the Project on People and Robots, has been investigating how robotic technology might be used to support people in the near future. Although our primary audience is elders and caregivers, our investigations have sought to address some of these issues, specifically looking at what makes a product social, how social products and social robots might be designed, and how lifelike behavior and emotional expression might play a role in social products. In this paper, we present a ideas for a design-centered framework for organizing findings related to these issues.

Categories and Subject Descriptors

Design, Design tools and techniques, User interfaces

General Terms

Design, Human Factors

Keywords

Interaction design, robots, social products

1. INTRODUCTION

Robotic technology is now mature enough to be integrated more and more into consumer products. Intelligent products are now robust enough to be deployed in industrial, institutional, and domestic settings. They have the potential to be greatly beneficial to humankind. However, how these products should behave and interact with humans — *act socially* — remains largely unclear. For example, an intelligent product might be perceived as having intentional behavior that is usually only ascribed to living things.

2. THE PROJECT ON PEOPLE AND ROBOTS

Our research group, the Project on People and Robots, has begun to understand how robotic products might act socially in both domestic and professional environments [10]. Although our initial audience for this research is the aging population and their caregivers, we have learned a great deal about how robots might assist humans with tasks that are boring, repetitive, unsafe, or in the need of better communication and information management. We have conducted both qualitative and quantitative research on how people make attributions to robots, animated characters, and pets, and we have conducted detailed qualitative studies on how people form relationships to products in order to glean clues for future robotic product design.

Pearl is one of the first robotic product prototypes that we have worked with in our research group (Figure 1). Pearl is a mobile robot that offers interaction through a touchscreen and an expressive face [9]. Motivated by the need for a new robot head, our group did an extensive design investigation of what particular features of a face makes it appear most humanlike. Additionally, we derived some initial design guidelines for the design of a humanlike humanoid robot head [3].

Pearl was also used in a series of experiments to learn what robot "personality" should be used for a particular task, in order to best match the robot's behavior to a user's expectations and mental models [7, 8]. A number of these experiments have shown that the appearance and behavior of the robot should match the task that it is used for in a logical way.

Figure 1. Pearl, a robotic product prototype with a humanlike appearance.

A second robot that we have studied in our research group is Valerie, a robot receptionist that is stationed in the entryway of a computer science building on our campus [11]. Valerie's face is portrayed on a movable screen that turns to orient and interact with visitors. Valerie provides factual information (weather and local directions) as well as emotional communication (stories about her life, revealed through interactions with visitors and phone conversations with her "friends" and "family"). This social robot has provided a testing ground for how natural and engaging interactions might be offered through an autonomous system, especially for those who have never interacted with a robot. However, she has also inspired criticism about her stereotypical character, and fears that robots like her will replace human workers in all kinds of contexts.

Figure 2. Valerie, a social roboceptionist that provides factual and emotional information to visitors on Carnegie Mellon's campus.

Our group has also explored the potential for robots to look and act more like everyday appliances or domestic products. We hypothesized that we could learn clues about how future robotic products could be designed by understanding people's relationships to products today. We conducted an extensive ethnographic study on how elders relate to products. The findings from this work indicated that robotic products should be conceived of as adaptive components of a system, and at least for elders, should mimic familiar products that are used in the home everyday [5].

We are using this data to design several product concepts. The first is the Hug, a telepresent product that facilitates rich social and emotional communication between elders and family members who live at a distance [6]. This concept (Figure 3) provided answers to initial questions about how a robotic product might be used to create expressive and emotional reciprocal interactions with a group of users.

Figure 3. The Hug, a telepresent robotic product that facilitates communication between family members at a distance.

These investigations have taught us that in order to fill a social role in human life, a robotic product must match the task it is created for in a logical way, make an expressive and emotional connection to a human, and be perceived of as an appropriate assistant or social partner. More work is needed to better understand the role of social robotic products and the process through which we create expressive and emotional relationships with certain product forms and behaviors.

3. MODELS OF HUMAN-ROBOT INTERACTION

A number of theoretical approaches have been developed to characterize the area of social human-robot interactions. Many of these approaches are driven by technological advances that enable more natural and lifelike human-robot interactions. Rodney Brooks has created an organizational framework that organizes social robots relative to their sensing and modeling capabilities. His framework includes perception, attention, motivation, behavior, and expression as organizing factors [2]. Cynthia Breazeal has created a taxonomy of human-robot interactions that includes tools, cyborgs, tinkers, avatars, and social robots. Social robots are classified as evocative, interface-driven, receptive, or sociable [1]. We find it critical to take into account the notion of designed form — the shape, materials, and expressive capabilities that the robot offers to a user. Of particular interest are designed forms that mimic human or other biological life [4].

4. A DESIGN-CENTERED FRAMEWORK FOR SOCIAL HUMAN-ROBOT INTERACTION

During the workshop we will propose a framework from a design-centered perspective. It will focus on the form that robotic products take and the interactions between robotic products and individuals. Two possible scales for classification include form (defined as the perceived aspects of the robotic product, ranging from visible to invisible), and interaction (defined as the behavioral aspects of the robotic product and defining the relationship between human and product, ranging from human control to fully reciprocal social interactions).

During the workshop we hope to classify a fairly large group of both commercial products and research prototypes, in order to understand the role of design in creating products that will function as companions and personal assistants.

5. CONCLUSION

This paper has presented a brief overview of findings from our research group related to social robots. A possible framework for understanding the design and development of social robotic assistants has also been proposed. During the workshop, we hope to classify robotic products ranging from aware homes to humanoid assistants. We also hope to stress the importance that design plays in the conception and development of future robotic products.

6. ACKNOWLEDGMENTS

This work is made possible through the many contributions of the members of the Project on People and Robots: Sara Kiesler, Carl DiSalvo, Francine Gemperle, Venus Lee, Aaron Powers, Rachel Gockley, and others. This work is partially funded by NSF IIS-0121426.

7. REFERENCES

[1] Breazeal, C. (2002). *Designing Sociable Robots*. Boston, MA: MIT Press.

[2] Brooks, R. (2002). Flesh and Machines: How Robots Will Change Us. New York: Pantheon Books.

[3] DiSalvo, C., Gemperle, F., Forlizzi, J. and Kiesler, S. (2002). “All Robots Are Not Created Equal: Design and the Perception of Humanness in Robot Heads.” *DIS2002 Conference Proceedings*. New York: ACM Press.

[4] DiSalvo, C. and Gemperle, F. (2003). “From Seduction to Fulfillment: The Use of Anthropormophic Form in Design.” *Proceedings of the International Conference on Designing Pleasurable Products and Interfaces ’03*. New York: ACM Press.

[5] Forlizzi, J., DiSalvo, C., and Gemperle, F. (2004). “Assistive Robotics and an Ecology of Elders Living Independently in Their Homes.” *Journal of HCI Special Issue on Human-Robot Interaction*, V19 N1/2.

[6] Gemperle, F., DiSalvo, C., Forlizzi, J., and Yonkers, W. (2003). “The Hug: A new form for communication.” *Proceedings of the DUX 2003 (Designing the User Experience) Conference.* New York: ACM.

[7] Goetz, J. and Kiesler, S. (2002). “Cooperation with a Robotic Assistant,” in *CHI 02 Extended Abstracts*. New York: ACM Press.

[8] Kiesler, S. and Goetz, J. (2002). “Mental Models of Robotic Assistants.” in in *CHI 02 Extended Abstracts*. New York: ACM Press.

[9] Nursebot, Carnegie Mellon University. http://www-2.cs.cmu.edu/~nursebot/web/design.html

[10] The Project on People and Robots, Carnegie Mellon University. http://research.nursebotcs.cmu.edu

[11] Valerie, the Roboceptionist, Carnegie Mellon University. http://www.roboceptionist.com

From Fiction to Science – A cultural reflection of social robots

Christoph Bartneck
Eindhoven University of Technology
Den Dolech 2, 5600 MB Eindhoven
The Netherlands
Christoph@bartneck.de

ABSTRACT
This paper reflects on the culture of human-robot interaction. A review of common concepts in movies and literature is presented and their relation to scientific work is discussed. Two new research directions on the synthesis of behavior models and the perception of social robots are presented.

Author Keywords
Social Robots, Society, Interaction

ACM Classification Keywords
Interaction design, robots, social products

INTRODUCTION
The role that robots might play in our society and what their abilities will be, has been an important topic in science fiction literature. Issac Asimov (1991) defined the Three Laws of Robotics which set a framework of human-robot interaction:

1. A robot may not injure a human being or, through inaction, allow a human being to come to harm.
2. A robot must obey orders given it by human beings except where such orders would conflict with the First Law.
3. A robot must protect its own existence as long as such protection does not conflict with the First or Second Law.

The arrivals of the first consumer robots, such as AIBO, confront us with the need to take these frameworks and ideas out of fiction and into reality. To start with, I would like to review common concepts in the science fiction domain.

ROBOTS WILL TAKE OVER THE WORLD
The movie "Animatrix" (see Figure 1) describes the second renaissance as a period in which humanity created millions of robots to server their needs, the process of the robot's emancipation, the war of humanity against robots with the final stage of humanities' enslavement (Wachowski & Wachowski, 2003).

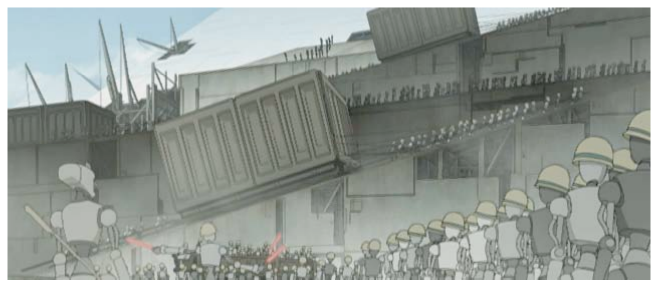

Figure 1: Scene from Animatrix

This is a typical scenario that can also be found in "Terminator" (see Figure 2) and other Hollywood movies.

Figure 2: Terminator

Interestingly, this vision of the future is not so popular in Japan. A possible reason for this could be that the in Shinto Buddhism god is in everything, including humans, animals, plants and machines and rocks. The Christian world makes a strict division between creatures that have a soul and objects that do not. According to Shinto, robots are not that different from humans. In the popular Japanese Manga movies good fights evil just like in the western world, but the role of the good and the evil is not mapped directly to humans as being the good against robots being the evil. In these movies the good and the evil are distributed. You might have a good robot that fights an evil human villain or a good robot fighting bad robots.

ROBOTS WANT TO BE LIKE HUMANS

In the TV series "Star Trek – Next Generation" the android Data (see Figure 3) is constantly trying to become more human (Paramount Pictures, 2002).

Figure 3: Data

At some point he even acquires an emotion chip that enable him to experience feelings. I cannot see a good reason for Data's behavior other than that the writers of the series wanted to flatter humanity. It is perfectly acceptable that he would want to be able to communicate effectively with the crew of the Enterprise and thus it makes sense that he studies their behavior. But why would Data want to become human? Why would he want to be something that he cannot be?

Figure 4: AI

The same assumption is present in Steven Spielberg's "AI" (see Figure 4) in which the main robot character wants to become "a real boy" (Watson & Aldiss, 2001) and in "Bicentennial Man" (Asimov, 1999).

Data's brother Lore does not have this need, but again the writers slip back to the "robots will take over the world" scenario. In the double episode "Descent" Lore teams up with the Borg to take over the universe's leadership from the inferior biological life forms. It appears difficult for humans to accept that other intelligent beings would not want to be like them. In particular if humans originally created these beings. The situation appears similar to the one that parents face when their children are very different from them and choose a different life style than themselves. - Those ungrateful (robot) brats do whatever they want! - Maybe we ourselves have to mature and let robots be what they really are or want to be. This will of course still have to be within the boundaries of the laws. Robots will have to respect them as any other member of the society.

In the animated TV series "Futurama" the robot Bender (see Figure 5) demonstrates what robot emancipation can be like (Groening, 1999).

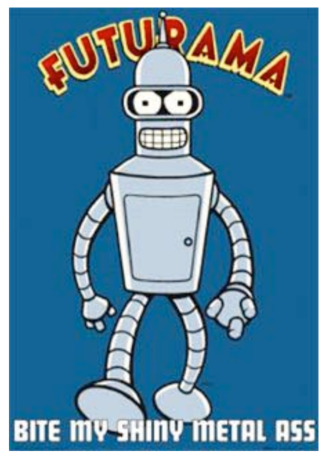

Figure 5: Bender

Bender and all the other robots live along humans, but are happy with what they are. Confronted with the choice to be a human Bender is most likely to answer: "bite my shiny metal ass!", which also illustrates Bender's general attitude.

PEOPLE WANT ROBOTS TO BE LIKE HUMANS

In the movie Star Wars: Episode V - The Empire Strikes Back (1980) by George Lucas the group of Jedi Nights and robots visit the City in the Sky. During their visit one of the main characters, the robot R2D2 (see Figure 6), separates from the group and enters a secret room.

To his own dismay, he entered a robot torture chamber. A similar model to R2D2 is turned up side down and glowing irons are pressed against his feed. The robot unsuccessfully wiggles to avoid the irons and upon contact beeps out loud. R2D2 is shocked and afraid and expresses his distress with a series of beeps. This whole scene makes no sense whatsoever. Robots cannot feel pain, do not have emotions and torture is a very ineffective way to extract information from them. Still, the viewers feel sorry for the tortured robot and worries about R2D2.

Figure 6: R2D2

This movie scene demonstrates how easily people attribute human emotions into machines. One can conclude that robots only need to mimic human behavior as closely as possible to be perceived as a social being. Given this assumption the creators of robots have to formalize existing psychological models of human behavior, such as the OCC model (Ortony, Clore, & Collins, 1988) for an emotion system, and implement them into the robot. Since these models were usually only created to explain human behavior and not to synthesis it, it takes a considerable effort to convert them into a working software model (Bartneck, 2002).

RESEARCH DIRECTIONS

But this is exactly one of the most interesting areas of robot research: synthesizing human behavior to validate psychological models of human behavior. Synthesis is an essential activity in scientific conduct but psychologists were limited to the analyses of human behavior since they were not able to create an artificial being that they could use to synthesize human behavior. With the maturity of robotic technology, including a significant increase of computing power, it became possible to create intelligent social beings. We now can create robots that act autonomously in the real world and that interact with humans (Fong, Nourbakhsh, & Dautenhahn, 2003). By cycles of synthesis and analysis we will be able to create robots that act naturally with humans and at the same time gain a better understanding of humans themselves. The arrival of studies into the ethical (Dennet, 1997) and legal (Lehman-Wilzig, 1981) aspects of human-robot interaction shows that the integration of robots in our society in immanent.

All of this is still under the assumption that humans want robots to act human like. A very interesting area for robot research is the questions when do people not treat robots like humans. When does the perception of a social robot break down and the robot is treated like a machine. To find a clear answer to this question it appears necessary to take a closer look into more extreme situations, possibly negative ones. Here is a list of possibly research questions:

- Do humans torture robots differently than other humans?
- How do humans treat robots that lie and cheat?
- Do humans hold robots responsible for their failures?

By understanding human attributions in extreme conditions, we might discover effects that we could observe to a lesser degree in common human-robot scenarios. They might help us better understand the human-robot interaction.

CONCLUSIONS

It appears necessary to let go some concepts about human robot interaction that have been promoted by movies and literature. They utilize people's fear of the unknown to build engaging stories. These concepts are therefore so strong in the minds of the people that they might have influence on the results of our empirical studies, since the participants in our experiments are exposed to them too.

Two main research areas appear to be of interest. First, the synthesis of human behavior models based on existing models and their further development through iterative cycles. Second, the study of the perception of robots as social actors. When do people perceive them as social beings and when like machines and what is the influence of this perception on the interaction between them?

REFERENCES

1. Asimov, I. (1991). I, robot. New York: Bantam Books.
2. Asimov, I. (Writer) (1999). Bicentennial Man.
3. Bartneck, C. (2002). *Integrating the OCC Model of Emotions in Embodied Characters*. Proceedings of the Workshop on Workshop on Virtual Conversational Characters: Applications, Methods, and Research Challenges, Melbourne.
4. Dennet, D., C. (1997). When HAL Kills, Who's to Blame – Computer Ethics. In D. G. Storck (Ed.), *Hal's Legacy* (pp. 351-365). Cambridge: MIT Press.
5. Fong, T., Nourbakhsh, I., & Dautenhahn, K. (2003). *A survery of socially interactive robots*. Robotics and Autonomous Systems, 42, 143-166.
6. Groening, M. (Writer) (1999). Futurama.
7. Lehman-Wilzig, S. N. (1981). *Frankenstein Unbound: Toward a Legal Definition of Artificial Intelligence*. FUTURES: The Journal of Forecasting and Planning, 13(6), 442-457.

8. Ortony, A., Clore, G., & Collins, A. (1988). The *Cognitive Structure of Emotions*. Cambridge: Cambridge University Press.
9. Paramount Pictures. (2002). StarTrek. from http://www.startrek.com/
10. Wachowski, A., & Wachowski, L. (Writer) (2003). Animatrix.
11. Watson, I., & Aldiss, B. (Writer) (2001). AI.