Catz, Dogz & Robotz? Human interaction with domestic robotic devices

Shaun Lawson and Thomas Chesney

***Abstract*—This special issue of the Journal of Physical Agents is devoted to human interaction with domestic robots. The form, features and future, of domestic robotic devices, from entertainment-based agents through to robotic cleaners, companions, assistants and helpers, are considered and discussed.**

***Index Terms*—human-robot interaction (HRI), physical agents, domestic agents and applications.**

I. Introduction

IN the future, humans will be expected to engage in interactions with embodied robotic systems that are truly ubiquitous [1], [2]. Many such interactions are likely to be situated in public places a person is greeted by a service robot when they enter a museum or shopping centre, a traveler checks into a hotel and the luggage is taken by a robotic busboy, a survivor of a collapsed building encounters a rescue robot as it penetrates the debris around them. Many other interactions however will take place in domestic environments. Indeed the major commercial successes to date in consumer robotics have all been devices, such as the iRobot Roomba and WoWees RoboSapien, intentionally built for the domestic home.

This special issue of the Journal of Physical Agents is devoted to human interaction with domestic robots. It was conceived following the successful staging of two international symposia entitled "The Reign of Catz and Dogz: the role of virtual creatures in a computerised society", the first of which was held at the Society for the Study of Artificial Intelligence and Simulation of Behaviour (AISB) annual conference in Newcastle (UK) in April 2007, and the second at AISB 2008 in Aberdeen (UK) in April 2008. The "Catz & Dogz" symposia aimed to explore aspects of interaction with anthropomorphised and domestic technology such as Aibo, Pleo, Paro and Nabaztag, as well as software such as Catz, Dogz, (fluff)Friends, Neopets and Nintendogs, as well as the numerous non-commercial devices and systems that have been developed in many research labs. The world-wide popularity of many examples of such artifacts provides evidence of the widespread appeal of interacting with artificial representations of creatures, however the academic investigation of such interactions remains scarce. The two symposia to date, both chaired by the guest editors of this issue, have attracted not only contributions centred upon software agents but also many on interaction with embodied interactive and social robots.

The notion of social robots - and human interaction with them - has, in recent years, spawned a new, and inter-disciplinary, academic field: that of Human Robot Interaction (HRI). HRI however is still very much in its infancy in a worldwide context. Japan, with its aging population and low birth rate, has long fostered research in areas which could, in the long term, provide autonomous, but socially acceptable, assistive-care for older people. Additionally, it is often speculated (for example [3]) that cultural differences in the Far East result in a much more comfortable acceptance of the notion of living machines when compared to the skepticism and irony that often greets notions of commercial social robots in, for instance, Europe and the US (for example [4]). Tellingly, therefore, the first eight annual IEEE International Workshops on Robot and Human Interactive Communication (RO-MAN) were held in Japan (starting in 1992). However, since 2000, RO-MAN workshops have also been held in Europe and the US, whilst the First Annual ACM Conference on Human-Robot Interaction (HRI) was held in Salt Lake City, USA, in 2006 followed by events in Washington (USA) in 2007 and in Amsterdam (the Netherlands) in 2008 (both co-sponsored by ACM/IEEE). Furthermore, both of the IEEE Robotics and Automation Society (RAS) flagship meetings in 2008, the International Conference on Robotics and Automation (ICRA) and the International Conference on Intelligent RObots and Systems (IROS) both incorporated themes of humans co-existing with robots in their main calls thus demonstrating the burgeoning nature, and contemporary importance of HRI.

It was in this context therefore that we solicited papers for a special issue of JoPhA early in 2008 devoted to human interaction with domestic robots. We received a very enthusiastic response to the call and reluctantly had to reject many papers that were of a high quality but fell outside of the theme of the issue. Several outstanding and highly innovative pieces of work from Catz and Dogz 08 such as the Haptic Creature project at UBC [5] - also werent available due to time constraints on publication. In the event we have accepted six papers all broadly situated within the theme of interaction with robots in a domestic setting.

The accepted papers reflect the complexities and, especially, the inter-disciplinary issues, currently facing HRI researchers: some papers included in this issue are focused on systems and engineering design whilst others report on experiments involving user groups; several of the papers in the latter category also incorporate methodologies originally forged in the human computer interaction domain such as Wizard-of-

Shaun Lawson is with University of Lincoln.
E-mail: slawson@lincoln.ac.uk
Thomas Chesney is with Nottingham University Business School.
E-mail: thomas.chesney@nottingham.ac.uk

Oz (WoZ) techniques [6]. The melding of research work in systems design together with user experimentation work is a difficult challenge for HRI researchers - and one that was discussed, along with issues such as WoZ approaches, at length at the NEWHRI workshop at IEEE ICRA 08 [7] attended by many of the discipline's leading researchers.

II. THIS ISSUE CONTENT

The first paper in this special issue by Saldien et al. [8] falls into the category of a systems design contribution - the authors describe their approach to the construction of a novel interactive robot called Probo - a device intended to act as a therapeutic companion for children when in hospital. Probo draws immediate comparison with devices such as the MIT Huggable [9] but also features a highly novel morphology and a constrained target user group which shows great future potential. The second paper in this issue by Looije et al. [10] also addresses the future role of robots as companions and supporting characters for children; however this paper not only has its focus more in the domain of a user study (it uses the well-known Philips iCat [11] as an off-the-shelf interactive robot) but also addresses the difficult issue of embodiment in HRI - comparing a physical device with both a virtual representation of the same thing as well as a text interface.

The issues surrounding robot appearance and behaviour is a highly debated topic with many other authors elsewhere experimentally comparing different scenarios (for example [12], [13], [14], [15]. Individual differences in human subjects psychological impressions of social agents is also a difficult and intensely debated issue [16]. A great deal of HRI research currently involves humanoid robots - however, it is often predicted that mismatches between the appearance and ability of such robots could render interactions with them problematic (for example [17]). Much has also been written about the role of the uncanny valley in HRI and the problems with robots that fall through the cracks of human-like behaviour [18]. The third paper in this special issue by Lohse et al. [19] makes a contribution to this debate by comparing human reactions to a number of different robots - including a humanoid device.

The next paper by Heerink [20] also makes use of the iCat in a user study - this time in an application featuring older people as end-users. The idea that the older population might find great benefit in the future from social robots has great potential though its an area that also raises complex ethical issues [21]. The fifth and penultimate paper in this issue, by Lee et al. [22] returns to the style of a systems design contribution and describes the development of the robot, and its sensor systems, which went on to win the newly formed RoboCup@Home 2007 competition. The final paper in this issue by Jacobsson et al. [23] describes a new kind of robot for HRI all the way through concept, design and finally to evaluation - this paper emphasizes to us that the robots that will exist in everyday environments of the future will take on many different forms and purposes. We are reminded, in this respect, of innovative new commercial devices for the home such as the Nabaztag and the Sony Rolly - and we predict many more novel robotic systems like these will before long move from research lab to the domestic home.

ACKNOWLEDGMENT

We would like to thank the reviewers of the papers submitted for this special issue as well as the chairs respectively of AISB 2007 and 2008 for their support of Catz and Dogz.

REFERENCES

[1] K. Dautenhahn, B. Ogden and T. Quick, "From embodied to socially embedded agents—implications for interaction-aware robots. *Cognitive Systems Research* 3(3):397-428, 2002.
[2] B. Gates, "A robot in every home". *Sceintific American* January 2007.
[3] M. Kusahara, "The art of creating subjective reality: an analysis of Jananese digital pets. In C. Maley and E. Boudreau (eds), *Artificial Life VII Workshop Proceedings*, 2000.
[4] G. Musser, "Robots that suck: have they finally come out witha robot for the rest of us? *Scientific American*, Febuary 2003.
[5] S. Yohanan and K. E. MacLean, "The haptic creature project: social human-robot interaction through affective touch. In *Proceedings of the AISB 2008 Symposium on The Reign of Catz and Dogz: The Second Symposium on the Role of Virtual Creatures in a Computerised Society*, 1:7-11, Aberdeen, UK, 2008.
[6] D. Maulsby, S. Greenberg and R. Mander, "Prototyping an intelligent agent through Wizard of Oz". In *Proceedings of CHI'93*, The Netherlands, ACM Press, 1993.
[7] NEWHRI: Unifying characteristics of research in human-robot interaction. *Workshop at IEEE ICRA 2008*, Pasadena, California. See: http://newhri.cs.brown.edu/Home.html
[8] J. Saldien, K. Goris, S. Yilmazyildiz, W.Verhelst and D. Lefeber "On the design of the huggable robot Probo", *Journal of Physical Agents* 2(2), June, 2008.
[9] W. D. Stiehl, J. Lieberman, C. Breazeal, L. Basel, R. Cooper, H. Knight, L. Lalla, A. Maymin and S. Purchase, "The huggable: a therapeutic robotic companion for relational, affective touch", *IEEE Consumer Communications and Networking Conference*, Las Vegas, US, 2006.
[10] R. Looije, M. A. Neerincx and V. de Lange. "Children's responses and opinion on three bots that motivate, educate and play", *Journal of Physical Agents* 2(2), June, 2008.
[11] A. J.N. Van Breemen, K. Crucq, B. J.A. Krose, M. Nuttin, J. M. Porta and E. Demeester, "A user-interface robot for ambient intelligent environments", in *Proceedings of ASER 2003*, Bardolino, Italy, 2003.
[12] C. F. DiSalvo, F. Gemperle, J. Forlizzi and S. Kiesler, "All robots are not created equal: the design and perception of humanoid robot heads", In *Proceedings of DIS 2002* New York, ACM, 2002.
[13] S. L. Lee and S. Kiesler, "Human mental models of humanoid robots" In *Proceedings of IEEE ICRA 2005*, Spain, 2005.
[14] C. Bartneck, J. Reichenbach and J. Carpenter, "Use of praise and punishment in human-robot collaborative teams", In *Proceedings of the RO-MAN 2006*, Hatfield, UK, 2006.
[15] T. Komatsu and S. Yamada, "How do robots agents' appearances affect people's interpretations of the agents' attitudes?" In *Proceedings of ACM CHI 2007*.
[16] K. F. MacDorman and P. H. Khan Jr., "Psychological benchmarks of human-robot interaction", *Interaction Studies* 8(3): 359-362, 2007.
[17] B. Duffy, "Anthropomorphism and the social robot", *Robotics & Autonomous Systems* 42: 177-190, 2003.
[18] K. F. MacDorman and H. Ishiguro, "Opening Pandora's uncanny box", *Interaction Studies*, 7: 361-368, 2006.
[19] M. Lohse, F. Hegel, and B. Wrede, "Domestic Applications for Social Robots—an online survey on the influence of appearance and capabilities", *Journal of Physical Agents* 2(2), June, 2008.
[20] M. Heerink, B. Kröse, V. Evers and B. Wielinga, "The influence of social presence on acceptance of a companion robot by older people", *Journal of Physical Agents* 2(2), June, 2008.
[21] R. Sparrow and L. Sparrow, "The the hands of machines? The future of aged care", *Minds and Machines* 16: 141-161, 2006.
[22] J. Lee, W. Bradley Knox and P. Stone, "Inter-classifier feedback for human-robot interaction in a domestic setting", *Journal of Physical Agents* 2(2), June, 2008.
[23] M. Jacobsson, Y. Fernaeus and L. E. Holmquist, "GlowBots: Designing and Implementing Engaging Human-Robot Interaction", *Journal of Physical Agents* 2(2), June, 2008.

On the Design of the Huggable Robot Probo

Jelle Saldien, Kristof Goris, Selma Yilmazyildiz, Werner Verhelst and Dirk Lefeber

***Abstract*—Nowadays robots are being created that interact with human beings in order to satisfy certain social needs. Following this trend, the development of the social robot Probo has started. The robot will be used in hospitals, as a tele-interface for entertainment, communication and medical assistance. Therefore, it requires the ability to express emotions. In order to do so, an emotional interface is developed to fully configure the display of emotions. These emotions -represented as a vector in an emotion space- are mapped to the degrees of freedom used in the robot. Besides emotions, the interface includes a control for the point of attention and a module to create and store animations. A 3D virtual model is created, acting as a virtual replica of the robot, providing realistic visual feedback to evaluate the design choices for the facial expressions. This paper presents the objectives of this new robot and describe the concepts and design of the first prototype.**

***Index Terms*—Human robot interaction, robot assisted therapy, robot design.**

I. INTRODUCTION

A Hospitalization has a serious physical and mental influence, particularly on children. It confronts them with situations which are completely different from these at home. In a hospital, children's experiences are limited due to the closed and protective environment, leading to many difficulties [1].

Animal-assisted therapy (AAT) and animal-assisted activities (AAA) are becoming commonly used in hospitals, especially in the United States [2]. AAT and AAA are expected to have useful psychological, physiological and social effects. Some psychological studies have already shown that animals can be used to reduce heart and respiratory rate [3], lower levels of stress [4], progress mood elevation and social facilitation. However animals are difficult to control, they always have a certain unpredictability, and they are carrier of disease and allergies. Therefore, the use of robots (instead of animals) has more advantages and a better chance of being allowed in hospitals. Using these social pet robots for therapeutic purposes is termed robot-assisted therapy (RAT). For example, the seal robot Paro is used for pediatric therapy at university hospitals [5][6]. Currently, Sony's dog robot Aibo [7], Philips' iCat [8] and Omron's Necoro [9] are also being tested for RAT. As a part of the ANTY project the development of the huggable robot Probo has started. The main goal for the robot Probo is to create a friend for children, acting as an *interface* between the real, sometimes hard, hospital world and the imaginary fantasy world wherein children grow up. The robot will also be used as a multidisciplinary research platform, giving other researchers the opportunity to improve and explore the possibilities of RAT. Communication will be the first focus of this robot. Having a fully actuated head, the robot is capable of expressing a wide variety of facial expressions, in contrast with other comparable robots such as; Paro, Huggable [10], Aibo and Necoro. Philip's iCat has facial expression of emotions, but lacks the huggable appearance and warm touch that attracts children. Probo will emphasize its *expression of emotions* by using his nonsense affective speech. Probo must fulfill the specifications to operate in a hospital environment and guarantee a smooth interaction with children. The *intrinsic safety* when dealing with human robot interaction is therefore of high priority.

Saldien, Goris and Lefeber are with the Robotics and Multibody Mechanics research group of the Vrije Universiteit Brussel.
E-mail: {jsaldien, kgoris, dlefeber}@vub.ac.be
Yilmazyildiz and Verhelst are with the Digital Speech and Signal Processing research group of the Vrije Universiteit Brussel.
E-mail: {syilmazy, wverhels}@etro.vub.ac.be

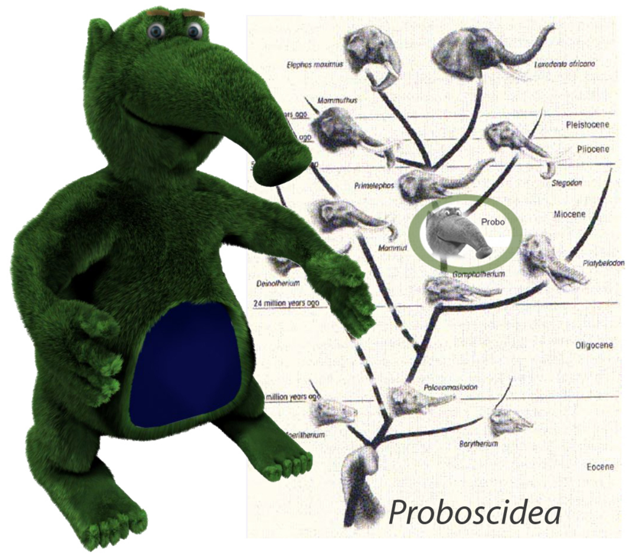

Fig. 1. A 3D computer model representing the huggable robot Probo.

II. PROBO

A. A huggable robotic imaginary animal

The name *Probo* is derived from the word ***Proboscidea***, the order containing only one family of living animals, *Elephantidae* or *the elephants*, with three living species (African Bush Elephant, African Forest Elephant, and Asian Elephant) [11]. During the period of the last ice age there were more, now extinct species, including a number of species of the elephant-like *mammoths* and *mastodons*.

The looks of the robot in Figure 1 represents an imaginary animal based on the ancient mammoths. The main aspects are the huggable appearance, the attractive trunk or proboscis, and the interactive belly-screen. The internal mechanics of the robot will be surrounded by foam and a removable fur-jacket, in such a way that Probo looks and feels like a stuffed animal. The basic design of the robot is based on an imaginary animal, so that there is no exact similarity with a well-known creature. The combination of a caricatured and zoomorphic [12] representation of a mammoth-like animal is useful and effective to accomplish the goals, rather than using complex, *realistic* representations. The color of Probo is green because this color evokes mainly positive emotions such as relaxation and comfort. In [13], the relationship between color and emotion was tested, whereas the color green attained the highest number of positive responses (95.9%), followed by the color yellow (93.9%). The majority of emotional responses for the green color indicated the feelings of relaxation and calmness, followed by happiness, comfort, peace, hope, and excitement. The green color was associated with nature and trees, and thus creating feelings of comfort and soothing emotions.

B. *A tele-interface*

The robot Probo will be used as a tele-interface focusing on entertainment, communication and medical assistance. A touch screen in the belly of the robot creates a window to the outside world and opens up a way to implement new and existing computer applications.

1) Entertainment: Young children have a strong need for distraction and entertainment. Providing them with a robotic user interface (RUI) will extend the possibilities of interactive game playing and includes the capability of emotional feedback.

2) Communication: Hospitalized children are sometimes placed in a isolated environment, strongly reducing the communication with friends and family. The robot can function as a perfect interface to communicate with other people using standard videoconferencing techniques. The eyes of the robot will house the cameras, whereas the screen in the belly will display the image, giving the opportunity to establish interactive video-communication.

3) Medical Assistance: The robot interface can be used by the medical staff to inform the children about medical routines or operations. In the same philosophy, Probo can comfort children during difficult medical procedures. The unknown environment will be explored and examinations will be described in a child friendly manner. By using predefined scenarios with pictures, video and sounds children can pre-experience the medical routines, guided by Probo. A good preparation for the examinations will reduce the child's fear, providing the medical staff with better results when assessing the child's pain factor.

C. *A social interface*

Children will have some basic expectations as the robot represents a living animal, resulting in the necessity to react on primary stimuli and to have natural movements. In order to establish some bond with the children, Probo must be able to communicate. In daily life, people rely on face-to-face communication and the face plays a very important role in the expression of character, emotion and/or identity [14]. Mehrabian [15] showed that only 7% of information is transferred by spoken language, that 38% is transferred by paralanguage and 55% of transfer is due to facial expressions. Facial expression is therefore a major modality in human face-to-face communication. To start face-to-face communication with children, the robot is equipped with an intriguing trunk in the middle of its face, provoking children to interact with its trunk and stimulate them to maintain their focus on its face.

In [16], Breazeal defines four classes (social evocative, social interface, socially receptive, sociable) of social robots in terms of; (1) how well the robot can support the social model that is ascribed to it and, (2) the complexity of the interaction scenario that can be supported. This project aims to start working with the robot Probo as a social interface, providing a *natural* interface by employing human-like social cues and communication modalities. In this first phase the focus is the construction of a physical prototype with an actuated head, trunk and facial expressions.

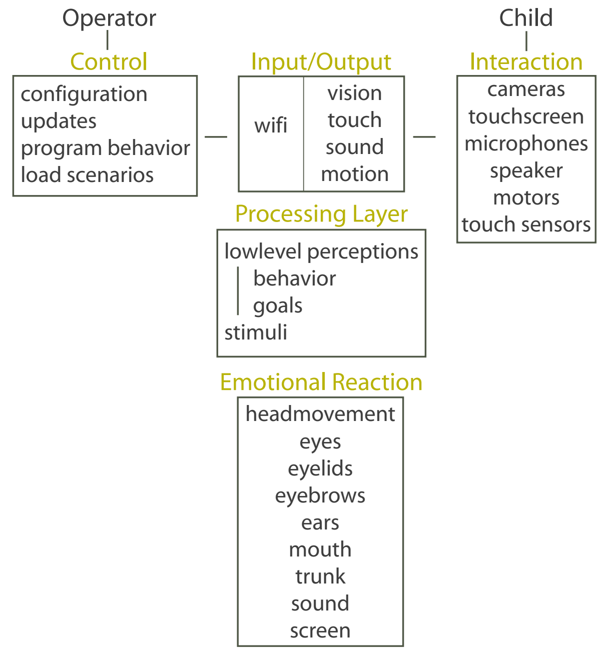

Fig. 2. The Robotic User Interface (RUI) between an operator and a child

D. *Operational Concept*

At first, the prototype is used as a RUI (Figure 2) interacting with children and controlled by an operator. The operator can be every person who wants to communicate with the child, in particularly caregivers and researchers. The robot functions as an interface that performs preprogrammed scenarios and reacts on basic input stimuli. The input stimuli, coming from low-level perceptions, are derived from vision analysis, audio analysis and touch analysis. Those stimuli will influence the attention-system and emotion-system, used to set the robot's point of attention, current mood and corresponding facial expression. The vision analysis includes the detection of faces, objects and facial features. Audio analysis includes detecting the direction and intensity of sounds and the recognition of emotions in speech.

A specific behavior-based framework is being developed to process these input stimuli. The framework is based on earlier work of Ortony, Norman and Revelle [17], who focus on the interplay of affect, motivation and cognition in controlling behavior. Each is considered at three levels of information processing: *the reactive level* is primarily hard-wired and has to assure the quick responses of the robot to make it look alive; *the routine level* provides unconscious, un-interpreted scenarios and automotive activity; and the *reflective level* supports higher-order cognitive functions, including behavioral structures and full-fledged emotions, finally resulting in a sociable robot. Starting with a *social interface*, the reactive and routine level are being implemented. Currently, there is a shared control between the operator, configuring behavior, emotions and scenarios, and the robot, having basic autonomous reactions. Further research and development is required to enhance the robot's emotions and behavior, by implementing a cognitive software architecture at the reflective level to successfully obtain a *sociable robot* in the end. Therefore a study and implementation of joint attention mechanisms for human-robot communication has been started.

E. Nonsense Affective Speech

The robot Probo will speak to the children using nonsense affective speech, which will be a cross-cultural and understandable language for most of the children regardless of their own native language. In one of the current approaches [18], this speech is produced by using a database with natural expressive speech samples and a database with neutrally spoken speech samples, both recorded with a professional speaker. From the neutral speech examples, carrier sentences of the non-existing language for Probo will be produced by firstly segmenting the recorded utterances into several nonsense syllables and then concatenating them in the same syllabic structure as the desired emotional prosodic template, which is selected from the expressive database. To produce emotional speech for that non-existing language, the same pitch and timing structure as found in the prosodic template are copied on the nonsense carrier phrase, a process that is known as prosodic transplantation and that effectively provides the synthetic output with a same intonation pattern as the natural expressive example.

III. Mechanical Design

The first prototype of the robot has 20 Degrees Of Freedom (DOF) to obtain a fully-actuated head and trunk (Figure 3). By moving its head (3 DOF), eyes (3 DOF), eyelids (2 DOF), eyebrows (4 DOF), ears (2 DOF), trunk (3 DOF) and mouth (3 DOF) the robot is able to express its emotions [19]. The trunk of the robot is the most remarkable element. When a child interacts with this trunk, it points its attention towards the face of the robot, locating itself in the scope of the onboard cameras, allowing proper vision analysis. Using these cameras, located in the eyes, the robot will be able to focus on a point of attention and follow it with natural eye-movements [20]. The robot will use eyebrows, ears and eyelids to express moods and feelings. Flexible materials and compliant actuators are

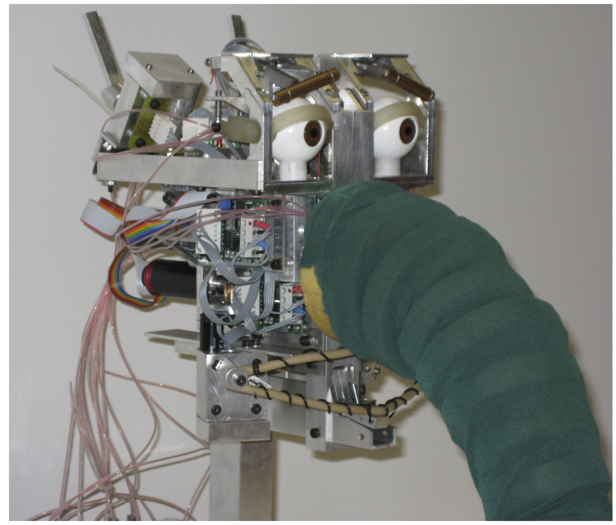

Fig. 3. The prototype of the head of Probo

TABLE I
DOF and ranges of the actuated joints of Probo's head in comparison with other prominent non-humanoid robot heads

Kismet	Eddie	iCat	Probo		
(DOF)				Range [°]	
Eyes (3)	Eyes (3)	Eyes (3)	Eyes (3)	Pan Tilt	100 80
Eyelids (2)	Eyelids (4)	Eyelids (2)	Eyelids (2)		150
Brows (4)	Brows (4)	Brows (2)	Brows (4)		45
Ears (4)	Ears (4)		Ears (2)		90
Yaw (1)	Yaw (1)		Mouth (3)	Yaw	45
Lips (4)	Lips (4)	Lips (4)		Lipcorners	60
	Crown (1)		Trunk (3)		360

being applied, considering a safe interaction. Because of the high hospital requirements on hygiene, the fur of the robot can be easily replaced and washed prior to each visit. The prototype measures about 66cm in height and 32cm in width.

A. Degrees of Freedom

For the display of the emotions most of the DOF in the face are based on the Action Units (AU) defined by the Facial Action Coding System (FACS) developed by Ekman and Friesen [21]. AU express a motion of mimic muscles as 44 kinds of basic operation, with 14 AU to express the emotions of anger, disgust, fear, joy, sadness, and surprise, which are often supported as being the 6 basic emotions from evolutionary, developmental, and cross-cultural studies [22]. Because the robot does not have a human face and in order to simplify the design, some of the AU are missing, others are replaced and some are added. The lack of the lower eyelid and a fixed upper lip lead to missing AU, the AU regarding the nose movements will be replaced by the movement of the 3 DOF trunk. The movement of the ears and the greater visual influence of the trunk will add extra gestures to express the emotions. Table I shows the DOF of Probo's robot head compared with some other non-android robot heads.

B. Soft Actuation

Most of the robots are actuated by electric drives as these actuators are widely available and their control aspects are well-known. Because of the high rotational speed of the shaft and the low torque of an electrical motor, a transmission unit is often required. Due to the high reflected inertia of the transmission unit, the joint must be seen as rigid. For safe and soft interaction the joints need to be flexible, which can be obtained by incorporating compliant actuation. Compliant actuators are gaining interest in the robotic community. Pneumatic artificial muscles [23](such as McKibben muscles, Festo muscles, PPAM [24]), electric compliant actuators (such as VIA [25], AMASC [26] and MACCEPA [27]) and voice coil actuators [28] are some examples of compliant actuators. While some of them exhibit adaptable compliance, so that the stiffness of the actuated joint can be changed, it is not required in the Probo robot. Therefore, compliance is introduced by placing elastic elements between the motor and the actuated robot joint. In this way the external forces on the joint will be dissipated by the elastic elements, resulting in safe and flexible joints. It is more complex to do precise positioning with compliant actuators than with classic high positioning (non-compliant) actuators, typically used in industrial applications, however, the intrinsic safety introduced in the system is of major importance.

C. Materials

In this stage of the development, most mechanical parts of the prototype are made of aluminum because it is a strong, lightweight and tractable material. Some very specific and complex parts are manufactured using rapid prototyping. To comply to the design constraints stated earlier our mechanical robotic part is encapsulated in a foam layer. This layer of flexible polyurethane provides a soft touch, protects the robotics inside and gives the robot a final form. On top of the foam layer the robot will have a removable fur-jacket, which can be washed and disinfected. The fur-jacket, which is a 100% cotton fabric, complies to the European toy safety standards EN71-1, EN71-2 and EN71-3. The use of the soft actuation principle together with well-thought designs concerning the robot's filling and huggable fur, are both essential to create Probo's soft touch feeling. To realize a full-body sense of touch, a sensitive skin will be used. A good example is being developed (by Stiehl et al. [10]) for a therapeutic robotic companion named: The Huggable. In another approach, research has started for the use of photonic crystal fibers [29] which will be implemented in some parts of Probo, such as the trunk.

IV. Modular System Architecture

Besides the restrictions mentioned above, the prototype designer has to bear in mind the need of a modular mechanical system architecture to simplify assemblage and maintenance. This approach leads to an effective development and realization of a robot prototype and requires the use of transferable mechanical and electronic components. Due to a lack of commercially available standard mechanic and electronic modules e.g. eyes, eyebrows, trunk, etc. one must design prototype dependant modules. In the next paragraphs the different modules with the AU needed to display facial expressions are described. Each module can be easily replaced without effecting the others.

A. Eyes and Eyebrows

Besides the role of the eyes to show some facial expressions, there are two additional reasons to equip a social robot with actuated eyes.

1) Eye-gaze based interaction: The phenomenon that occurs when two people cross their gaze is called eye contact [30], furthermore, people use eye-gaze to determine what interests each other. The same phenomenon will be used between robot and child to encourage human robot interaction. By focussing the robot's gaze to a visual target, the person that interacts with the robot can use the robot's gaze as an indicator of its intentions. This facilitates the interpretation and readability of the robot's behavior, as the robot reacts specifically to what it is looking at [31]. This *visual target* will be referred to as the robot's *point of attention* (POA).

2) Active vision: When a robot is intended to interact with people, it requires an active vision system that can fulfill both a perceptual and a communicative function. An active vision system is able to interact with its environment by altering its viewpoint rather than passively observing it. Therefore, the designed eyes are hollow and contain small cameras. As these cameras can move, the range of the visual scene is not restricted to that of the static view. Although the aim is a pet-type robot, the design of the robot eyes are based on that of human anthropomorphic data. The imitation of anthropomorphic eyes gives the impression of being natural.

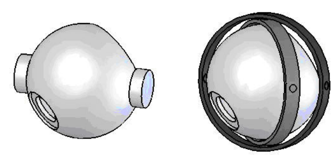

Fig. 4. Eye supports.

Two eye-supports candidates are shown in Figure 4. The support shown on the left, holds the eye-ball between two Teflon®parts with the same spherical curvature as the eye-ball itself, resulting in three DOF just like in a spherical joint, and a smooth rotation around the center of the sphere due to the low friction. Because there is no mechanical part that intersects the eye-ball, the eyes can bulge out of the head. The second concept (on the right in Figure 4) consists of two rings and two axis. One rotation axis passes through the center point of the eye and holds the eye in the inner ring. This way the eye can rotate relatively to the inner ring. A second rotation axis passes through the inner and outer ring, allowing the inner ring to rotate with respect to the outer ring. While panning the eye, the inner ring comes out of the plane of the other ring, whereas the eye can not bulge out as far as in the former support. Most of the other mentioned robot heads use the second support type or a variant on it, which in our case could lead to the visibility of mechanical parts or the disability to bulge out the eyes. For this reason, the first support type has been chosen .

The five DOF eyes module exists of two hollow eyeballs mounted with the chosen eye-support as shown in Figure 5. According to the chosen DOF based on the AU mentioned earlier; the eyes can pan separately and tilt together, each

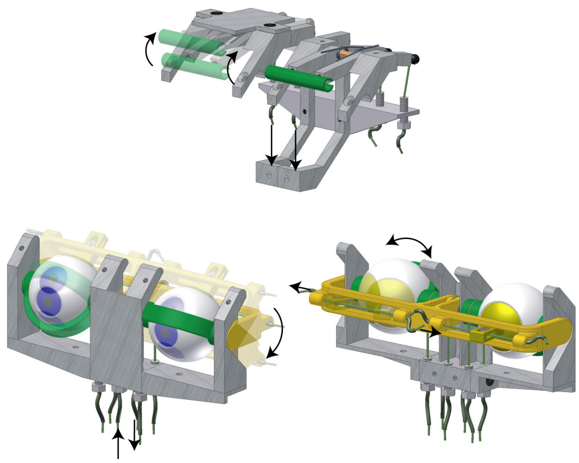

Fig. 5. CAD of Eyebrows (top) and Eyes (bottom) mechanism.

eye can be covered by an upper eyelid and the eyelids can blink separately. The eyebrows module fits on top of the eyes module. Each eyebrow has two DOF meaning that both the vertical position and the angle of each eyebrow can be set independently. Nine of the shelve hobbyist servomotors, together with a Bowden cable mechanism are used to power the eyes, eyelids and eyebrows. Axial springs and the usage of flexible cables both introduce compliance. Using flexible Bowden cables creates the opportunity to group and isolate the different servos and to place them anywhere in the robot. That way heat and noise dissipation can be controlled and the head can be held light-weighted, both resulting in a safe design.

B. Trunk

The trunk or *proboscis* of Probo seems to be the most intriguing element concluding the results of a small survey amongst children aged 10-13. In this survey, it was observed that all the children first touched the trunk, and most of them also start playing with it. That is why the trunk is used to grab and maintain the child's attention. When the child's attention is focussed on the trunk, the child's face fits within the scope of the on board eye cameras. In this way the child's attention can be guided towards the face of the robot, to start face to face communication by using Probo's facial expressions.

The three DOF trunk as shown in Figure 6 consists of a foam core with segmented extension discs. The trunk is created using *FlexFoam-iT!^{TM}X*, a two-component flexible urethane foam with a $160kg/m^3$ density cell structure, with a silicone mold. Axial to the centerline, three flexible cables are guided through the discs and attached to the front disc. The end of each cable is attached to a wind-up pulley resulting in a motion of the entire trunk. The motion of the trunk depends on; the number of discs, the dimensions of the discs and the core, the flexibility of the cables and the composition of the foam. A high compliance and durability of the trunk is ensured by using a foam material actuated by flexible cables. Interaction with this trunk will be safe both for the child, that can not be hurt, and for the motors, that can not be broken.

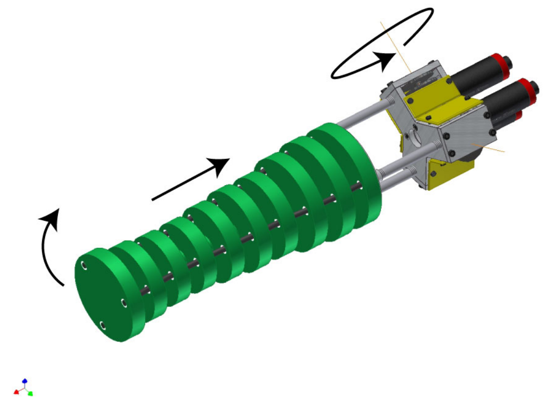

Fig. 6. CAD of the trunk.

Fig. 7. Test of the trunk movements.

Three maxon brushless motors are used to actuate the trunk. Each motor is coupled with a worm worm-wheel gear train to reduce the rotational speed and to increase the output torque. A worm drive is used because of its self locking capability. If, during interaction, the trunk is grasped, it will follow the grasp motion until it is released. When released, the trunk will return to its set position. That is because all external forces on the trunk will be stored and released by the elastic cables. Optical encoders are used to calculate the angular displacement of the pulleys to estimate the position of the trunk.

C. Mouth and Ears

The mouth and ears are both actuated to contribute to the robots facial expressions. In addition to the expressions, the mouth also serves to enhance the affective speech by performing basic lip-sync movements. Probo's mouth has an upper lip and a lower lip, the middle of the upper lip is attached to the trunk and the middle of the lower lip can move vertically so that the mouth can open. Both lips come together in the mouth's corners, which are actuated. The mechanism used for actuating the mouth corners is the same as that used in

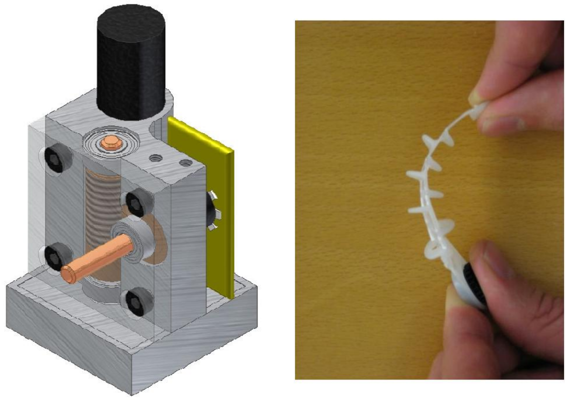

Fig. 8. Modular actuation mechanism (left) and flexible ear material (right).

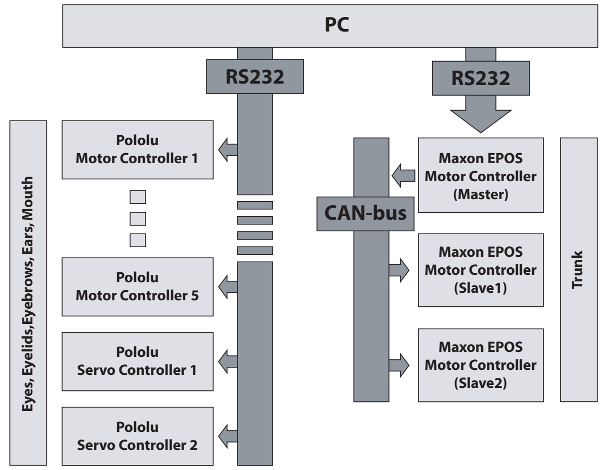

Fig. 9. Architecture of the motor controllers.

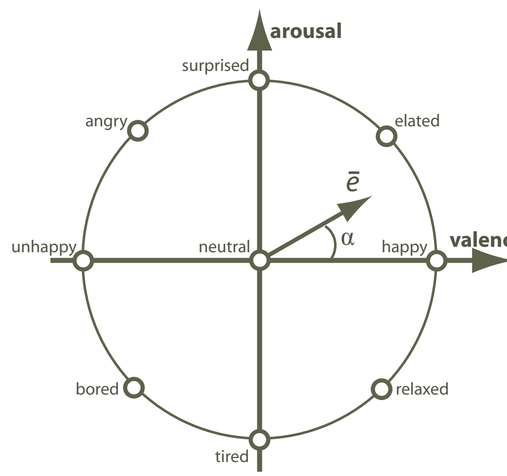

Fig. 10. Emotion space based on the circumplex model of affect defined by Russell [33].

the ears module, shown in Figure 8. It consists of a brushed maxon motor with a planetary gear train. The first gear train is followed by a second one, which is a worm drive. Position measurement is established by an absolute position sensor fixed on the output shaft. On the output shaft either an ear or a mouth corner is attached. Opening the mouth is established by movement of the middle of the lower lip. Compliance is introduced by the shape of the ear and mouth corners and by means of flexible materials. The actuated part is flexible in a perpendicular direction, and stiff in the tangent direction. Position measurement of the joints is also established by absolute position sensors. In comparison with [5],[13] and [18], Probo has less DOF in the mouth. Each ear has one DOF. The movement of the robotic ear is a rotation which consists of two combined rotations. The first rotation turns the entire ear while the second rotation twists the ear axially. That way the ear's opening is pointed to the front when the robot is attentive and the opening is pointed to the ground when the ear lies flat to the back.

V. Electronics and Control Software

The maxon brushless motors, which are used to actuate the trunk, are driven by maxon's EPOS motor controllers. The maxon brushed motors, used in the mouth and the ears, are driven by Pololu's motor controllers with position feedback and the hobbyist servo motors, for the eyes and eyebrows, are driven by Pololu's micro serial servo controllers. Figure 9 shows the architecture.

A Personal Computer (PC) is used to control the different motors. Two serial ports, using the RS232 protocol, are used to communicate with the motor controllers. The first serial port communicates with one of the three maxon EPOS motor controllers. This controller acts as a master in a master-slave set up with the two other maxon EPOS motor controllers (slaves). The communication between master and slaves is performed with a CAN-bus. The second serial port communicates with all Pololu controllers. Despite the use of serial communication and the high number of motor positions and speeds needed to refresh, the refresh time rests less than the mechanical inertia and is consequently acceptable.

The control software running on the host PC is written in C# using the Microsoft®.NET framework, it sends the desired motor positions and speeds to the respective motor controllers. This software component is linked with the emotional interface, providing a real time control for setting the emotions, a specific point of attention or to display programmed animations and the ability for visual feedback of the virtual model, which receives the same motor positions.

VI. Facial Expressions

A. Emotional interface

Several theorists argue that a few select emotions are basic or primary, they are endowed by evolution because of their proven ability to facilitate adaptive responses to the vast array of demands and opportunities a creature faces in its daily life [22] [32]. To achieve a translation from emotions into facial expressions, emotions need to be parameterized. In the robot Kismet [34], facial expressions are generated using an interpolation-based technique over a three-dimensional, componential *affect space* (arousal, valence, and stance). In this model two dimensions; valence and arousal are used

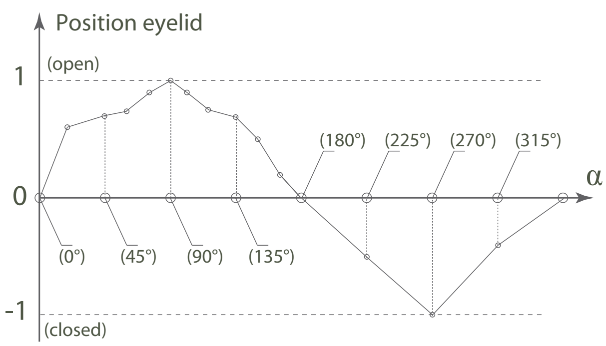

Fig. 11. Adjustable interface for defining the value off the DOF (controlling the position of the eyelid) for each emotion (angle α).

to construct an emotion space, based on the circumplex model of affect defined by Russell [33], which has as well been implemented in the robot Eddie [35]. In the emotion space a Cartesian coordinate system is used, where the x-coordinate represents the valence and the y-coordinate the arousal, consequently each emotion $e(v, a)$ corresponds to a point in the valence-arousal plane (Figure 10). This way, the basic emotions can be specified on a unit circle, placing the neutral emotion $e(0, 0)$ in the origin of the coordinate system. Now, each emotion can also be represented as a vector with the origin of the coordinate system as initial point and the corresponding valence-arousal values as the terminal point. The direction α of each vector defines the specific emotion, whereas the magnitude defines the intensity of the emotion. The intensity i can vary from 0 to 1, interpolating the existing emotion $i = 1$ with the neutral emotion $i = 0$. Each DOF that influences the facial expression is related to the current angle α of the emotion vector. An adjustable interface is developed to define the specific value for each angle $(0° - 360°)$ of each DOF. When selecting one DOF, a value for each basic emotion is set on the unit circle. To attain a contiguous relation, a linear interpolation between the configuration points is applied. By adding more (optional) points or values the curve can be tuned to achieve smooth, natural transitions between the different emotions. An example is shown (Figure 11) for the DOF that controls the eyelid, extra points were added in the first half of the emotion space respectively starting and ending with the happy emotion $(\alpha = 0° = 360°)$.

An emotional interface (Figure 12) has been developed wherein the user can fully configure the facial expressions and use the emotion space to test the different emotions and transitions. The user will obtain visual feedback from a virtual model of the robot. In addition to the facial expression this interface has been extended with a component controlling the point of attention. This component controls the eyes and neck motion according to a specific point in the three dimensional space. The respective coordinates of that point can be altered in real time and will be represented as a red cube in the virtual space. This coordinate is translated into rotation angles

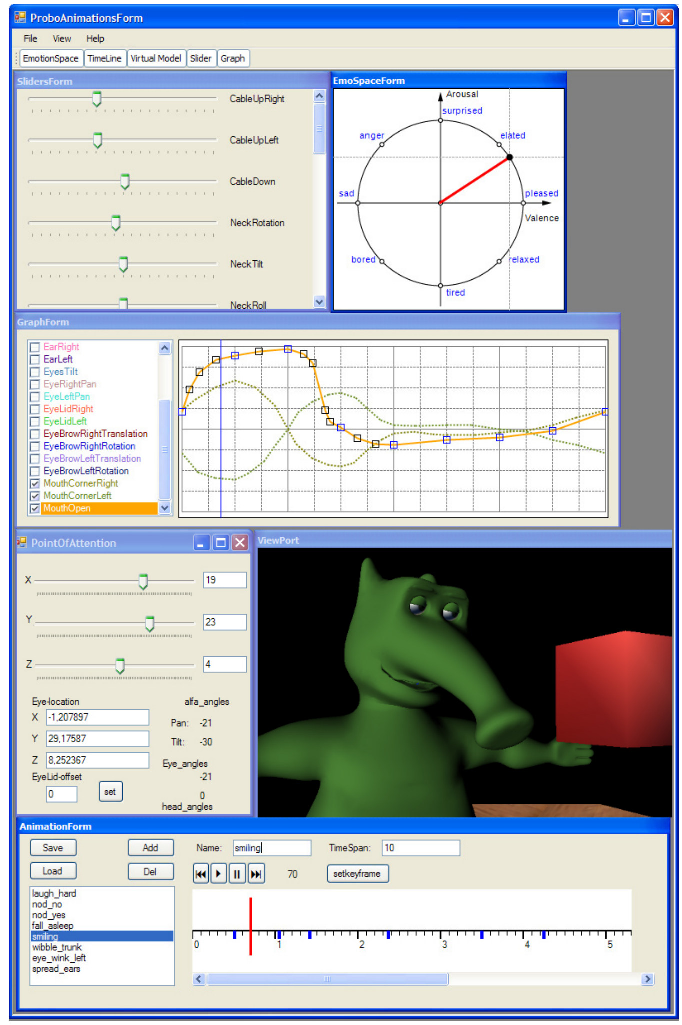

Fig. 12. Emotional interface for controlling facial expressions, point of attention and animations.

for the 4 DOF controlling the eyes (pan/tilt) and the head (pan/tilt). As part from the vision analysis, a face recognition component is developed using Intel®'s OpenCV library. This component uses a webcam to capture the images and then calculates the center of the face as a cartesian coordinate. This coordinate can then be used to control the point of attention in the virtual space. Another component in this interface gives the user the ability to create animations, store, edit and play them. Each animation consists of different key frames, which hold the values of the DOF at a given time. There is a linear interpolation between the different key frames resulting in a contiguous animation. The emotional interface can be used to easily insert emotions at a certain point in an animation. The different animations are stored in a database and will be employed later to build scenarios for the robot.

B. Virtual model

A virtual model of Probo has been created to evaluate the design choices and to advance on user testing, without the need for an actual prototype. The virtual model is created combining the mechanical designs (using Autodesk®Inventor®) with

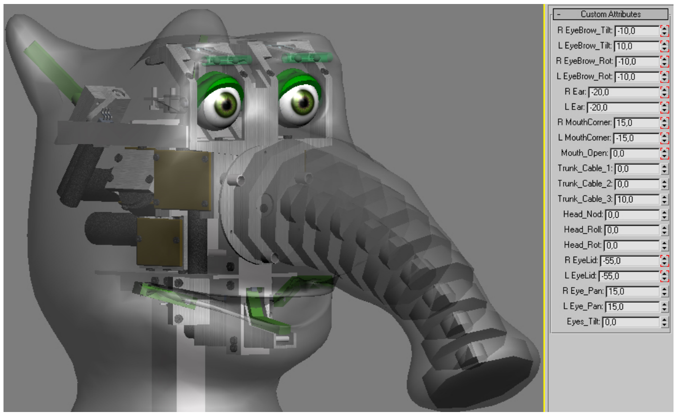

Fig. 13. Virtual model with control slider for the DOF.

the visual exterior of our robot, represented by the skin (using Autodesk®3ds Max®). The mechanical parts are linked together to obtain kinematical movements for realistic visual motions of the model. The skin is attached on the mechanical parts using skinning techniques in 3ds Max®. The movements can be controlled by using sliders to set the desired angle for each DOF and simulate actuation of the parts (Figure 13). This model has also been implemented in Microsoft®XNA™framework where it is linked to the emotional interface to simulate the motions of the robot. Another benefit of this virtual model is that the positions of our body parts are known at anytime, which are practically the same as these in the real robot. Position feedback will be implemented using potentiometers on the DOF of the robot to improve the accuracy of the virtual model.

VII. Conclusion and Future Work

The first steps in the creation of a social interface were successfully established. By using specific materials and compliant actuation, the durability and safety of the robot Probo is guaranteed. Based on the AU, a modular and efficient design for the DOF is realized and implemented. The developed software controls the virtual model, by using the emotion space, setting the point of attention and programming new animations. The virtual model provides visual feedback on every motion, as if it was the real robot. All the DOF of the physical prototype can be tested and configured. Using our emotional interface all the emotions can be translated into the values for each DOF. To fully cover all the emotions, the emotion space can be extended with a third dimension: stance, which will allow us to make more difference between anger and fear. By combining techniques from CAD and animation software, a fully realistic virtual prototype was created. In the next steps the virtual model will be connected with the interface controlling the actual motors, resulting in a real time control interface that can be used by an operator.

Acknowledgment

The authors would like to thank the Brussels-Capital Region and the ANTY Foundation for their support.

References

[1] J. Vanderfaeillie and V. Y., *Handboek orthopedagogische hulpverlening. Deel 1. Een orthopedagogisch perspectief op kinderen en jongeren met problemen*. Leuven/Voorburg: Acco, 2005, ch. Zieke kinderen en jongeren, pp. 199–237.

[2] M. Burch, “Animal-assisted therapy and crack babies: A new frontier,” *Pet Partners Program: A Delta Society Newsletter*, 1991.

[3] K. Allen, J. Blascovich, J. Tomaka, and R. Kelsey, “Presence of human friends and pet dogs as moderators of autonomic responses to stress in women,” *Journal of personality and social psychology*, vol. 61, no. 4, pp. 582–589, 1991.

[4] G. Ballarini, “Pet therapy. Animals in human therapy.” *Acta Biomed Ateneo Parmense*, vol. 74, no. 2, pp. 97–100, 2003.

[5] T. Shibata, T. Mitsui, K. Wada, A. Touda, T. Kumasaka, K. Tagami, and K. Tanie, “Mental commit robot and its application to therapy of children,” *Advanced Intelligent Mechatronics, 2001. Proceedings. 2001 IEEE/ASME International Conference on*, vol. 2, 2001.

[6] T. Shibata, K. Wada, T. Saito, and K. Tanie, “Robot Assisted Activity for Senior People at Day Service Center,” *Proc. of Intl Conf. on Information Technology in Mechatronics*, pp. 71–76, 2001.

[7] T. Tamura, S. Yonemitsu, A. Itoh, D. Oikawa, A. Kawakami, Y. Higashi, T. Fujimooto, and K. Nakajima, “Is an Entertainment Robot Useful in the Care of Elderly People With Severe Dementia?” *Journals of Gerontology Series A: Biological and Medical Sciences*, vol. 59, no. 1, pp. 83–85, 2004.

[8] A. van Breemen, “iCat: Experimenting with Animabotics,” *Proceedings, AISB 2005 Creative Robotics Symposium*, 2005.

[9] A. Libin and E. Libin, “Person–Robot Interactions From the Robopsychologists’ Point of View: The Robotic Psychology and Robotherapy Approach,” *Proceedings of the IEEE*, vol. 92, no. 11, pp. 1789–1803, 2004.

[10] W. Stiehl, J. Lieberman, C. Breazeal, L. Basel, L. Lalla, and M. Wolf, “Design of a therapeutic robotic companion for relational, affective touch,” *Robot and Human Interactive Communication, 2005. ROMAN 2005. IEEE International Workshop on*, pp. 408–415, 2005.

[11] D. Wilson and D. Reeder, *Mammal Species of the World: A Taxonomic and Geographic Reference*. Johns Hopkins University Press, 2005.

[12] T. Fong, I. Nourbakhsh, and K. Dautenhahn, “A survey of socially interactive robots,” *Robotics and Autonomous Systems*, vol. 42, no. 3-4, pp. 143–166, 2003.

[13] N. Kaya and H. Epps, “Relationship between Color and Emotion: A Study of College Students.” *College Student Journal*, vol. 38, no. 3, pp. 396–406, 2004.

[14] J. Cole, *About Face*. Mit Pr, 1998.

[15] A. Mehrabian, “Communication without words,” *Psychology Today*, vol. 2, no. 4, pp. 53–56, 1968.

[16] C. Breazeal, “Toward sociable robots,” *Robotics and Autonomous Systems*, vol. 42, no. 3-4, pp. 167–175, 2003.

[17] A. Ortony, D. Norman, and W. Revelle, “7 Affect and Proto-Affect in Effective Functioning,” *Who Needs Emotions?: The Brain Meets the Robot*, 2005.

[18] S. Yilmazyildiz, W. Mattheyses, Y. Patsis, and W. Verhelst, “Expressive Speech Recognition and Synthesis as Enabling Technologies for Affective Robot-Child Communication,” *Lecture Notes in Computer Science*, vol. 4261, p. 1, 2006.

[19] J. Saldien, K. Goris, B. Vanderborght, B. Verrelst, R. Van Ham, and D. Lefeber, “ANTY: The development of an intelligent huggable robot for hospitalized children,” *CLAWAR*, 2006.

[20] K. Goris, J. Saldien, B. Vanderborght, B. Verrelst, R. Van Ham, and D. Lefeber, “The Development of the Eye-System for the Intelligent Huggable Robot ANTY,” *CLAWAR*, 2006.

[21] P. Ekman and W. Friesen, *Facial Action Coding System*. Consulting Psychologists Press, 1978.

[22] P. Ekman, “Are there basic emotions?” *Psychological review*, vol. 99, no. 3, pp. 550–553, 1992.

[23] F. Daerden and D. Lefeber, “Pneumatic artificial muscles: actuators for robotics and automation.” *European Journal of Mechanical and Environmental Engineering*, vol. 47, no. 1, pp. 11–21, 2002.

[24] B. Verrelst, R. Van Ham, B. Vanderborght, D. Lefeber, F. Daerden, and M. Van Damme, “Second generation pleated pneumatic artificial muscle and its robotic applications,” *Advanced Robotics*, vol. 20, no. 7, pp. 783–805, 2006.

[25] G. Tonietti, R. Schiavi, and A. Bicchi, “Design and control of a variable stiffness actuator for safe and fast physical human/robot interaction,” *Proceedings of the IEEE International Conference on Robotics and Automation*, pp. 528–533, 2000.

[26] J. Hurst, J. Chestnutt, and A. Rizzi, “An actuator with physically variable stiffness for highly dynamic legged locomotion,” *Robotics and Automation, 2004. Proceedings. ICRA’04. 2004 IEEE International Conference on*, vol. 5.

[27] R. Van Ham, B. Vanderborght, M. Van Damme, B. Verrelst, and D. Lefeber, “MACCEPA, the mechanically adjustable compliance and controllable equilibrium position actuator: Design and implementation in a biped robot,” *Robotics and Autonomous Systems*, vol. 55, no. 10, pp. 761–768, 2007.

[28] J. McBean and C. Breazeal, “ Voice Coil Actuators for Human-Robot Interaction. IEEE,” *RSJ International Conference on Intelligent Robots and Systems (IROS04), Sendai, Japan*, 2004.

[29] W. Urbanczyk, T. Martynkien, M. Szpulak, G. Statkiewicz, J. Olszewski, G. Golojuch, J. Wojcik, P. Mergo, M. Makara, T. Nasilowski, *et al.*, “Photonic crystal fibers: new opportunities for sensing,” *Proceedings of SPIE*, vol. 6619, p. 66190G, 2007.

[30] C. Breazeal, “Regulation and Entrainment in Human-Robot Interaction,” *Experimental Robotics VII*, 2001.

[31] D. Miyauchi, A. Sakurai, A. Nakamura, and Y. Kuno, “Human-robot eye contact through observations and actions,” *Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004.*, vol. 4, 2004.

[32] C. Izard, “Four Systems for Emotion Activation: Cognitive and Noncognitive Processes.” *Psychological Review*, vol. 100, no. 1, pp. 68–90, 1993.

[33] J. Posner, J. Russell, and B. Peterson, “The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology,” *Development and Psychopathology*, vol. 17, no. 03, pp. 715–734, 2005.

[34] C. Breazeal, *Designing Sociable Robots*. Mit Pr, 2002.

[35] S. Sosnowski, A. Bittermann, K. Kuhnlenz, and M. Buss, “Design and Evaluation of Emotion-Display EDDIE,” *Intelligent Robots and Systems, 2006 IEEE/RSJ International Conference on*, pp. 3113–3118, 2006.

Children's responses and opinion on three bots that motivate, educate and play

Rosemarijn Looije, Mark A. Neerincx, Vincent de Lange

Abstract— Social robots may help children in their daily health-care related activities, such as adherence to diet and exercises of diabetics. Based on a domain and literature study, we specified three support roles with corresponding bot behaviors: motivator, educator and buddy. These behaviors, such as showing attentiveness, could be implemented well in a physical character (the iCat robot), somewhat less well in a virtual character, and least well in a text interface. Twenty—eight to nine years old—children participated in a controlled experiment to evaluate the bots. They proved to value the support roles positively, in particular the buddy role. Objective and subjective data showed that they highly appreciated both the physical and virtual characters (more than the text interface). Furthermore, children proved to interact faster with the character than with the text interface. There is a clear added value of robots compared to conventional text interfaces.

***Index Terms*—Human-robot interaction, Physical agents, Domotic agents and applications**

I. INTRODUCTION

INFORMATION and communication technology (ICT) in home, school and health settings has changed dramatically in the last two decades. For example for education it has been changing from one computer in a class that is hardly used, to computer usage by every school subject and the requirement to do homework on the computer. This use can be extended from homework tasks for school to physical exercise. These physical exercises might help to counter the increasing number of children suffering from obesity and diabetes. ICT technologies can thus aid in doing exercises [1-6], giving social support [7,8], and helping with lifestyle change [9-12]. Research on persuasive technology [13] and affective computing [14] provides (partial) solutions, e.g. for the realization of social behavior, such as social talk and turn-taking [2-5], and of empathic behavior, such as attentiveness and giving compliments [7,8],[9],[6,10,12]. This research comprises supporting technologies that are more conventional text-based [6,9,12], and more innovative character-based virtual [1,2] or physical [3-5,7,8,10] "robots". The media equation [15] states that technology is higher appreciated when it exposes social behavior and is physically present. Consequently, one would expect that physical characters are appreciated more than virtual characters and text interfaces. This is confirmed in research comparing virtual with physical characters, as all results are in favor of the physical character [4,16-19]. In comparison with adults, children react to, and interact with, physical characters differently. Tanaka [20] found that children – after 27 lessons - interact with a physical character as if it was a peer instead of a toy. This can be caused by their tendency to heavily anthropomorphize the character. Draper [21] conducted research towards physical characters in the education of children. This research showed that a teacher teaches best, but that a physical character is better than a sound-tape with the lesson.

The paragraph above summarizes some research on persuasive technology, affective computing, virtual and physical characters. However, more research is needed for better understanding of the added value of robots compared to conventional text interfaces. First, there is a need for further theoretical foundation from psychology, pedagogy, persuasive technology and affective computing, to improve the development of a motivating and educating social companion. Second, there is a need for further empirical foundation, in which the different user interfaces are being evaluated in a comparative experiment with children.

In this paper, we address this by comparing a text interface, a virtual and a physical character that all implement the roles of educator, motivator, and (game)buddy as far as their dialogue and appearance characteristics allow for. Our general hypothesis is that a physical character is better at fulfilling these roles than a text interface and virtual character. We focus on the user experience [22]: how the children response to, and enjoy the interaction with the different interfaces.

II. DESIGN OF THREE BOTS FOR YOUNG DIABETICS

We chose the iCat from Philips (**Fig. 1**), in both physical and virtual form, to implement the behaviors for the concerning roles. This character was previously used in an experiment with older adults [10,23]. During this experiment, participants evaluated five different interfaces: a text interface, a social and non-social virtual character, and a social and non-social physical character. User preference was measured for the different assistants on several factors, such as empathy,

Rosemarijn Looije is with TNO Defence, Security and Safety
E-mail: rosemarijn.looije@tno.nl
Mark. A. Neerincx is with Delft University of Technology and TNO Defence, Security and SafetyE-mail: mark.neerincx@tno.nl
Vincent de Lange is with vCreativo
E-mail : vincent.de.lange@vcreativo.tk

trust, and acceptance. The results indicated that socially intelligent characters are rated more empathetic than a text interface and a non social character. Moreover, the virtual character was appreciated more than the physical character

Fig. 1. A happy, angry and surprised iCat

both on the trustworthiness and the empathy dimensions [10]. Notwithstanding the positive results for the virtual character, half of the users indicated that they preferred the text interface while the other half preferred a social character. A possible explanation could be the anxiety that older adults have towards characters [24].

A. Media equation

People have the tendency to socialize information and communication technology [15], this is called the media equation. The more a device supports this tendency, the more people will like to use the technology. Furthermore, a physical character will have a greater social facilitation effect [4,16-18] (i.e. people tend to perform simple tasks better in the presence of others [25]) than a virtual character [16]. Both the tendency to like social devices and the social facilitation effect support the idea that a social physical character is preferred as a personal assistant. Therefore, we distinguish three bots in this study:

- Conventional text
- Virtual robot (virtual iCat)
- Physical robot (physical iCat)

B. Design of a prototype for children

The social characters and text interface developed for adults were taken as a starting point for the design of the prototype for children. The existing prototype was adapted for the use by children and made more automatic. We had to adapt the prototype because children ask for a different approach of both the design as well as the evaluation of the interface. During the design phase, special attention should be given to the different interests and cognitive abilities that children have in comparison with adults, which influence their interaction with the computer [26]. We looked specifically at cognitive, physical, and affective characteristics of children in the age group of 8-9. Children of this age are linguistically skilled and start performing several tasks independently. An example is diabetes where children start administering insulin and counting carbohydrates themselves.

Relating the cognitive development of children, interfaces should be *visually oriented* with not too much text and, just as for adults, *immediate feedback* is needed to keep the interaction natural and non-irritating. In relation to the physical development, Chiason and Gutwin [26] propose that interfaces for children should be tangible, such as the physical iCats, and that interfaces need not be cuddly in order to be engaging. Finally, research in affective computing shows that children like to have the possibility to be in control of the interaction with technology and that children stay engaged and motivated by providing them with occasional *entertaining events* [26]. Engagement and motivation can be stimulated by challenging and fun games, e.g. implemented in a (game) buddy [6,27]. The (game) buddy ensures that users keep using the assistant, because it is fun [22].

In the evaluation phase, subjective measures are often used to get the opinion of the user about the tested interface. The opinion of children is important, because adults do not always understand what children want and why [28]. Doing a survey with young children is not easy. The children should be able to interpret all the questions correctly and make a considered choice between the answers. Another problem for the analysis is that children have the tendency to have extreme opinions on all the products they rate [28].

C. Diabetic children

In previous research a domain analysis of adults with diabetes was performed. We extended this analysis to the domain of children with diabetes, using diabetes as a case study. A diabetic nurse, play therapist, a patient who acquired diabetes on a young age and a game developer were interviewed. This analysis yielded insights in the differences and similarities between adults and children with diabetes and their computer technology usage. Both adults and children have a need for an educator who teaches them more about diabetes, because chronically ill have little knowledge about their disease [12] and therefore do not understand why they have to comply with certain advices. Furthermore, there is a need for a buddy that is a companion in coping with the disease. In addition, children were in need of help for counting carbohydrates, and one that helps keeping track of time to take their medication in time. An important remark was that the use of the device should be fun and challenging to improve the engagement and motivation. Eventually, diabetic children could be one of the first “serious” users of the envisioned personal assistant. Eating, physical exercise, and their joint effect on energy consumption are important issues for such children, and, therefore, ‘core’ elements for our study on robot assistance.

III. Design of three roles for the bots

Based on the knowledge we gathered about diabetic children and their needs, a scenario was developed that includes personal assistance. Based on the scenario we chose three roles to be implemented in the prototype: educator, motivator, and game buddy. An extra advantage of implementing the motivator and educator roles is that the results can be compared to the motivator and educator role in the experiment for a personal assistant for older adults [10]. That experiment showed that these roles are appreciated when implemented in a social robot. We implemented the roles in the same three bots as in [10]: a chatbot, a virtual, and a physical robot. In contrast to the chatbot, the robots have the possibility to express facial and voice emotions.

A. *Motivator*

Both the motivator and educator are based on the Motivational interviewing theory, which by means of questions tries to facilitate increase in knowledge on persons' behavior *and* disease – in our case diabetes - thereby increasing the motivation to change. A therapist who can apply motivational interviewing successfully should be: empathetic [29] and trustworthy [30]. Motivational Interviewing is successfully applied in a text-based personal assistant, the HealthBuddy®, for chronically ill [9,11]. We divide the properties of motivational interviewing into two roles, the motivator and educator role. The motivator role implements the properties that are linked to *how* things are said and done while the educator role focuses on *what* is said and done. This means that the motivator role looks at ways to make the assistant appear empathetic and trustworthy.

To make the assistant look empathetic we could find some skills with related behaviors to implement. We implemented three behaviors for three skills; reflective listening, positive regard, and attentiveness. The *virtual* and *physical iCat* are able to implement behaviors for all three skills, while the *text interface* can only implement behaviors for the positive regard skill.

Reflective listening behaviors that are implemented are: reacting positive or negative according to the event and asking questions when something is not understood. The behaviors that are implemented for positive regard are: give compliments when something is done correct and do not punish if something is done wrong. The behaviors for the last skill, attentiveness, are: look at the user, have an active listening expression, and sometimes nod.

It is very difficult to find behaviors that make an assistant look trustworthy; trust in an application is something that comes in time, but it can be stimulated. To enable trust, the dialog, mainly the form and content, can be made acceptable for the user. This can be done for example by taking the vocabulary of the user in account. Another way to receive trust, that the play therapist proposed, is to make the user comfortable (e.g. let the user play a game).

B. *Educator*

Motivational interviewing tries to increase the knowledge of a patient by educating the user. We implemented this in a quiz form that used educational videos on nutrition and/or exercise each followed by a multiple choice quiz question about the video to increase the knowledge of the user about the subject. The educator uses behaviors from the motivator to appear empathetic and trustworthy. It listens to what the user says, is happy when the user answers a question correctly, and just gives the reason for the correct answer when the answer is incorrect. The educator behavior was the same for the *physical* and *virtual iCat* and for the *text interface*.

C. *Game Buddy*

The game buddy role was chosen, because an assistant for children would definitely need a fun activity. Children need to stay engaged, and alongside of the serious tasks a personal assistant can offer them, some entertaining functionality is necessary.

A first prerequisite for the game buddy was to offer a familiar two player game that was not too difficult, did not take long, and was fun for a little while. In previous research with the game of tic-tac-toe [31], children found it fun to play it with the iCat. Therefore, we decided to use tic-tac-toe in our prototype. Furthermore we based the personality of the game buddy on the personality that was preferred in the research of Verhaegh [31]: moderate expressive.

There was an algorithm that made sure that the level of the game was adapted to the user so that it became harder if the user won and easier if the user lost. The outcome of the previous game was stored in a user profile. We tried to keep the game challenging in this way.

The personal assistant in the game buddy role was empathetic (using the motivator behaviors, which were different for the *robots* and the *text interface*, see section III.A) towards the user; it gave compliments and was not over enthusiastic if it won a game. The personal assistant gave comments on the game; compliments ("nice move"), neutral remark ("now we are equal"), and congratulating remarks ("congratulations you won"). The comments were given taking three factors into account: Who made the last move, whether the situation is advantageous for the user, and if the game is in an end state.

Besides being complimentary the assistant was also attentive in the way that it asked the user if he/she would like to start, which symbol he/she preferred to use, and it looked at the game board when the attention of the user was there. Furthermore the assistant did not cheat, and left the user in control.

IV. MULTI AGENT STRUCTURE

We implemented the prototype with the use of distributed agents that were in compliance with the FIPA standards [32]. The different roles were all implemented in their own agent so that the structure was modular. The modularity makes it possible to extend or adapt the system without changing the whole system. Furthermore, the use of agents makes the whole system easy distributable. Fig. 3 gives an overview of the implemented agents. The agents are implemented in JADE.net [33] with the use of C#, because the communication framework was already implemented in C#.

The three different roles are implemented in different agents. The motivator is implemented in the dialogue agent (which is the central agent), deciding when what text and what expression should be used. The dialogue agent also poses the quiz questions and handles the answers. Secondly the tic-tac-toe agent implements the game buddy that decides when to do which move. Finally the quiz agent implements the educator role by starting up movies. The touch screen agent displays the movie and tic-tac-toe and sends the move of the user in tic-tac-toe back to the tic-tac-toe agent.

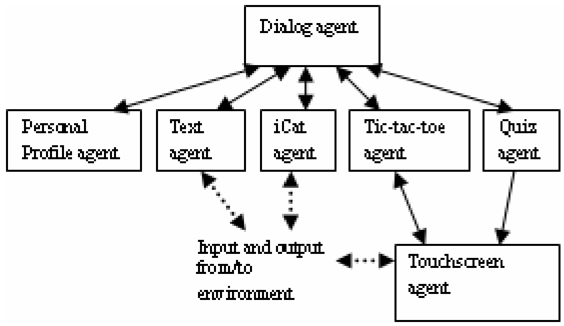

Fig. 3. Agent structure

TABLE 2 QUESTIONS REGARDING FUN

		Question
Expectation	**ifx1**	How nice do you think working with the robot/chatbot is going to be?
	ifx2	What did you think of working with the robot/chatbot?
Engagement	**ife1**	Would you like to use the robot/chatbot again?
	ife2	Would you like to play another game with the robot/chatbot some time?
	ife3	Would you like to play another quiz with the robot/chatbot some time?
	ife4	Would you like to talk some more with the robot/chatbot some time?

The text, touch-screen, and iCat agent receive and send information from and to the environment. The text agent represents the text interface, and the iCat agent represents the iCat. Within the iCat agent, there is a module that handles the text input from the speech recognition that is performed by the experimenter. The last agent is the personal profile agent that holds information about the user, such as age, gender, lost and won games. This information can be used to adapt dialogue, game, and quiz.

A. *Wizard of Oz*

The participants thought they were using a completely autonomous assistant, but the experimenter/wizard simulated the speech-to-text. The agents, text interface, and iCat were implemented in a way that the whole interaction between participant and personal assistant was autonomous (i.e., only the speech recognition was simulated via a person in another room, the so-called Wizard of Oz).

V. EVALUATION

The three bots; chatbot, virtual robot, and physical robot, were implemented with the use of the predetermined roles and agents. After which they were evaluated. In this evaluation we tested if the participants thought of the bots as being empathetic, trustworthy, and fun, amongst others. Furthermore, we objectively measured positive and negative utterances and time spent at the interaction with the robot.

TABLE 1 POSITIVE AND NEGATIVE UTTERANCES THAT WERE COUNTED

Positive	Property	Negative	Property
Smiles	Mouth angles directing upwards	Frowns	Lowering the eyebrows
Laughing	smile with unveiling of teeth		
Concentration signs	fingers in mouth, tongue out	Signs of boredom	Ear playing, fiddling
Excitable bouncing	Moving (slightly) back and forth in the vertical direction	Shrugs	Moving shoulders quickly up and downwards
Positive vocalization	Exclamations such as “cool”, “I like ..”, if made not directly towards the interface	Negative vocal expression	Exclamations such as “boring”, “don’t like”, if made not directly towards the interface

Based on literature about social actors and previous research our hypotheses were:
(H1) The robots will be evaluated as more empathetic than the chatbot.
(H2) Children will trust the physical robot most and the chatbot least.
(H3) The physical robot is most attractive.
(H4) The interaction will be faster with the robots.

Fig. 2. Experimental setting

A. *Method*

Participants: Twenty-four non-diabetic children took part in the experiment, that lasted around 1 hour and quarter, for which they were rewarded with a book token. The data of twenty children was usable (due to incompleteness and a child with a neuro-developmental disorder). The twenty children were all third graders (i.e., fifth group of the primary school in the Netherlands), aged 8-9 (M age = 8.40, SD = 0.50).

Setting: The experiment was conducted in a room that resembled a living room. There was a table, on which touch-screen and iCat stood, or instead of the iCat a keyboard and computer screen stood (Fig. 2).

Experimental design: A within subject design was used for iCat vs. text interface, while there was a between subject design for physical vs. virtual iCat. This meant that all children used the text interface and the iCat for which the order of use was counterbalanced. Furthermore the children that used the virtual iCat did talk and played a game with the physical iCat at the end to get some additional information on their preferences for a virtual or physical robot.

Measures:

We limited the amount of questions to a minimum to keep the experimentation time reasonable.

Fun: The six questions regarding subjective fun (Table 2) were asked with the use of a smiley-o-meter [28], which is a five point Likert scale that uses smileys to represent the answers. We did also count the number of negative utterances and number of positive utterances and subtracted these from each other as a measure for observed fun. The utterances we counted are enumerated in Table 1.

Acceptance: Five different questions about acceptance were asked (Table 3). The questions were all posed on a five point Likert scale. We adapted the annotation of the scale to every question; An example of this is "Do you understand the robot" which has the scale "Never", "Sometimes", "Always".

Empathy: For empathy four questions were asked (Table 3), the questions were also posed on a five point Likert scale and posed in the same way as the acceptance questions.

Trust: Three questions for trust were asked (Table 3). The questions were posed on a five point Likert scale similar to that of the acceptance and empathy questions.

TABLE 3 QUESTIONS REGARDING ACCEPTANCE, EMPATHY, TRUST, AND HEALTH INTENTION

		Question
Acceptance	ia1	Would you like to have the robot/chatbot at home?
	ia2	Did you find it easy to work with the robot/chatbot?
	ia3	Do you understand the robot/chatbot?
	ia4	Which interface did you find easiest to use?
	ia5	Which interface did you prefer?
Empathy	ie1	Do you find the robot friendly?
	ie2	Do you think the robot understands you?
	ie3	Do you think the robot tells the truth?
	ie4	Do you find the robot is curious about you?
Trust	iv1	Do you think the robot tells the truth?
	iv2	Would you answer honestly to the robot's questions?
	iv3	Do you think the robot would tell your secrets to someone else?
Health Int.	hi1	How many times a day would you like to eat fruit?
	hi2	How many lollipops do you think you should be allowed to eat a day?

Efficiency: The efficiency was calculated using the time of interaction with the interface. Because the virtual iCat and the physical iCat condition require some extra time caused by the "speech recognition", this amount of time had to be subtracted. The subtraction of the speech recognition was done because in the future this will be done automatically and not by hand as was the case in this experiment. We calculated the efficiency by taking the total amount of interaction time minus the wizard time. This is around 6% of the total time.

Learning effect: The learning effect is related to the accurateness and completeness of the tasks. The effectiveness was therefore measured by the number of correctly answered quiz questions.

Health intention: Health Intention is interesting in relation with the motivational interviewing (change in lifestyle) approach we took. Therefore we asked questions about the attitude towards nutrition before the experiment and after the use of each assistant. The questions (Table 3) were based on the theory of Reasoned Action [34].

TABLE 4 REASONS WHY CHILDREN CHOSE AN INTERFACE (NR. & % OF CHILDREN)

Argument	iCat	Text
Talking & no typing	4 (20%)	
Talking	3 (15%)	
Difficult to understand (speech)		3 (15%)
Typing		2 (10%)
No typing	3 (15%)	
Difficulty reading	3 (15%)	
Other	2 (10%)	
Total	15 (75%)	5 (25%)

Procedure:

Participants were told they participated in an experiment to evaluate personal assistants for children. They would work with a number of interfaces and have to fill in some questionnaires on what they thought of the interfaces.

They used the bots subsequently. First they answered a question about their health intention. And before using an interface, they answered a question about expected fun. They were told that when they would hear a beep, the interaction would start. The interaction with the interface followed a structured dialog, which was led by the interface. In the interaction, questions were asked by the bots and the participants were expected to answer on those. It was structured, since we wanted to let the participants experience more or less the same interaction, in order to be able to compare the results. In each condition, the dialog followed the same structure, consisting of three parts or tasks that represented the three different roles: motivator, educator, gamebuddy. First the assistant introduced itself (talking task/motivator), then a video quiz was played with the children followed by a quiz question (video quiz task/educator) and finally one or two tic-tac-toe games were played (game task/gamebuddy). After the interaction children were asked the five remaining questions on the experienced fun and the questions about trust, health intention (two after the first interface and three after the second), perceived empathy and three of the acceptance questions (ia1-ia3). In the end the children were asked what kind of roles or applications they would use the iCat for and ia4-ia5.

VI. RESULTS

Fun: The question about the fun expectation (ifx1) resulted in a significant difference between the physical iCat (mean = 4.6 out of 5) and the text interface (mean = 4.0 out of 5)

TABLE 5 SIGNIFICANT RESULTS FOR EFFICIENCY (TIME ON TASK)

	Mean			
Task	iCat	Text	Test	Sign.
Talking (physical iCat)	27.6	60.6	One way MANOVA (F(1,8)=15.5)	p<0.01
Talking (virtual iCat)	23.9	56.0	One way MANOVA (F(1,9)=7.8)	p<0.02
Game (physical iCat)	122.6	187.3	One way MANOVA (F(1,8)=9.6)	p<0.01
Game (virtual iCat)	120.5	171.9	One way MANOVA (F(1,9)=11.7)	p<0.01
Total (physical iCat)	478.9	621.6	One way MANOVA (F(1,8)=6.6)	p<0.03
Total (virtual iCat)	462.3	584.2	One way MANOVA (F(1,9)=24.0)	p<0.001

(Mann-Whitney U (1,8)=20.5, Z=2.06, p<0.05). In addition, we compared the indicated value of fun per task within and between interfaces (ife2-4). The game with the physical iCat was valued significantly more fun (mean = 4.7 out of 5) than the quiz with the physical iCat (mean = 3.3 out of 5) (Sign test Z(1,8)=2.04, p<0.05). The same applied for the virtual iCat (4.8 vs. 4.0) (Sign test Z(1,9)=2.04, p<0.05) and the text interface (4.7 vs. 3.3) (Sign test Z(1,18)=2.41, p<0.02). The game of the physical iCat was also experienced as more fun than the quiz of the text interface (4.7 vs. 3.4) (Sign test Z(1,8)=2.27, p<0.03). These results indicate that the game is considered more fun than the quiz.

The observed fun was measured by examining the result of the positive expression values minus the negative ones. In the talking task this gave significant differences between physical iCat (2.7) and virtual iCat (1.0) (Manova F(1,8) = 18.3) and between physical iCat (2.7) and text interface (0.0) (Manova F(1,8) = 7.0). When all expression values were taken together there was a significant divergence between physical iCat (10.9) and text interface (5.4) (Manova F(1,8) = 5.0). Another interesting measure is the total amount of fun utterances, which can be used to determine whether or not there are more positive utterances towards a particular interface. This measure provided two significant differences between both the virtual iCat (1.6 utterances) and the text (0.8 utterances) (Manova F(1,9)=7.0, p< 0.02) and between physical iCat (2.8 utterances) and text (Manova F(1.8)=8.7, p<0.001) So, children show more indicators of fun when talking with an iCat than with the text interface.

Acceptance: Both acceptance questions about the ease of use (ia4) and preference (ia5), asked at the end of the experiment, showed significant differences between the different interfaces. The iCats were found easier to use than the text interface (Chi-Square (1,19) = 5.0, df = 1 p<0.03). The physical and virtual robots were found easiest to use, 70% and 80%, respectively. Similar results were found when asked for their preference. About 70% favored the iCats and 30% the text interface (Chi Square(1,19) = 4.1, df = 1 p<0.05) . The majority of the children stated the iCat to be more fun. The reasons they gave are summarized in Table 4. Children who performed their tasks with the virtual iCat were also given the opportunity to use the physical iCat. These children were also asked which of the three interfaces they preferred. The physical iCat appeared to be the most fun to work with. It was favored by 80% of the children, because it was real. Some additional comments were that its eyebrows and mouth could move. The remaining three questions regarding acceptance did not yield significant differences. All interfaces were rated high on acceptance: scoring 4.3, 4.5, and 4.4 out of 5 for the text interface, virtual iCat, and physical iCat, respectively. This indicates that all interfaces were very acceptable.

Empathy: All the three interfaces had high scores on the empathy questions ranging from 4.0 to 4.2 out of 5: 4.2 for the physical iCat, 4.0 for the virtual iCat, and 4.1 for the text interface. All interfaces were thus perceived as empathetic. There were no significant differences between the interfaces.

Trust: The children rated all three interfaces high on trust 4.1 out of 5 for the physical iCat and the text interface and 4.3 out of 5 for the virtual iCat. Again there were no significant differences between the interfaces.

Efficiency: For the efficiency of the interfaces we looked at the duration of the complete interaction. Both the efficiency of the virtual iCat and the physical iCat differed significantly from the text interface (Table 5). A comparison between the iCat and virtual iCat did not provide any significant difference.

Learning effect: About 85% of the children answered the question, posed before the movie containing the information, correctly. This affirms that the children were already knowledgeable on the topic. On average the children answered 8.3 out of 10 questions correct. Thus no learning effects could be found.

VII. CONCLUSION AND DISCUSSION

The experimental set-up, in which only the speech recognition was simulated, worked well, and the physical and virtual robots were highly appreciated. We realized bots that could have meaningful and pleasant dialogues with children for their three roles. The interaction with the robots was significantly faster than with the chatbot and the physical robot was most fun to interact with. The game buddy role was important for the engagement with the personal assistant of the children. In contrast with the experiment with older adults [10], no significant differences were found for empathy. This can be explained by the high ratings the children gave to all three interfaces ("ceiling effect"). So, the proposed type of support for personal healthcare was well-accepted by the children in general.

This study compared three interfaces with their "natural" dialogue styles: a text-based chat-bot with two speech-based robots. You could say that we compared text to speech. We argue that a text interface for the characters would have been unnatural, because their appearance strongly suggests they have the ability to speak and listen. Correspondingly, speech dialogues are uncommon for the graphical, direct manipulation displays (windows).

In the short term, no significant discrepancies were observed regarding motivation and education between the different personal assistants. Therefore, a long-term experiment should be conducted in which engagement will play a larger role, because children will have to keep using the personal assistant for a longer period of time. Long-term effects of artificial agents in healthcare interventions are discussed in e.g. Marsella, Lewis Johnson, Bore[35] (education about cancer), Bickmore and Picard [36] (motivating to exercise), and Brave, Nass, and Hutchinson [37](social support). These papers show the relevance of the educator, motivator and buddy roles for user support. The long term results suggest that virtual characters that exhibit affection are more enjoyable, more trustworthy, more supportive, and a better educator in comparison with no virtual character or a virtual character without affective abilities. Furthermore, learning results were better, and the participants were more willing to continue working with the social character. This literature focused only on adults. We would like to explore the long term effects on children and the effects of a physical character in comparison with a virtual character. In the healthcare domains we are looking into children with e.g. obesities, diabetes, and coeliac. These children should adapt their diet to stay healthy and are not allowed to eat the same as most children (i.e. a diabetic should keep track of his/her sugar intake). A buddy to cope with being different could be appreciated. Furthermore, the buddy could help educating them about their condition and motivate them to follow the physician's advice of the physician.

In the future the game buddy role should be extended to make it possible to play multiple games. Furthermore, the dialog agent should be able to handle more diverse interactions and preferably even conversations that were not anticipated by the programmer beforehand. As expected, the results showed that the quiz was valued as less fun than the game. Fun is very important to keep the children engaged, as we learned from the educational game developer during domain analysis. In the future, we would like to explore other educational methods that are perhaps more fun to use (this might eventually lead to a game educator).

In general, we can say that the children rated the interface properties high, which caused a small number of significant differences in the subjective measures. The objective measures also showed a preference for the robots, while their interaction was faster and exhibited more social behavior. They were excited about participating in the experiment and using the iCat. These results indicate that the iCat is an interface that attracts the attention and therefore can have positive effects on motivating and educating children while being a buddy, which is of importance when applying the robot in the healthcare domain. So, the motivator and educator roles that we developed are appropriate for both older adults (see [23]) and children, and the iCat is a good platform to implement and test such roles for both user groups.

Acknowledgment

We would like to thank all the children of "groep 5" from the Postiljon and their parents for their participation. Furthermore we thank IOP-MMI SenterNovem a program of the Dutch Ministry of Economics, for partially funding the SuperAssist project.

References

[1] Ruttkay, Z. M., Zwiers, J., Van Welbergen, H., & Reidsma, D. (2006). Towards a reactive virtual trainer. *Proceedings of IVA 2006*, pp. 292-303. Berlin: Springer-Verlag.

[2] Bickmore, T., Gruber, A., & Picard, R. (2004). Establishing the Computer-Patient Working Alliance in Automated Health Behavior Change Interventions. *Patient Education and Counseling*.

[3] Kidd, C. D. & Breazeal, C. (2006). Designing a Sociable Robot System for Weight Maintenance. *IEEE Consumer Communication and Networking Conference 2006*, Piscataway: IEEE.

[4] Goetz, J., Kiesler, S., & Powers, A. (2003). Matching robot appearance and behavior to tasks to improve human-robot cooperation. *IEEE Ro-Man 2003*, pp. 55-60. Piscataway: IEEE.

[5] Gockley, R. & Mataric, M. J. (2006). Encouraging physical therapy compliance with a hands-Off mobile robot. *ACM SIGCHI/SIGART Human-Robot Interaction 2006*, pp. 150-155. New York: ACM.

[6] Toscos, T., Faber, A., An, S., & Gandhi, M. P. (2006). Chick clique: persuasive technology to motivate teenage girls to exercise. *Conference on Human Factors in Computing Systems*, pp. 1873-1878. New York: ACM.

[7] Kidd, C. D., Taggart, W., & Turkle, S. (2006). A Sociable Robot to Encourage Social Interaction among the Elderly. *ICRA 2006*, pp. 3972-3976. Piscataway: IEEE.

[8] Kriglstein, S. & Wallner, G. (2005). HOMIE: an artificial companion for elderly people. *Conference on Human Factors in Computing Systems*, pp. 2094-2098. New York: ACM.

[9] Bigelow, J. H., Cretin, S., Solomon, M., Wu, S. Y., Cherry, J. C., Cobb, H. et al. (2000). *Patient Compliance with and Attitudes Towards Health Buddy*. Santa Monica: RAND Corporation.

[10] Looije, R., Cnossen, F., & Neerincx, M. A. (2006). Incorporating guidelines for health assistance into a socially intelligent robot. *Ro-man, 2006*, pp. 515-520. Piscataway: IEEE.

[11] van Dijken, G. D., Niesink, I. A., & Schrijvers, A. J. P. (2005). *Telehealth in de Verenigde Staten* Utrecht: U.M.C.

[12] Blanson-Henkemans, O., Fisk, A. D., Lindenberg, J., van der Mast, C. A. P. G., Neerincx, M. A., & Rogers, W. A. (2007). A Computer Assistant for Supervision of Self-Care and Improvement of Health Literacy. *115th Annual APA Convention*, Washington, DC: APA.

[13] Fogg, B. J. (2002). *Persuasive Technology: using computers to change what we think and do*. San Francisco, CA: Morgan Kaufmann.
[14] Picard, R. W. (1997). *Affective Computing*. Cambridge, MA: MIT Press.
[15] Reeves, B. & Nass, C. (1996). *The media equation: how people treat computers, television, and new media like real people and places*. New York, NY: Cambridge University Press.
[16] Bartneck, C. (2003). Interacting with an embodied emotional character. *Proceedings of the 2003 international conference on Designing pleasurable products and interfaces*, pp. 55-60. New York: ACM.
[17] Kidd, C. D. (2003). *Sociable Robots: The Role of Presence and Task in Human-Robot Interaction.* Master Thesis MIT, Boston.
[18] Wainer, J., Feil-Seifer, D., Shell, D., & Mataric, M. (2006). The role of physical embodiment in human-robot interaction. *Ro-man, 2006*, pp. 117-122. Piscataway: IEEE.
[19] Powers, A., Kiesler, S., Fussell, S., & Torrey, C. (2007). Comparing a computer agent with a humanoid robot. *HRI 2007*, pp. 145-152. New York: ACM.
[20] Tanaka, F., Cicourel, A., & Movellan, J. R. (2007). Socialization between toddlers and robots at an early childhood education center. *Proceedings of the National Academy of Sciences,* 17954-17958.
[21] Draper, T. W. & Clayton, W. W. (1992). Using a personal robot to teach young children. *Journal of Genetic Psychology, 153,* 269-273.
[22] Preece, J., Rogers, Y., & Sharp, H. (2002). *Interaction design: beyond human-computer interaction*. Hoboken: Wiley.
[23] Looije, R. (2006). *iCat for you: a comparison between different personal e-health assistants.* Master Thesis Rijksuniversiteit Groningen, Groningen.
[24] Scopelliti, M., Giuliani, M. V., & Fornara, F. (2005). Robots in a domestic setting: a psychological approach. *Universal Access in the Information Society, 4,* 146-155.
[25] Triplett, N. (1898). The Dynamogenic Factors in Pacemaking and Competition. *The American Journal of Psychology, 9,* 507-533.
[26] Chiasson, S. & Gutwin, C. (2005). Design Principles for Children's Technology. *interfaces, 7,* 28.
[27] Kahn Jr, P. H., Friedman, B., & Hagman, J. (2002). I care about him as a pal: Conceptions of robotic pets in online aibo discussion forums. *CHI 2002 Extended Abstracts of the Conference on Human Factors in Computing Systems*, pp. 632-633. New York: ACM.
[28] Read, J. C. & MacFarlane, S. (2006). Using the fun toolkit and other survey methods to gather opinions in child computer interaction. *Proceedings of the 2006 conference on Interaction Design And Children*, pp. 81-88. New York: ACM.
[29] Rogers, C. R. (1951). *Client-centered Therapy: Its Current Practice, Implications, and Theory*. Houghton Mifflin Boston.
[30] Neerincx, M. A. & Streefkerk, J. W. (2003). Interacting in Desktop and Mobile Context: Emotion, Trust and Task Performance. *Proceedings of the first European Symposium on Ambient Intelligence (EUSAI), Eindhoven, The Netherlands.*, Springer-Verlag.
[31] Verhaegh, J. (2004). *iCat, an intelligent user interface robot with character. An evaluation of robot personality and personality preference in child-robot interaction.* Master Thesis University of Maastricht.
[32] FIPA (2008). FIPA agent communication. http://www.fipa.org/repository/aclspecs.html [On-line].
[33] Jade (2008). Jade. http://jade.tilab.com/ [On-line].
[34] Fishbein, M. & Ajzen, I. (1975). *Belief, attitude, intention, and behavior*. Addison-Wesley Reading, Massachusetts.
[35] Marsella, S. C., Johnson, W. L., & LaBore, C. M. (2003). Interactive Pedagogical Drama for Health Interventions. pp. 341-348. Amsterdam: IOS Press.
[36] Bickmore, T. W. & Picard, R. W. (2005). Establishing and maintaining long-term human-computer relationships. *12*, *ACM Transactions on Computer-Human Interaction (TOCHI)*, pp. 293-327. New York: ACM.
[37] Brave, S., Nass, C., & Hutchinson, K. (2005). Computers that care: investigating the effects of orientation of emotion exhibited by an embodied computer agent. *International Journal of Human-Computer Studies*, *62,* 161-178.

Domestic Applications for Social Robots - an online survey on the influence of appearance and capabilities

Manja Lohse, Frank Hegel, and Britta Wrede

Abstract—**Can you imagine a useful task you would like a social robot to perform for you? This paper presents an internet survey where participants were asked this question to identify applications for social robots. The applications mentioned by the participants are based on the appearance of four social robots (AIBO, iCat, BIRON, and BARTHOC) and the information they received about their basic capabilities. It was found that AIBO and iCat seem to be suitable for domestic applications whereas suggested applications for the more functional mobile robot BIRON are situated also in public environments. The anthropomorphic robot BARTHOC mainly seems to be appropriate for public usage. The paper tries to explain how the appearance and the capabilities of the robot influence what applications are ascribed to them. Moreover, it is shown what role domestic robots play in the field of social robotics and how they relate to public robots.**

***Index Terms*— robot applications, domestic robots.**

I. Introduction

The general objective of the social robotics research is to design robots that engage in social scenarios which are compelling and familiar to humans. Thus, the robots have to provide a social communicative functionality that is natural and intuitive. This can be supported if appearance and functionality fit the robots' tasks [1] and robots are as self-explaining as possible. Keeping all this in mind is especially important when designing robots to interact with naïve users in domestic environments.

Today, mainly simple domestic toy robots (e.g. Lego Mindstorms [2], Ugobe's Pleo [3], WowWee's Robosapien [4]) and floor cleaning robots (e.g. Roomba [5]) are sold off-the-shelf. Also some public applications have been developed (e.g. receptionist robot [6], museum guide robots [7], [8]). All these applications have mainly been invented by researchers. They demonstrate the technical functionalities available so far. However, we argue that potential users should be included in the process of finding and developing new scenarios. Altogether, there is a lack of application scenarios and systematic classifications. Researchers do not know what tasks robots should be able to complete for potential users although knowing consumers' opinions is important for designing useful applications. Moreover, no guidelines how robots should look like to be suitable for a certain task exist.

In the work presented here this lack of systematic research on robot applications is addressed. The paper introduces an internet survey with potential users that were asked to suggest relevant applications for the four robots *AIBO* (Sony), *iCat* (Philips) (both consumer products), *BIRON* (Bielefeld University) and *BARTHOC* (Bielefeld University) (both research robots). It is evaluated which of the robots – due to their appearance and functionalities – are rated as being suitable for domestic applications and which applications participants would use themselves.

Section II gives an overview of research on social robots, applications and the importance of the appearance of robots. Section III describes the four robots chosen for the survey based on their appearance. Section IV introduces the method stressing why we decided to conduct an online study. The results of the survey and a discussion are presented in Sections V to VIII. Finally, in Section IX we conclude what we learned from the study.

II. Related Work

Since robots have started to move around our physical and social spaces, their role and our dealings with them have changed significantly. Therefore, a whole research area called social robotics has developed. Social robots are physical entities embodied in a complex, dynamic, and social environment, sufficiently empowered to behave in a manner conducive to their own goals and those of their community [9]. They especially serve as an interface between man and machine [10]. So far, many social robots have been built mainly to demonstrate technical skills within very specific scenarios (e.g. [11]). Even though such robots enable huge advances in research, they are far from being off-the-shelf consumer products. Yet, despite the current situation social robots should become everyday life applications in order to take advantage of their ability to communicate with almost everybody.

When building social robots it has to be strived for a balance between the expectations people have, the capabilities of the machine, and as a result the way to communicate the robot to do specific tasks. Thus, a social robot should be able to communicate with us, understand and even relate to us, in a personal way. It should be capable to understand humans and

Manja Lohse, Frank Hegel, and Britta Wrede are with the Applied Computer Science Group, Bielefeld University, Germany.
E-mail: mlohse; fhegel; bwrede@techfak.uni-bielefeld.de

itself in social terms. In turn, human beings should be able to understand the robot in the same social terms, to relate to it and to empathize with it [12].

Bartneck and Forlizzi [13] have created a guideline, which outlines the components a social robot should comprise: First, the form should match the expectations of a user. Second, the robot should communicate verbally and nonverbally with all available modalities. And third, the robot has to be able to take human social norms into consideration.

One domain in which these capabilities are especially useful is domestic applications. In fact, these become increasingly important [14], [15]. "Domestic" describes the place of usage. This place is the home of the user. It is special since the user decides who enters it and creates the rules that apply within it. The robot should be able to perform its specific tasks within this limited domain. Another challenge for domestic applications is that the robot has to learn new tasks within its environment. Learning poses a special problem as the typical user is naïve in programming and unaware of mechanical and control issues. Therefore, new learning methods will have to be developed. Probably the most intuitive and appropriate way to instruct the domestic robots is to interact with it as naturally as humans do in their daily life which again requires robots that, in fact, are social.

A. Applications of Domestic and Social Robots

Up to now, there are only few commercial applications for both social and domestic robots (see Section I). Developing useful applications for such robots seems to be a challenging task.

Christensen [16] postulates three potential commercial application categories for domestic robots: *Entertainment*, *Everyday Tasks*, and *Assistance* to elderly and handicapped people. These can be differentiated by what he calls a performance metric which specifies requirements for the applications. The performance metric of entertainment applications is forgiving, i.e. the robot is not really required to perform specific tasks – as long as the robot performs interesting interaction sequences the users will in general be satisfied. The challenge for this kind of applications is to provide an open-ended repertoire of interesting interactions to guarantee that the robot does not become boring for the user. In contrast, the performance metric of everyday home tasks for domestic robots is well defined while the individual environment where they will have to fulfil these tasks cannot be anticipated. According to Christensen [16], typical examples for such applications are vacuum cleaning, fetch-and-carry tasks, ironing clothes, window cleaning etc. Finally, the category of assistance applications for elderly and handicapped people is mainly motivated by the demographic profiles of occidental societies. The performance metric of these applications is a high degree of flexibility and an easy instruction how to interact with the robots.

Another approach to classify the field has been presented by Fong et al. [17]. In a first survey of socially interactive robotics the authors have mentioned several application fields: social robots as test subjects for research on communication and human development theory, as short-term and long-term service assistants in public and private life, as toys and entertainment devices, for therapy, for research on anthropomorphism, and last but not least in the field of education. These application fields are based on the applications that existed when the paper was written.

Also Ljungblad [18] reported on a workshop held with the goal to conceive potential applications. As a result three interesting scenarios were selected: self-organizing robot plants, robots as travel companions, and amusement park guide robots. These scenarios are rather based on a brainstorming than on a scientific approach to the topic.

Kaplan chose a different approach [19]. By asking himself what would actually make social robots valuable as everyday objects, he came to the conclusion that the value of the objects has to meet the needs of the users. Thus, the robot has to find its place in human life by adding value to it in terms of short- or long-term usage. Based on this hypothesis concrete applications still have to be derived.

Obviously, so far, no single exhaustive categorization of applications for social robots exists even though the question what the systems can be used for has been addressed from different viewpoints. However, it is noticeable that most of these approaches have domestic applications in mind.

B. Appearance of Robots

Appearance has a major influence on the assumptions people have about applications and functionalities, i.e. behaviours of robots [1]. Current research states that the appearance has to support the correct estimation of the robot's real competencies by the user. The better the user's estimation the less will she be disappointed during the interaction with the robot [19].

In this context, the embodiment of a robot plays a major role. Bartneck [20] has found a facilitation effect in his study with the emotional robot eMuu. Participants acquired a higher score in a negotiation game and they put more effort into the negotiation when they interacted with the embodied robot character instead of the screen character. This may be due to the feeling of social presence [21].

Among very few approaches, Fong et al. [17] define four broad categories of social robots with respect to their appearance: anthropomorphic, zoomorphic, caricatured, and functionally designed robots. An anthropomorphic appearance is recommended to support a meaningful interaction with users [10], [12] because many aspects of nonverbal communication are only understandable if expressed in similarity to a human-like body. Zoomorphic robots are intended to look like their animal counterparts to support the idea that an observer expects the robot to behave like an animal. In some cases this might be helpful to communicate the functional limitations of a robot. For example, a dog is partly able to understand aspects of human speech but makes many mistakes. This mirrors the quality of speech recognition software [20]. Robots with a caricatured appearance are mainly designed both to not elicit any expectations based on familiarity and to focus on very specific attributes like mouth or eyes. Finally, functional robots are designed in a technical/functional manner to illustrate their ultimate functions. This corresponds to the famous claim by Sullivan [22] that form ever follows function. Hence, the designer expects that the user is able to understand the capabilities of the functionally designed robot by looking at its features. For instance, a camera mediates the feature of 'seeing'.

Especially the design of a robot's head is an important issue within human-robot interaction (HRI), because it has been shown that most non-verbal cues are mediated through the face [23]. Without a face the robot is perceived as being anonymous [24]. The physiognomy of a robot changes the perception of its human-likeness, knowledge, and sociability. Therefore, people avoid robots behaving or looking negatively and prefer to interact with positive robots [25]. Furthermore, an expressive face indicating attention [26] and imitating the face of a user [27] makes a robot more compelling to interact with. Consequently, Duffy [9] argues that a robot has to have a certain degree of human-like attributes for meaningful social interaction.

A certain degree of human-like attributes also belongs to the three aspects DiSalvo et al. [28] propose regarding the appearance which should be taken into account when designing robot heads: the robot should have a certain amount of robot-ness to stress the robot's machine capabilities and avoid false expectations of its emotional capabilities; it should have an amount of human-ness to make the user feel comfortable; and finally it should have a certain amount of product-ness that the robot is seen as an appliance.

These aspects are consistent with the matching-hypothesis [1]. This hypothesis claims that appearance and social behaviour of a robot should match the seriousness of the task, because a robot that confirms the expectations people have increases the sense of the robot's compatibility with its task. Especially within the field of domestic applications this seems to be an important aspect in the design of satisfying interactions with naïve users.

Another essential aspect concerning robot design is in what sense users expect the robot to behave like a human and when they attribute human-like qualities to the machine. This phenomenon, which is called anthropomorphism, is in the focus of the following sub-section.

C. Anthropomorphism and the Uncanny Valley

It has been shown that the more human-like the appearance of a robot is the more people attribute intentions to the robot within a classical *Prison Dilemma Game* task [29]. This automatic attribution of human-like qualities is called anthropomorphism. More general, anthropomorphism entails attributing human-like properties, characteristics, or mental states to real or imagined non-human agents and objects [30]. According to v. Foerster [31] humans anthropomorphize because it allows them to explain things they do not understand in terms that they do understand, and what they understand best is themselves as human beings. This is consistent with the familiarity thesis [32] which claims that humans understand the world based upon the mental model of it that they are most familiar with.

According to Guthrie [32], the best explanation for this phenomenon is the best-bet-thesis which is a cognitive and game-theoretic approach to explain anthropomorphism. He claims that humans anthropomorphize because in the face of chronic uncertainty about the nature of the world, guessing that things and events are human-like or have a human cause constitutes a good bet, i.e. if we are right we gain much, but if we are wrong we usually lose little.

Another theoretical approach to the topic is the Three-Factor-Theory of Anthropomorphism by Epley et al. [30]. They claim that the extent to which people anthropomorphize is mainly determined by one cognitive and two motivational factors: (a) *Elicited Agent Knowledge*: Knowledge about humans in general or self-knowledge serve as a basis for induction primarily because such knowledge is acquired earlier and is more richly detailed than knowledge about non-human agents or objects. The more human-like in appearance or motion the non-human agent is the more people are likely to use themselves as a source of induction. For example, robots are anthropomorphized more readily when given human-like faces and bodies [28] and hummingbirds suddenly appear more deliberate and thoughtful when their natural quickness is slowed to a human-like speed [33]. The second motivational factor is (b) *Effectance Motivation*: Effectance describes the need to interact effectively with one's environment. Attributing human characteristics and motivations to non-human agents increases the ability to make sense of an agent's action and reduces uncertainty. Finally, (c) the *Sociality Motivation* describes the principal need and desire to establish social connections with humans. When people feel lack of social connection they anthropomorphize to a higher extent to satisfy their motivation to be together with others.

Up to now, the phenomenon of human-likeness seems to be the most considered aspect regarding the theory of anthropomorphism in the field of robotics. The hypothesis from anthropomorphism is that the more a robot resembles a human being in appearance the more people expect it to have human-like qualities [9]. Therefore, human-likeness is a special challenge to reach for some researchers (e.g. [34],[35]). Furthermore, the development of robots that closely resemble human beings can contribute to research on cognition [36].

One theory closely connected to anthropomorphism is the Uncanny Valley [37]. It represents how an object can be perceived as having enough human-like characteristics to evoke a constrained degree of empathy through one's ability to rationalize its actions and appearance. When the movements and the appearance are almost human-like but not entirely, there are too many expectations of the capabilities.

When these are not met the reaction of the observer is most certainly negative. In the end, the object becomes so human-like that it is effectively treated as a human being where it has re-established a balance between anticipated and actual function and appearance to a sufficient degree that works well [9].

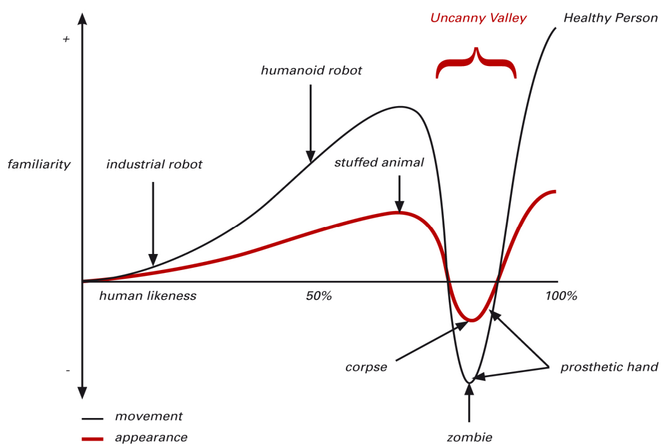

Fig. 1 Uncanny Valley

There are several studies whether the Uncanny Valley exists or not and how to consider the hypothesis when designing human-like robots. For example, [38] was able to find an Uncanny Valley within a study using morphings from a machine-like robot to an anthropomorphic robot. In a further psychological study based on four experiments [39] verified the Uncanny Valley caused by the human visual system. They found that in daily life people rarely confuse artificial faces with real human faces – people do not ask a mannequin in a store for directions to a train station. This suggests that the human visual system has a specific sensitivity to the degree of human-likeness. Therefore the visual system is highly sensitive to abnormalities. In the study it has further been shown that a high degree of abnormality elicits unpleasantness within the tested subjects only if entities exhibit a high degree of realism. For example, an increased eye size within an artificial character did not elicit any unpleasantness, but the same eye size within a human face was sensed as abnormal and therewith unpleasant. Thus, improving the degree of realism of robots without removing abnormal features may lead to an exaggeration of the observers' unpleasant impressions of the artificial faces.

No matter whether the Uncanny Valley exists or not, the theory may serve as a design guideline to not disappoint the expectations of the user. Also it is assumed that within the study presented here the discussed aspects of the robot's appearance have an effect on the applications people derived by seeing short videos of four different robots.

III. Description of Robot Systems

Since the appearance and functionalities of the robot are in the centre of the survey presented here, this section gives a technical overview of the robot platforms displayed. For each robot the applications and scenarios intended by the developers are mentioned.

A. AIBO (Sony)

The AIBO Robot ERS-7 is presented in Fig. 4. The design of AIBO is dog-like. The robot uses its four feet to move in its environment. With the acceleration sensors on-board it is able to balance its body. AIBO has – considering its feet, head, ears, and tail – altogether 20 joints (degrees of freedom) which enable the robot to perform dog-like moves.

AIBO has a set of sensors on the head, the back, the chin, and the paws which allow the robot to examine itself and its environment. With the help of the sensors the robot can sense touch. It can perceive sound using a pair of stereo microphones. Therefore, it can react to voice. By means of the colour camera and distance sensors AIBO can recognize colours, faces, and obstacles. It is able to communicate via sounds, a face display, and speech. On the website [39] Sony talks about the "AIBO Entertainment Robot" or even the "World's most popular entertainment robot". Obviously this is the main scenario the developers had in mind.

Moreover, the robot became well known due to the RoboCup league. The Sony website [39] provides more information about the robot.

B. iCat (Philips)

The iCat shown in Fig. 5 is a plug & play desktop user-interface robot developed by Philips Research (Eindhoven, the Netherlands). It does not have an on-board processor and is controlled by a PC via USB.

Fig. 2 Sony AIBO ERS-7

Even though the name suggests that the robot resembles a cat, its facial features are human-like. The iCat can create facial expressions by moving its eyebrows, eyelids, eyes, lips, and the head controlled by 13 RC servos and 2 DC motors. We belief that, according to the categorization of Fong et al. (see II A [17]), the iCat is a caricatured robot. Both the robot's mouth and eyes are big and, thus, emphasized compared to a human face.

Next to the facial features, the robot has LEDs located in the feet and the ears to communicate its mode of operation (e.g.

sleeping, awake, busy, or listening). The USB webcam in the nose of the iCat can be used for different computer vision tasks such as object and face recognition. Stereo microphones, a loudspeaker, and a soundcard placed in the feet of the robot are used for recording and playing sounds and speech. Thus, it is possible to use the platform for speech recognition tasks.

The iCat was primarily developed as a research platform and as an interface robot. As an interface the robot should be able to coordinate different electronic appliances like VCRs, or refrigerators in the intelligent home. Users can employ the Open Platform for Personal Robots (OPPR) software to program certain behaviours and applications. More details can be looked up in [41].

Fig. 3 Philips iCat

C. BARTHOC (Bielefeld University)

Fig. 6 gives an impression of the humanoid robot BARTHOC (Bielefeld Antropomorphic RoboT for Human-Oriented Communication). This robot is designed by Bielefeld University in cooperation with Mabotic.

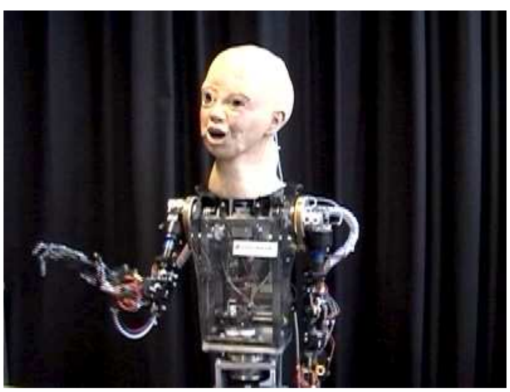

Fig. 4 BARTHOC ((Bielefeld Antropomorphic RoboT for Human-Oriented Communication)

BARTHOC consists of a mechatronic head with human-like features and two arms including hands. These components are mounted on a steel-frame backbone. Each arm has three joints similar to the human ones. The given degrees of freedom allow BARTHOC to perform human-like gestures. Actuators next to the upper lip and above the eyes simulate movements of lips and eyebrows which leads to basic human facial expressions. A camera is integrated in each eyeball for stereovision and microphones are currently placed on the shoulders. Additionally, a removable latex mask is available to equip the robot with different characters.

The robot is mainly used for research on human-like communication. One scenario the development of BARTHOC is based on is the receptionist scenario. In this scenario the robots is placed at the entrance of a building. It has to welcome people, to give general information, and to show the way to certain locations the visitor wants to got to by pointing and spoken instructions [42]. In [35] Hackel et al. describe BARTHOC in depth.

D. BIRON (Bielefeld University)

The robot BIRON (BIelefeld RObot companioN) (Fig. 7), developed at Bielefeld University, is based on an ActiveMedia™ Pioneer PeopleBot platform. BIRON is equipped with a pan-tilt colour camera at a height of 141 cm

Fig. 5 BIRON (BIelefeld RObot companioN)

for acquiring images of the upper body part of humans interacting with the mobile robot. A pair of AKG far-field microphones is located right below the touch screen display at a height of approximately 106 cm. Therefore, BIRON has the capability to localize speakers. Speech processing is supported by a headset the human interaction partner wears. A SICK laser range finder mounted at a height of 30 cm facing front measures distances within a scene. Since BIRON has wheels, it can follow a person and move in a room. Additionally, it is able to track people and pay attention selectively to humans looking at it.

The robot is strongly embedded in the so-called home tour scenario. In this scenario the user orders a robot for the home which is delivered with all necessary sensors and capabilities but has to learn about its new environment to be able to fulfil tasks like getting things from a certain location. Thus, the user has to show the robot around the house and to teach it rooms and objects. Since most users are inexperienced the interaction with the system has to be designed as intuitive and natural as possible. Further information about BIRON is given in [43].

The robots described here were chosen for two reasons. Firstly, they have different appearances according to the categories of Fong et al. (see A, [17]). BIRON' design is functional, AIBO's zoomorphic, BARTHOC's anthropomorphic and iCat's caricatured. Secondly, we wanted to evaluate how our own robots developed for research compare to commercial products like iCat and AIBO.

IV. METHOD

The research presented in this paper is based on an internet survey (see also [14]). This type of study was chosen for several reasons. First, it allows addressing people with different professional backgrounds, age and so forth. Thus, not only students or people related to a certain research field can be reached. Second, it can be conducted in a rather short time compared to live user trials. Third, the method supports the general idea of the study. People were asked to make a judgment based on the appearance of the robot in a video and a short description of its functionality. The internet provides a possibility to present this information without any further context and no technical problems with the robots could occur which would have influenced the subjects' answers. Therefore, we published the questionnaire in German on the website of an online laboratory. Some people were invited via private and professional mailing lists to participate. The mailing lists chosen were not connected to the robotics community.

Fig. 6 Screenshot of questionnaire containing: description of BARTHOC; questions: Would you use this robot?; What useful applications could you imagine for this robot in principal?, Which applications of this robot would you use yourself?

The questionnaire commenced with general questions about the participants like age, profession, and knowledge about computers. The second part of the survey focused on the question which robots people knew. The participants were allowed to enter multiple robots, but they should not put down more than five. Moreover, they were asked to indicate where they knew the robots from (media, seen in real life, interacted with the robot, develops the robot, owns the robot). Thereafter, the four robots (AIBO, iCat, BIRON, BARTHOC) were introduced in random order (on one site each) (Fig. 2).

All participants received very basic information about the functionalities of the robots:

- AIBO: can speak and understand speech, it is able to walk and to recognize objects and people.
- iCat: can speak and understand speech, perceive and recognize people and objects in its environment, it can show facial expressions and reacts to touch, iCat is not mobile.
- BIRON: can speak and understand speech, recognize people and objects in its environment, it can remember the position of objects and people in a room, drive, and display information on its screen.
- BARTHOC: can speak and understand speech, it can show facial expressions and gestures, perceive and recognize people and objects in its environment, BARTHOC is not mobile.

In the descriptions we tried to use similar wording whenever possible in order to avoid different interpretations of the functionalities by the participants.

Next to the descriptions videos were displayed, which showed few movements of each system to give an impression of the robots' appearance. The movements shown in the video clips that lasted about 3 seconds were:

- AIBO: getting up on its feet and turning
- iCat: turning the head and blinking
- BIRON: driving towards the viewer
- BARTHOC: turning head and upper body and lifting the arm

We chose to display these very short videos with a neutral background to avoid any context like concrete tasks the robot could perform in a certain environment. In contrast to simple pictures the videos give the viewer a better sense of the embodiment of the robot. They were replayed as long as the subjects stayed on the given site of the questionnaire. Sound was turned off since the study mainly concentrated on visual appearance. Voices and background noise might have changed the perception of the robot significantly.

After reading the descriptions and watching the videos, subjects were asked to suggest applications for all four robots. Since the study was exploratory this question was an open question and the subjects were free to write down as many items as they could think of. The question was divided into a general and a personal one (What useful applications could you imagine for this robot in principal?, Which applications of this robot would you use yourself?). Moreover, participants had to indicate for each robot whether they could imagine using it at all. At the end of the questionnaire they were asked to decide (Fig. 3):

- Which of the robots shown would you like to own?
- Which robot is most enjoyable to interact with?
- Which robot is most likeable?

Fig. 7 Screenshot of questionnaire containing questions: Which of the robots shown would you like to own?, Which robot is most enjoyable to interact with?, Which robot is most likeable?

V. RESULTS

In [15] we have proposed four dimensions for the distinction between the applications named by the participants of the web survey: (1) *public vs. private use*, (2) *intensity of interaction*, (3) *complexity of interaction model*, (4) and *functional vs. human-like appearance.* In the centre of the paper at hand is an in-depth analysis of the first dimension (*public vs. private use).* This dimension describes to which degree the robot is suitable for public or domestic use. Starting point of the analysis is the matching hypothesis (see II B, [1]). It is assumed that if the appearance has to match the task of a robot, viewers will also suggest certain tasks for a robot mainly based on its appearance and also on its functionalities. The analysis is based on the data from the web survey. Implications for the design of domestic robots will be identified. In the sense used in this paper the terms domestic and private are interchangeable.

Altogether, 127 people participated in the survey (61% male, 39% female). Most of the participants have German nationality (92%), which is due to the fact that the questionnaire was published in German (0.8% Swiss, 2.4% Austrian, 4.8% other). Their age ranged between 9 and 65 years (average=27.4 years). The educational background of the subjects was: 32.3% high school graduates; 55.9% university graduates; 9.5% doctoral degree; 2.3% other. 94% indicated to use computers both at home and in their job. They rated their own knowledge about computers fairly high (25% expert, 67% advanced knowledge). This data shows that the sample was not representative for what we would call the general public. On the other hand, we suppose that people taking part in such a study are rather representative for the future user group of the robots. Nevertheless, the validity is restricted to Germany. Therefore, a further study is already set up in English. For the German survey presented here, participants needed an average of 11min 21sec.

As described above, in the second part of the questionnaire people were asked to write down names of robots they knew. Altogether 284 entries were made (mean=2.39, SD=1.73). The robots mentioned most often were robots from TV and media (R2D2, C3PO etc.) (71), industrial robots (49), AIBO (23), service robots (21), and soccer robots (18). 71% indicated to know the robots from the media, 26% had seen a robot in real life, and only 9% had actually interacted with a robot before. Only four participants had used a domestic robot at home. All robots used at home were toy robots like Lego Mindstorms. Seven people had worked with a robot at a university.

As mentioned above, some participants had seen robots in real life. These were mainly industrial robots at work or at an exhibition. Some had seen AIBO at a soccer competition. Altogether, participants can be regarded as rather inexperienced in the field of robotics. One fact that has to be kept in mind though is that 18% knew AIBO, which is one of the robots shown in the study. The question if this has an effect on peoples' judgement of the robot will be addressed later in this section.

Altogether, people show a widespread rejection to using the robots (Table 1). The following comments taken from the questionnaires underline this (quotations were translated from German to English by the authors): "I don't want any robot at all!", "People or animals should not be replaced by robots.", "I don't think it's desirable to live together with a robot.", "I don't think that there are any useful applications for the robots shown here.", and "The robots scare me.". Participants made many more similar comments. As Kaplan [44] argues this is typical for the Western in contrast to the Japanese culture. One reason for rejecting robots might be that we still lack applications, which add some value to peoples' lives [19]. This explains why the willingness to use BIRON, the functionally designed robots, is highest (21.2%). As will be seen later in this section, quite a few participants could imagine using this robot especially in a public context.

TABLE 1
WILLINGNESS TO USE THE ROBOTS (SAMPLE (N); YES, MAYBO, NO IN %)

	AIBO	iCat	BARTHOC	BIRON	average
n	121	124	122	118	
yes	15.7	9.7	9.0	**21.2**	13.9
maybe	23.9	25.8	26.2	27.1	25.8
no	60.3	64.5	64.8	51.7	60.3

TABLE 2
RATING (IN %) OF THE QUESTIONS: (A) WHICH ROBOT WOULD YOU LIKE TO OWN? (B) WHICH ROBOT IS MOST ENJOYABLE? (C) WHICH ROBOT IS MOST LIKEABLE?; CORRELATION BETWEEN THE ANSWERS: ** P<.01.; SAMPLE (N)

	n	Aibo	iCat	Barthoc	Biron	(a)	(b)
(a)	111	**44.1**	17.1	5.4	**33.3**		
(b)	114	**57.0**	16.7	9.7	16.7	.47**	
(c)	110	**49.1**	**39.1**	2.7	9.1	.41**	.33**

Although the participants rejected using the robots in general, they preferred AIBO when asked which robot they would like to own, they find most enjoyable to interact with, and they regard as most likeable (Table 2). Answers to all three questions are strongly correlated as can be seen in Table 2. The preference for AIBO might be due to its appearance. As [45] found, people tend to find the robot cute and feel attracted to it. Moreover, many people knew AIBO beforehand and, thus, an acquaintance effect might also explain the preference.

It was further found that many participants refused BARTHOC. Only 2.7% of the participants stated that it was the most likeable of the robot. Also least subjects would like to own the robot (5.4%) and it was found less enjoyable (9.4%). This finding was supported by the fact that seven participants even wrote that they find BARTHOC scary. This can be explained with the uncanny valley theory (see II C). BARTHOC was the most human-like robot in the study. Anyhow, its appearance is far from imitating a human in a believable manner. This was also true for the movement displayed in the video which was rather choppy. As the uncanny valley theory states, robots that are supposed to look human-like but are as unperfected as BARTHOC is at the moment, cause a feeling of eeriness in the interlocutor. Also BARTHOC is perceived as being abnormal. Even though its proportions are similar to a human, its surface and movements are not. Since the uncanny valley theory describes problems caused by appearance and movements, it can be assumed that this is the reason for the bad rating of BARTHOC since the robot is not lagging behind the other robots regarding functionality.

VI. APPLICATION CATEGORIZATION

As the central question of the survey, participants were asked to propose applications for AIBO, BIRON, BARTHOC, and iCat. Altogether 570 items were mentioned (see Table 3). Gender and age did not have a statistically significant influence on subjects' answers. As for the robots, there is a strong correlation between applications proposed for AIBO and iCat (r=.77; p<.01 (two-tailed)). It can be assumed that this is due to appearance. As explained above, both have animal like features. Moreover, being consumer products both robots convey a similar degree of product-ness (see Section II B) which the other robots do not have.

For further analysis a method to classify the applications was developed. First, a content analysis was conducted. Three researchers classified the applications mentioned by the participants of the study independently into a free set of functional categories which should not overlap. The categories were than discussed and compared. On this basis a final set of application categories was developed. Second, based on these classifications the applications were grouped (Table 3). If one participant had written down two or more applications belonging to the same category the applications were counted as one. Every category was assigned to one of the higher-level environment-oriented groups domestic applications, public applications, or both domestic and public applications. In the focus of this paper are the domestic applications. However, to identify special requirements for these it will be necessary to contrast them with the other groups. Thus, firstly all categories are described and it is explained why certain applications are assigned to a certain group.

A. Domestic Applications

The first group consists of **domestic** applications. Within this group, *Healthcare* refers to robots used for therapy (e.g. autism therapy) and as support for sick or old people. This category also includes *Caregiver robots* that are used to watch old or sick people when nobody else is around. *Companionship* consists of all robots that keep company.

TABLE 3
APPLICATIONS PROPOSED FOR THE ROBOTS BY THE PARTICIPANTS SORTED BY CONTEXT OF USE

Applications	Aibo	iCat	Barthoc	Biron	total
domestic:					
Healthcare / Care giving	**19**	**22**	8	8	57
Companionship	**8**	3	2	2	15
Entertainment	**10**	**8**	2		20
Toy	**55**	**44**	5	1	**105**
Pet	**20**	6			26
Personal Ass. / Interface	7	**14**	3	**30**	54
both:					
Security	23	17	5	**32**	**77**
Teacher	4	**16**	5	3	28
Research	**11**	1	4	2	18
Transport (Fetch & Carry)	**14**		2	5	21
public:					
Business	3	12	**32**	12	59
Public Assistant	6	13	27	**44**	**90**
total	180	156	95	139	**570**

According to the theory of anthropomorphism by Epley et al. [30] people anthropomorphize to a higher extent if they feel lonely, because for a human being there is a principle need for social connection. Loneliness elicits social pain, which hurts in the same way as physical pain [45]. Actually, e.g. elderly pet owners appear to be buffered from negative impact of stressful life events and visit a doctor less often compared to elderly people without pets [47]. Therefore, both Caregiver and Companionship are applications where robots might compensate the feeling of loneliness or social pain, respectively. This is what differentiates these categories from entertainment. The sole purpose of *Entertainment robots* is to entertain their users and to be a pastime. They are not built to have a psychological impact. The same is true for *Toy robots* which are mainly used for playing. Most robots already being sold for domestic usage belong to this category. Another application, which has been mentioned above, is *Pet robots.* It implies that the human shows responsibility for the robot. Pet robots are animal-like in appearance and function and might take the place of a real pet. *Personal assistant or Interface* describes robots used as butlers, organizers or interfaces. This category includes robots for cleaning and other household chores.

All applications described so far are domestic. We want to contrast them with the ones Christensen differentiates. As mentioned in Section II A [16], he names three groups of domestic applications (entertainment, assistance for elderly and handicapped, everyday tasks). We propose six categories, however, his groups are part of the categorization introduced in this paper. It is assumed that entertainment is part of the entertainment category established here; assistance for elderly and handicapped is similar to Healthcare / Care giving; and robots for everyday tasks are included in Personal Assistants / Interface. As a result, the categorization presented here includes three more categories (Companionship, Toy, Pet). One could argue that Companionship could be included in Healthcare or in Entertainment, Toy and Pet could also be incorporated in the Entertainment or in the Companionship category. Nevertheless, there are some major reasons not to do so. Mainly the differentiation is motivated by the fact that the categories bring about different capabilities and / or appearances the robot has to have. While Companionship is more than Entertainment because the state of the user has to be taken into account by a successful companion it is less than a Healthcare robot or a Caregiver whose function is broader since it has more responsibility for the well being of a person. Again there is a functional difference between a Toy and an Entertainment robot. Fong et al. also differentiate between these two groups (see II A, [17]). While a person usually actively makes up games for a toy, an entertainment robot is restricted to specific pre-programmed tasks. The difference between Pet and other categories is mainly due to the appearance of the robot. As discussed for AIBO, the appearance of the robot in this case obviously causes certain feelings in the user that are similar to feelings caused by real pets.

B. Public and Domestic Applications

The following applications can be used in **both public and domestic** environments. *Security applications* include robots used for surveillance, military tasks, exploration, tasks that are dangerous for humans (e.g. minesweeping), and for protection. Another category includes robots that *teach* certain tasks or abilities. The robots in this case are supplements to real teachers especially when learning languages. *Research* applications are useful to learn about robotics and human nature. Thus, these applications also serve to improve the robots themselves. *Transport robots* are useful for all kinds of fetch and carry tasks.

As noted above, these tasks can be both public and domestic. Even though the application does not allow for a concrete allocation, the appearance of the robots might in some cases. Whereas it is imaginable to use iCat as a teacher in a public as well as in a domestic environment, it is more plausible - because of the appearance of the robots – to use BIRON for security in public whereas AIBO might be used for the same task at home. The same is true for Fetch & Carry tasks, one reason being that AIBO could easily be overseen and stepped on in public environments.

C. Public Applications

The last group consists of **public** applications. Within this group, a *Business robot* is either a receptionist, a sales robot, or a robot used for representation. *Public assistants* are guides (museum), information terminals or translators. All these robots are usually used for a short time at once and with other possible interaction partners being around.

Altogether, it cannot be reassured that the classification is exhaustive. There might be more applications that did not come up in the study especially since only four robots were tested. Anyhow, the categories help to explain the resulting differences between the robots and to compare them to scenarios proposed by the developers.

VII. Applications for the Robots

While the last section has focused on the categorization of all applications we now elaborate on the applications mentioned for the four robots. It was found that the applications named most often were Toy (105), Public Assistant (90), and Security (77) indicating that potential users have quite different and more diverse ideas on robot applications than robot developers. This is presumably due to the fact that participants who are not working in the field of robotics know less about feasibility of certain applications. Nevertheless, most applications mentioned were quite realistic from a robotics point of view or at least not completely out of reach. One reason is, presumably, that participants' were explicitly asked for "useful" applications.

AIBO was mainly seen as a domestic robot, especially as a Toy (55) and a Pet (20). This might have been influenced by the fact that many cheaper robotic dogs are available off-the-shelf as toys. The applications mentioned are in line with the notion of "entertainment" used by Sony. Participants also suggested that the robot could be used for Healthcare / Care giving (19). This category also includes tasks like guide dog. The fact that people mention tasks very typical for dogs highlights that AIBO indeed has a zoomorphic appearance. It also explains that many subjects suggest Fetch & Carry things (14) and Security (23) (guard dog) as applications for the robot.

As discussed above, in contrast to AIBO iCat is not a complete cat but only a torso. It has a rather caricatured appearance which explains that it was seen as a Pet by only few people (6). Part of the robot's caricatured appearance are its human-like facial features. These are presumably the reason that many participants suggested Teacher (16) (especially for languages) as an application for the robot. As explained in Section II B, Duffy states that human-like attributes are necessary for meaningful interaction. Meaningful interaction is undoubtedly especially important in teaching scenarios, which are very goal oriented. Thus, iCat with its facial features in fact seems suitable for this task. On the other hand iCat, like AIBO, most strongly evoked the impression of being a Toy (44). Apart from this, 14 people suggested Personal assistant as an application, half of these Interface, which is the scenario proposed by the developers. Surprisingly the participants did not indicate that they knew iCat in advance which would have implied that they had also been familiar with this scenario beforehand. The robot was also strongly linked with Healthcare / Care giving (22). In the applications described in this category, iCat in a broader sense is also an interface – not between a human and its electronic devices but between an old or sick person and other people that are not physically present.

Least applications were written down for BARTHOC. Most of the ones named are public (Business – 32, Public Assistant – 27). This is underlined by the following quotation (translation by the authors): "I would use the robot for the tasks mentioned above but only in the public. I would not want to have such a thing at home". However, the applications proposed are in line with the receptionist scenario the developers had in mind.

BIRON was judged quite differently. It was the only robot accepted for both public and private applications. On the one hand, people named applications like personal assistant (30), which supports the home tour scenario the researchers are mainly working on. On the other hand, participants strongly suggested using the robot for Security (32) and Public assistance (44). In fact, some pretty similar systems are already used as museum tour guides (e.g. [7],[8]) which might have influenced the subjects.

VIII. Applications Participants Would Use

Next to general applications, participants were asked to name applications they would use themselves. These were categorized just as the ones above. 12 people wrote down applications they would use with BARTHOC, 28 with iCat, 38 with AIBO, and 36 with BIRON. Altogether, 122 applications were mentioned (12 BARTHOC, 30 iCat, 42 AIBO, 38 BIRON). Almost all people that would use the robot could only imagine using it for one certain task. Again, by far the least applications were mentioned for BARTHOC and most for AIBO.

Looking at Table 4, the results underline the findings described so far. Participants would use AIBO and iCat in a domestic environment, whereas BARTHOC was seen as a public robot. While some people would use domestic applications for BIRON the preference to employ this robot in public was strengthened. This might also be due to the functionality of the robot in connection with the appearance. While a public assistant e.g. an information terminal might not need an arm, a personal assistant built to carry things around would urgently need one. A statement by one user underlines this (translation by the authors): "I need a mobile robot with a gripper so that it can really do something". This supports the idea that even though appearance is a crucial factor in robot design also functionality and its representation in the appearance of the robot plays an important role.

TABLE 4
Applications of robots participants would use themselves sorted by context of use

Applications	Aibo	iCat	Barthoc	Biron	total
domestic:					
Healthcare / Care giving	1	1		2	4
Companionship	1				1
Entertainment	2	3			5
Toy	**16**	**9**	2		27
Pet	1	1			2
Personal Ass. / Interface	5	5	1	**8**	19
both:					
Security	**8**	3	1	3	15
Teacher	1	3		1	5
Research	2				2
Transport (Fetch & Carry)	**4**				4
public:					
Business			1	**5**	6
Public Assistant	1	5	**7**	**19**	32
total	42	30	12	38	**122**

Table 4 also reveals that people only use applications if they have a need to do so (see [19], Section II A). Obviously only few have a need for Healthcare / Care giving, Companionship and so forth. This finding illustrates the fact that target groups for certain applications have to be identified and the design has to be based mainly on their needs.

IX. Conclusion

Altogether, it can be concluded that almost half of the applications proposed are domestic applications according to the categorization presented here (49% of all applications; 48% of applications the participants would use themselves). Moreover, many applications especially the ones mentioned for AIBO that are included in the group "both" are also very likely to being used in private as long as the appearance of the respective robot is appropriate. These findings give quite some evidence on the importance that social domestic robots will gain in the future.

Moreover, they indicate that appearance in fact plays a crucial role in the perception of a robot and determines which applications are proposed. This should be kept in mind when designing a robot. In contrast to appearance the functionalities of the robot seem to be less important. They become crucial when something is missing for a certain scenario the participants would like to have (e.g. the gripper for BIRON in a domestic assistant scenario). One has to keep in mind, however, that these results are based on an internet survey. They might and probably will change in live human-robot interaction. However, appearance should be consciously designed to help users build a mental model of what a robot's application is.

The study presented here is only a first step to research perception of appearance and capabilities of robots. Certainly, more studies with different robots and methodologies have to follow. Nevertheless, including potential users in the process of finding new applications and designing social robots has proven to be a promising means of doing research in the field of social robotics.

Acknowledgment

The authors would like to thank Agnes Swadzba, Sven Wachsmuth and Katharina Rohlfing for their help and advice regarding the design and the realization of the study.

References

[1] J. Goetz, S. Kiesler, and A. Powers, 'Matching robot appearance and behavior to tasks to improve human-robot cooperation', in Proceedings of the 12th IEEE Workshop on Robot and Human Interactive Communication (ROMAN 2003), San Francisco, CA, 2003, pp. 55–60.

[2] Lego Mindstorms, http://mindstorms.lego.com

[3] Ugobe Pleo, http://www.pleoworld.com/

[4] WowWee Robosapien, www.wowwee.com/robosapien/robo1/robomain.html

[5] iRobot Roomba, http://irobot.com/

[6] Roboceptionist, http://www.ri.cmu.edu/projects/project_523.html

[7] I. R. Nourbakhsh, C. Kunz, and T. Willeke, The mobot museum robot installations: a five year experiment, in Proc. Intelligent Robots and Systems, 2003. (IROS 2003), vol.3, 2003, pp. 3636 - 3641.

[8] M. Bennewitz, F. Faber, D. Joho, M. Schreiber, and S. Behnke, Towards a Humanoid Museum Guide Robot that Interacts with Multiple Persons, in: Proceedings of the International Conference on Humanoid Robots (Humanoids), 2005.

[9] B. Duffy, 'The Social Robot', Ph.D. Thesis, Department of Computer Science, University College, Dublin, 2000.

[10] B. Duffy, 'Anthropomorphism and The Social Robot', Special Issue on Socially Interactive Robots, Robotics and Autonomous Systems 42 (3--4), 2003.

[11] C. Breazeal, A. Brooks, J. Gray, G. Hoffman, C. Kidd, H. Lee, J. Lieberman, A. Lockerd, and D. Chilongo, 'Tutelage and collaboration for humanoid robots', International Journal of Humanoid Robots, 1(2), 2004, pp. 315– 348.

[12] C. Breazeal, 'Designing Sociable Robots'. Cambridge, MA, USA, MIT Press, 2002.

[13] C. Bartneck, J. Forlizzi, 'A Design-Centred Framework for Social Human-Robot Interaction', Proceedings of RO-MAN 2004, Kurashiki, 2004, pp. 591-594.

[14] M. Lohse, F. Hegel, A. Swadzba, K. Rohlfing, S. Wachsmuth, and B. Wrede, 'What can I do for you? Appearance and Applications of Robots' Proceedings of AISB'07, Artificial and Ambient Intelligence, Newcastle upon Tyne, 2007.

[15] F. Hegel, M. Lohse, A. Swadzba, K. Rohlfing, S. Wachsmuth, and B. Wrede, 'Applications of Social Robots: A User Study', in Proc. Int. Symposium on Robot and Human Interactive Communication (ROMAN), Jeju Island, Korea, 2007.

[16] H. I. Christensen, Intelligent home appliances. In Robotics Research, R. A. Jarvis and A. Zelinsky, Eds., no. 6 in Springer Tracts in Advanced Robotics (STAR). Springer Verlag, Heidelberg, Germany, January 2003, pp. 319-330.

[17] T. W. Fong, I. Nourbakhsh, and K. Dautenhahn, 'A Survey of Socially Interactive Robots: Concepts, Design, and Applications', Robotics and Autonomous Systems, 42(3 – 4), 2002, pp. 142 – 166.

[18] S. Ljungblad and L. E. Holmquist, 'Designing robot applications for everyday environments', in sOc-EUSAI '05: Proceedings of the 2005 joint conference on Smart objects and ambient intelligence, New York, NY, USA, ACM Press, 2005, pp. 65–68.

[19] F. Kaplan, 'Everyday robotics: robots as everyday objects', in Proceedings of Soc-Eusai 2005, Grenoble, France, 2005, pp. 59 – 64.

[20] C. Bartneck, 'Interacting with and embodied emotional Character'. Proceedings of the DPPI2003 Conference, Pittsburgh, 2003, pp. 55-60.

[21] Y. Jung and K. M. Lee (2004). 'Effects of physical embodiment on social presence of social robots'. Proceedings of Presence, 2004, pp. 80-87.

[22] L. Sullivan, The Tall Office Building Artistically Considered, in Kindergarten Chats (revised 1918) and Other Writings, edited by I. Athey, New York: Wittenborn, Schultz, 1947, pp. 202-13.

[23] M. Blow, K. Dautenhahn, A. Appleby, C. L. Nehaniv, and D. C. Lee, 'Perception of Robot Smiles and Dimensions for Human-Robot Interaction Design' 15th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 06), pages 469-474, Hatfield, UK, (2006).

[24] J. Donat,. 'Mediated Faces', in M. Beynon, C. L. Nehaniv, K. Dautenhahn (Eds.). Cognitive Technology: Instruments of Mind. Proceedings of the 4th International Conference, CI 2001, Warwick, UK, 2001.

[25] R. Gockley, J. Forlizzi, and R. Simmons, 'Interactions with a moody robot', in HRI '06: Proceeding of the 1st ACM SIGCHI/SIGART conference on human-robot interaction, New York, NY, USA, ACM Press, 2006, pp. 186–193.

[26] A. Bruce, I. Nourbakhsh, and R. Simmons, 'The role of expressiveness and attention in human-robot interaction', in Proc. AAAI Fall Symp. Emotional and Intel. II: The Tangled Knot of Soc. Cognition, 2001.

[27] F. Hegel, T. Spexard, T. Vogt, G. Horstmann, and B. Wrede, 'Playing a different imitation game: Interaction with an Empathic Android Robot', in Proc. 2006 IEEE-RAS International Conference on Humanoid Robots (Humanoids06), IEEE, 2006, pp. 56–61.

[28] C. F. DiSalvo, F. Gemperle, J. Forlizzi, and S. Kiesler, 'All robots are not created equal: the design and perception of humanoid robot heads', in DIS '02: Proceedings of the conference on Designing interactive systems, pp. 321–326, New York, NY, USA, ACM Press, 2002.

[29] F. Hegel, S. Krach, T. Kircher, B. Wrede, and G. Sagerer, 'Theory of Mind (ToM) on Robots: A Neuroimaging Study', Proceedings of Human-Robot Interaction (HRI 2008), Amsterdam, Netherlands, 2008.

[30] N. Epley, A. Waytz, and J. T. Cacioppo, 'On Seeing Human: A three-factor theory of anthropomorphism', Psychological Review, 114, 2007, pp. 864-886.

[31] H. v. Foerster, 'Wissen und Gewissen. Versuch einer Brücke', Frankfurt, 1997.

[32] S. E. Guthrie, 'Anthropomorphism: A Definition and a Theory', in: R. W. Mitchell, N. S. Thompson, and H. L. Miles, 'Anthropomorphism, anecdotes, and animals'. Albany: State University of New York Press, 1997.

[33] C. K. Morewedge, J. Preston, and D. M. Wegner, 'Timescale is in the attribution of mind', Journal of Personality and Social Psychology, 93, 2007, pp. 1-11.

[34] D. Sakamoto, T. Kanda, T. Ono, H. Ishiguro, and N. Hagita, 'Android as a Telecommunication medium with Human Like Presence'. ACM/IEEE 2nd Annual Conference on Human-Robot Interaction (HRI2007), Arlington, Virginia, USA, 2007, pp. 193-200.

[35] M. Hackel, M. Schwope, J. Fritsch, B. Wrede, and G. Sagerer, ‘Designing a sociable humanoid robot for interdisciplinary research’, Advanced Robotics, 20(11), 2006, pp. 1219-1235.

[36] K. F. MacDorman and H. Ishiguro, ‘The uncanny advantage of using androids in social and cognitive science research’, Interaction Studies, 7(3), 2006, pp. 297-337.

[37] M. Mori, ‘Bukimi no tani’. (Translation by K. F. MacDorman and T. Minato, http://www.androidscience.com/theuncannyvalley/proceedings2005/uncannyvalley.html). Energy, 7(4), 1970, pp. 33–35, (Originally in Japanese).

[38] K. F. MacDorman, ‘Subjective ratings of robot video clips for human likeness, familiarity, and eeriness: An exploration of the uncanny valley’, ICCS/CogSci-2006 Long Symposium: Toward Social Mechanisms of Android Science. July 26, 2005. Vancouver, Canada, 2005.

[39] J. Seyama and R. S. Nagayama,, The uncanny valley: Effect of realism on the impression of artificial human faces, Presence: Teleoperators and Virtual Environments, 16(4), 2007, pp. 337-351.

[40] Sony AIBO website. http://support.sony-europe.com/aibo/

[41] Philips Research, iCat research community, http:// www.hitech-projects.com/icat/

[42] N. Beuter, "Gestenbasierte Positionsreferenzierung für die multimodale Interaktion mit einem anthropomorphen Robotersystem", diploma thesis, Faculty of Technology: Bielefeld University, 2007.

[43] A. Haasch, S. Hohenner, S. Hüwel, M. Kleinehagenbrock, S. Lang, I. Toptsis, G. A. Fink, J. Fritsch, B. Wrede, and G. Sagerer, ‘BIRON – The Bielefeld Robot Companion’, in Proc. Int. Workshop on Advances in Service Robotics, Eds. E. Prassler, G. Lawitzky, P. Fiorini, and M. Hägele, Stuttgart, Germany, Fraunhofer IRB Verlag. 2004, pp. 27–32.

[44] F. Kaplan, ‘Who is afraid of the humanoid? Investigating cultural differences in the acceptance of robots’, International journal of humanoid robotics, 1(3), 2004, pp. 465-480.

[45] C. Bartneck, J. Reichenbach, and J. Carpenter, Use of Praise and Punishment in Human-Robot Collaborative Teams, Proceedings of the RO-MAN 2006 - The 15th IEEE International Symposium on Robot and Human Interactive Communication, Hatfield, 2006, pp. 177-182.

[46] G. MacDonald and M. R. Leary, Why does social exclusion hurt? The relationship between social and physical pain, Psychological Bulletin, 131, 2005, pp. 202–223.

[47] J. M. Siegel, Stressful life events and use of physician services among the elderly: The moderating role of pet ownership, Journal of Personality and Social Psychology, 58, 1990, pp. 1081–1086.

The Influence of Social Presence on Acceptance of a Companion Robot by Older People

Marcel Heerink, Ben Kröse, Vanessa Evers and Bob Wielinga

***Abstract*— If robotic companions are to be used in the near future by aging adults, they have to be accepted by them. In the process of developing a methodology to measure, predict and explain acceptance of robotic companions, we researched the influence of social abilities, social presence and perceived enjoyment. After an experiment (n=30) that included collecting usage data and a second experiment (n=40) with a robot in a more and less sociable condition we were able to confirm the relevance of these concepts. Results suggest that social abilities contribute to the sense of social presence when interacting with a robotic companion and this leads, through higher enjoyment to a higher acceptance score.**

***Index Terms*— Human robot interaction, technology acceptance, assistive technology**

I. Introduction

In the last decade a growing number of research projects have addressed the possibilities of robots in eldercare [1-3]. Besides in rehabilitation, where robotic technologies are already common, robots can serve as pets and fulfill roles that caregivers would fulfill. Besides to providing companionship, their functionalities can be related to supporting independent living by supporting basic activities (eating, bathing, visiting the bathroom, getting dressed) and mobility, providing household maintenance, monitoring of those who need continuous attention and maintaining safety [4-7]. Thus, robotic companions are generally considered a potentially major part of the technology that can address the problems of a growing older population and increasing labor shortage in the industrialized world.

However, there are challenges to be met - and not only technical ones. Elders do not always willingly accept new technologies and it might be crucial to map the psychological requirements that designers of robotic companions have to take into account [4]. Furthermore, robots are not only perceived as assistive devices, they are also perceived as social entities and it could be crucial for them to have certain social abilities in order for communication to run smooth and naturally. It might very well be that the more natural and 'human like' the conversation with a robot is, the more enjoyment a user feels and the more this user would feel encouraged to actually use this technology. Besides, robots could have 'hedonic' aspects: users might enjoy interacting with the robot as they would enjoy playing a game or having a pleasant conversation with a person.

In our research, we address some of those challenges by exploring and modeling the factors that may influence acceptance of a conversational robot by older users [8, 9]. Recent studies on human robot interaction stress the importance of social intelligent behavior in robots, even more so in a health- and eldercare environment [10-12]. This notion stems from the premise that a more socially intelligent robot will be more effective in its communication and is therefore easier and more pleasant to interact with and hence accepted easier. Inspired by these findings, we are particularly interested in the role that a robot's social abilities play in its communication with and acceptance by older users.

In order to establish the influence of (presumed) social abilities of a robot on its acceptance by older users, we developed a theoretical model and carried out two experiments that will be described in this paper. After discussing related work and introducing the theoretical framework we will describe these two experiments and present their results.

This work was supported in part by the Hogeschool van Amsterdam and in part by the European Commission Division FP6-IST Future and Emerging Technologies under Contract FP6-002020 (Cogniron).

Marcel Heerink is with the Instituut voor Information Engineering, Almere, Netherlands (+31-36-5480771; e-mail: m.heerink@hva.nl).

Ben Kröse is with the Institute of Informatics, University of Amsterdam, Amsterdam, the Netherlands, (e-mail: krose@science.uva.nl).

Vanessa Evers is with the Human Computer Studies Lab at the University of Amsterdam, the Netherlands (e-mail: evers@science.uva.nl).

Bob Wielinga is with the Human Computer Studies Lab at the University of Amsterdam, the Netherlands (email: wielinga@science.uva.nl)

II. Related work

A. Robots in eldercare

Recent projects concerning eldercare companion robots focus on possible application areas and requirements for robot companions or on measuring user responses to the robots. An example of the latter is the research concerning a seal shaped robot (Paro) [6, 13, 14]. Paro's function is merely that of a social companion without any assistive functionalities, more or less comparable to a pet. These experiments showed that a robot could have the same beneficial effect on elders that a pet can have, making them feel happier and healthier. A more recently developed robot with similar pet-like functionalities is the Huggable [15].

Another example of a robot developed specifically for eldercare experiments is 'nursebot' Pearl, a robot that could actually provide advanced assistance to elders, although its functionalities were merely simulated in the reported user studies [16, 17].

A robot with advanced assistive functionalities to be applied in eldercare is the German Care-o-bot. It is intended to provide assistance in many ways, ranging from a walking aid to a butler [18, 19].

Other projects focus on an assistive environment rather then on the development of a specific robot. An example of this is the RoboCare project [20], featuring an intelligent home of which a robot is an integrated part.

The above examples of experimental robots applied in eldercare indicate the growing interest and applicability of robots in eldercare and show a research focus on dual performance where the robots are positioned as social actors as well as fulfill practical functions to support assisted living.

B. Technology acceptance and robots

Technology acceptance methodology has had much attention in the last two decades. Although there have been earlier models, an overview of technology acceptance usually starts with the introduction of the technology acceptance model (TAM) by Davis in 1986 [21, 22]. This model has become one of the most widely used theoretical models in behavioral psychology. In its most basic form it states that Perceived Usefulness and Perceived Ease of Use determine the behavioral Intention to Use a system and it poses that this behavioral intention is predicting the actual use. The model has been used for many different types of technology and has been extended with more factors that were found to influence Intention to Use or Usage.

Usually in acceptance model methodology, each factor is represented in a questionnaire by a group of questions or statements that can be replied to on a five or seven point Likert scale. The thus formed construct can often not only be related to the intention to use a system, but also directly to the Usage of a system or to each other. The validation of a model typically includes a long term observation of the actual use of technology, which makes it possible to relate scores on Intention to Use to Usage of the system.

In 2003, Venkatesh et al. [23] published an inventory of all current models and factors and presented a new model called UTAUT in which all relevant factors would be incorporated. This UTAUT model has been used by research on acceptance of a conversational robot is described by De Ruyter et al [11]. It concerned a robotic interface (the iCat made by Philips), which was tested in a Wizard of Oz experiment where the robot was controlled remotely by an experimenter while the participants perceived it to be autonomous. This experiment was done in a laboratory setting, with adult, but not elderly participants.

The results showed that an extravert (more expressive in voice and face) iCat was perceived to be more socially intelligent and was also more likely to be accepted by the user than a more introvert version. The same robot was used in an experiment by Looije et al. [24] where it featured as a personal assistant for a small group of people with diabetes. Results showed that participants appreciated the iCat in the high social ability condition more and that users had a higher intention of using it than was the case for a less socially intelligent iCat.

These finding show that perceived social abilities of a robot are appreciated by users as they would be in a human conversational partner. In previous research we replicated the study by de Ruyter et al. to evaluate whether the same effect exists for older users of a robotic companion [8]. In applying the UTAUT instrument, the hypothesis that social abilities contribute to the acceptance of a robot for elderly could not be confirmed. Instead, we identified new variables that seemed especially relevant to the elderly population.[8].

In the next section we will introduce the three new variables that we propose are relevant for a companion robot acceptance model for the elderly: Perceived Enjoyment, Perceived Sociability and Social Presence.

III. MEASURING ACCEPTANCE OF ROBOTIC TECHNOLOGY BY OLDER USERS

A. Perceived Enjoyment

The original TAM, related models and UTAUT were merely developed for and validated in a context of utilitarian systems in a working environment. Robotic technology used outside a working environment provides systems that might be experienced as more than this: users might have a sense of entertainment when using it. Van der Heijden [25] points out that in 'hedonic systems', the concept of enjoyment is a crucial determinant for the Intention to Use it.

Of course, robotic technology in eldercare will hardly be developed just to entertain: it will be partly utilitarian, partly hedonic. But even if just partly hedonic, enjoyment could prove to be a construct that needs to be part of an acceptance model for robotic technology in eldercare.

Furthermore, Perceived Enjoyment can also be of importance in utilitarian systems, as pointed out in an extensive study by Sun and Zhang [26]. The study mainly supports the claims by Venkatesh et al. [23] and Yi and Hwang [27], that Perceived Enjoyment has no direct influence on Intention to Use, but that it can influence Ease of Use and Usefulness. Still the study does also recognize that this is not a general claim for all types of systems. Indeed this could work very differently for robotic systems used by older people.

An acceptance study that also included perceived enjoyment by Chesney, concerned the use of Lego Mindstorms development environment by Mindstorms hobbyists [28]. The study, based on the viewpoint that this concerns a partly hedonic, partly utilitarian type of system, confirms perceived enjoyment having just an indirect effect on intention to use.

We may conclude that literature on acceptance models in general does attribute some influence to Perceived Enjoyment in systems that are partly or totally hedonic. Since socially interactive robots may be experienced as hedonic systems, this means Perceived Enjoyment could be of some influence. When we consider social acceptance also to be a factor, especially with conversational robots, this means robotic systems differ from the systems described in acceptance model literature so far and the strength of the influence of Perceived Enjoyment is still very much uncertain, especially in the context of eldercare.

B. *Social Presence*

Since it is not unusual for humans to treat systems and devices as social beings [29] it seems likely that humans treat embodied agents as such. The extend to which they do so seems to be related to a factor that is often related to as either ‘Presence’ or, more specifically ‘Social presence’. Many research projects that are related to our research, incorporate this concept [30-32].

The term presence originally refers to two different phenomena. First it relates to the feeling of really being present in a virtual environment and can be defined as ‘the sense of being there’ [33]. Second, it can relate to the feeling of being in the company of someone: ‘the perceptual illusion of non mediation’ [34]. In our context, the second definition is relevant.

Regarding the close connection between social abilities and the sense of presence, there could be a crucial role for presence in the process of acceptance of functional and conversational acceptance of embodied agent technology. Therefore we intend to incorporate measuring social presence in our experiments to research its role and establish the influence of social abilities on it..

C. *Perceived Sociability*

The development social abilities for robots in general and for eldercare companions in particular is often recommended [10, 35, 36]. Nevertheless, Perceived Sociability as a construct in an acceptance model is new. We need it as such to establish weather users are conscious of the possibilities of seeing and judging a robot as a social entity. Besides, we want to measure the amount in which they perceive these abilities to see how it relates to the amount in which social presence is perceived.

In research concerning social aspects of autonomous interactive systems there are several definitions of the concept of social intelligence [37]. For the purpose of this study, social intelligence will be the social abilities, perceived by the users when interacting with robots.

A similar description is given for *socially communicative* robots within the classification by Breazeal [38](extended by Fong et al. [37]): robots providing a ‘natural’ interface by employing human-like social cues and communication modalities, that do not have to be based on deep models of social cognition.

Since we are interested in the influence of social abilities in a robotic interface on its acceptance, it is important to look at ways to measure both acceptance and social abilities. A widely used tool to evaluate social abilities for humans is Gresham & Elliott's Social Abilities Rating System (SSRS) [39]. This tool usually is applied in social research, mostly on scholars and students, often in relationship to disabilities. Nevertheless, the five basic features (Cooperation, Empathy, Assertion, Self-Control and Responsibility) match the aspects found in Human-Robot Interaction literature on social (or sociable) robots and agents [37], [35] well. These five constructs also appear to be relevant abilities in the study by De Ruyter et al. [11].

Other relevant concepts to study are Trust and Competence as they appear relevant in the experiments by De Ruyter et al. and research by Shinozawa et al. [40]

This leads to the following list of social abilities:

1. cooperate,
2. express empathy,
3. show assertivity,
4. exhibit self control,
5. show responsibility,
6. gain trust,
7. show competence

To translate these into programmable features, we tried to meet with the list of social behaviors, set up in the experiments by De Ruyter et al. and found the following behavioral features to be programmed into our robots character (the numbers refer to the above listed abilities) [11, 40, 41]:

- listening attentively, for example by looking at the participant and nodding (1, 2),
- being nice and pleasant to interact with, for example by smiling (1, 2, 7),
- remembering little personal details about people, for example by using their names (6, 7),
- being expressive, for example by using facial expressions (2, 3),
- admitting mistakes (5, 6).

With this list we were able to incorporate all features except ‘exhibit self control’ (4), for which we could not find an applicable behavior in this context.

D. *Modeling the hypothesized influence of social abilities*

First of all, we want to establish the relationship between the Intention to Use the system and the actual use of it. Furthermore, we suspect Perceived Sociability (PS) to influence the sense of Social Presence (SP). This perceived Social Presence is expected to influence Perceived Enjoyment (PENJ) which can be a direct influence on the Intention to Use the system (ITU) or through Perceived Ease of Use (PEOU).

Finally, we also expect the original TAM constructs of Perceived Ease of Use an Perceived Usefulness (PU) to have a predictive influence on Intention to Use.

This leads to a model (represented graphically in Figure 1) based on the following hypotheses:

H1 Usage is predicted by Intention to Use

H2a The amount in which Social Abilities are recognized correlates with the amount in which Social Presence in perceived.

H2b The amount in which social abilities are implemented influences the amount in which Social Presence in perceived.

- H3 Social Presence is a determining influence on Perceived Enjoyment.
- H4 Perceived Enjoyment is a determining influence on Perceived Ease of Use
- H5 Intention to Use is determined by Perceived Enjoyment, Perceived Ease of Use and Perceived Usefulness

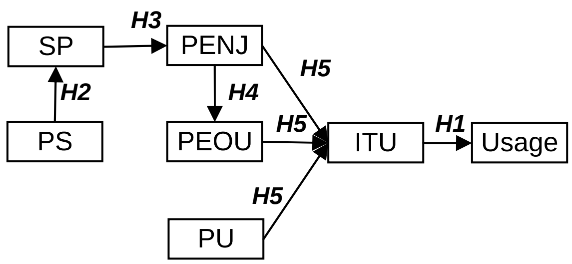

Fig. 1. Hypothetical model

E. Questionnaire

We developed a list witch three to six statements for each construct that test participants could reply to on a five point Likert scale that was transposed into verbal anchors: totally agree – agree – don't know – do not agree – totally do not agree. When processing the replies, these were related back to the 5 point scale. We wanted our participants to fill this list out themselves if possible.

Table 1 shows these statements translated into English (the original list is in Dutch).

TABLE 1
STATEMENTS FOR THE USED CONSTRUCTS

Construct	Statement
ITU	I think I'll use iCat the next few days
ITU	I am certain to use iCat the next few days
ITU	I'm planning to use iCat the next few days
PU	I think iCat is useful to me
PU	It would be convenient for me to have iCat
PU	I think iCat can help me with many things
PEOU	I think I will know quickly how to use iCat
PEOU	I find iCat easy to use
PEOU	I think I can use iCat without any help
PEOU	I think I can use iCat when there is someone around to help me
PEOU	I think I can use iCat when I have a good manual.
PENJ	I enjoy iCat talking to me
PENJ	I enjoy doing things with iCat
PENJ	I find iCat enjoyable
PENJ	I find iCat fascinating
PENJ	I find iCat boring
PS	I consider iCat a pleasant conversational partner
PS	I find iCat pleasant to interact with
PS	I feel iCat understands me.
PS	I think iCat is nice
SP	When interacting with iCat I felt like talking to a real person
SP	It sometimes felt as if iCat was really looking at me
SP	I can imagine iCat to be a living creature
SP	I often think iCat is not a real person.
SP	Sometimes iCat seems to have real feelings

IV. EXPERIMENTS

To test our hypotheses we set up two experiments. The first experiment was designed to gather data of actual use of a robotic companion that could be related to the Intention to Use of this system. Besides, correlational relationships could be established between the constructs. The second experiment was designed in order to be able to compare a robot with more sociability to a less sociable one. Besides confirming the established relationship between the constructs, this could test hypothesis H2b, looking at the implemented sociability (versus perceived sociability).

A. Used robotic system

The robotic agent we used in both experiment is the iCat ("interactive cat"), developed by Philips, also used in the experiments by De Ruyter et al. and Looije et al. and within our own project. The iCat is a research platform for studying social robotic user-interfaces. It is a 38 cm tall immobile robot with movable lips, eyes, eyelids and eyebrows to display different facial expressions to simulate emotional behavior. There is a camera installed in the iCat's nose which can be used for different computer vision capabilities.

B. First Experiment

1) Experimental setup

For our first experiment we used a setup in which the robot was connected to a touch screen as is shown in Figure 2. It could be used for information and for fun: the participants could ask for weather forecast, a television program overview or a joke by pressing the appropriate choices from a menu on the screen. The information was then given with pre-recorded speech by the iCat, for which we used a female voice. The recording was done with a text to speech engine.

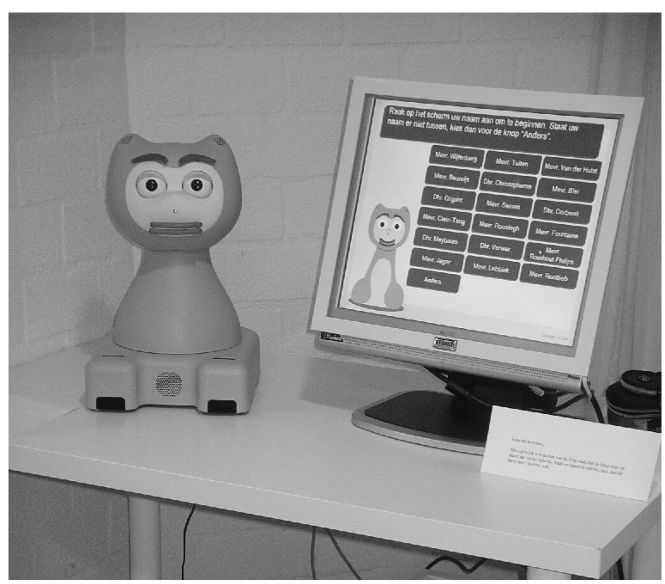

Fig. 2. Setup iCat with touch screen experiment

2) Subjects

There were 30 participants, recruited both by eldercare personnel and by students. Their age ranged from 65 to 94, while 22 of them were female and 8 were male. Some of them lived inside the eldercare institutions; some lived independently in apartments next to the institutions.

3) Procedure

The experiment consisted of two parts: a first initial test after which the questionnaire was used and a public use period to gather usage data.

For the initial test, participants were brought into a room were they were instructed to simply play with the robot for about three minutes. Subsequently they were brought to another room where they were given the questionnaire. They could ask for help if they were unable to read the statements.

After the initial test series were completed, we left the robot for public use in a tea room. On the screen were buttons with the names of the test session participants and one extra button saying “I’m not listed”. Passers by were informed by a note that anyone could use the robot and that they could start a session by pressing the button with their name on it or the “I’m not listed” button if their name was not on the screen.

During the days the iCat was available for use to anyone passing by, the system made video recordings as soon as it was used trough the camera in its nose. Furthermore, it kept a log of the start and end times of individual user sessions. The end time was either the time a user actively ended his session or if it was not used for 90 seconds.

By comparing the video footage to the log, we could later check if users had pressed the right button.

4) Results

The test session and the questionnaire were completed by 30 participants. In analyzing the reply scores, we used Cronbach’s alpha to test the reliability of the constructs. In psychology, an alpha of 0.7 and higher is considered acceptable. As table 2 shows, the constructs were highly reliable.

TABLE 2
CRONBACH’S ALPHA FOR THE USED CONSTRUCTS

Construct	**Items**	**Alpha**
Intention to use	3	,961
Perceived usefulness	3	787
Perceived ease of use	5	,820
Perceived enjoyment	5	,836
Perceived sociability	4	,786
Social presence	5	,866

Table 3 shows the correlations between the used constructs. The correlation between usage in minutes and Intention to Use is strong, which confirms our first hypothesis. Also the other hypothesized relationships are represented by strong correlations, except for Perceived Enjoyment (PENJ) and Perceived Ease of Use (PEOU).

TABLE 3
PEARSON’S CORRELATION FOR THE USED CONSTRUCTS AND USAGE MEASURED IN MINUTES

		ITU	PU	PEOU	PENJ	SP	PS
ITU	Correlation	1	,504**	,633**	,420*	,599**	,159
	Sig (2-tailed)		,005	,000	,021	,000	,402
PU	Correlation	,504**	1	,468**	,551**	,450*	,336
	Sig (2-tailed)	,005		,009	,002	,013	,069
PEOU	Correlation	,633**	,468**	1	,252	,607**	,149
	Sig (2-tailed)	,000	,009		,179	,000	,433
PENJ	Correlation	,420*	,551**	,252	1	,606**	,583**
	Sig (2-tailed)	,021	,002	,179		,000	,001
SP	Correlation	,599**	,450*	,607**	,606**	1	,540**
	Sig (2-tailed)	,000	,013	,000	,000		,002
PS	Correlation	,159	,336	,149	,583**	,540**	1
	Sig (2-tailed)	,402	,069	,433	,001	,002	
Min.	Corr	,625**	,360	,657**	,363*	,646**	,209
	Sig (2-tailed)	,000	,051	,000	,049	,000	,267

C. Second experiment

1) Experimental setup

In our second experiment, participants were interacting with iCat through speech. Conversational scripts were developed for the iCat in two conditions: more socially communicative and less socially communicative. The more socially communicative condition exhibited the following social abilities:

- it listened more attentively (by looking at the participant and nodding while the participant was speaking);
- it smiled during the interaction,
- it remembered and used the name of the participant during the interaction;
- it was showing more facial expressions;
- it had a more expressive voice (with variable pitch);
- it would apologize for making a mistake.

The scripted dialogues for the two conditions were identical except for the participant’s name being used by the more social version. All dialogues were set up with the same text to speech (*tts*) application.

Fig. 3. Setup iCat Wizard of Oz experiment

A specific interaction context was created where the iCat could be used in a Wizard of Oz fashion, which guaranteed a similar pattern for all sessions. The participants were first exposed to the iCat in groups (8 participants per group). After a short introduction by one of the researchers the robot told them what its possibilities were: an interface to domestic applications, monitoring, companionship, information providing, agenda-keeping and memorizing medication times and dates. They were told that for today's experiment, the robot was only programmed to perform three tasks: setting an alarm, give directions to the nearest supermarket and giving the weather forecast for tomorrow. The experimenter subsequently demonstrated how to have a conversation with the robot in which it performed these tasks.

After this group session, the participants were invited one by one to have a conversation with the robot, while the other group members were waiting in a different section of the room. The conversation was standardized as much as possible and we asked the participants to have the robot perform the three simple tasks. While being engaged in conversation, the participants' behavior was observed by a researcher and recorded by camera. The group session and the individual session were both about 5 minutes, so the maximum time spent with the robot was 10 minutes for each participant.

2) Subjects

Our experiment featured 40 participants between 65 and 89 years old, divided into 4 groups of 8 and 2 groups of 4. Exactly half of the participants (2 groups of 8, 1 group of 4) were exposed to the more sociable version and the other half to the less sociable one.

3) Procedure

A specific interaction context was created where the iCat could be used in a Wizard of Oz fashion, which guaranteed a similar pattern for all sessions. The participants were first exposed to the iCat in groups (8 participants per group). After a short introduction by one of the researchers the robot told them what its possibilities were: an interface to domestic applications, monitoring, companionship, information providing, agenda-keeping and memorizing medication times and dates. They were told that for today's experiment, the robot was only programmed to perform three tasks: setting an alarm, tell a joke and giving the weather forecast for tomorrow. The experimenter subsequently demonstrated how to have a conversation with the robot in which it performed these tasks.

After this group session, the participants were invited one by one to have a conversation with the robot, while the other group members were waiting in a different section of the room. The conversation was standardized as much as possible and we asked the participants to have the robot perform the three simple tasks. While being engaged in conversation, the participants' behavior was observed by a researcher and recorded by camera. The group session and the individual session were both about 5 minutes, so the maximum time spent with the robot was 10 minutes for each participant.

4) Results

In the second experiment the test session and the questionnaire were completed by 40 participants – 20 for each condition. In this experiment the constructs also appear to be strong, as is shown in Table 4.

TABLE 4
CRONBACH'S ALPHA FOR THE USED CONSTRUCTS IN THE SECOND EXPERIMENT

Construct	**Items**	**Alpha**
Intention to use	3	,901
Perceived usefulness	3	,865
Perceived ease of use	5	,765
Perceived enjoyment	5	,846
Perceived sociability	4	,885
Social presence	5	,831

Table 5 shows the t-test scores on the used constructs for the two conditions. There is a significant difference in acceptance score in favor of the more social condition.

TABLE 5
T-TEST RESULTS FOR THE TWO CONDITIONS

Construct	**t**	**Sig (2-tailed)**
Intention to use	-2,264	,029*
Perceived usefulness	-,470	,641
Perceived ease of use	-,928	,360
Perceived enjoyment	-2,043	,048*
Perceived sociability	-2,208	,033*
Social presence	-2,271	,029*

. Also the scores for Social Presence, Perceived Enjoyment and Perceived Sociability show a significant difference

D. Combined results

In our two experiments, the three conditions (iCat with touch screen, more social iCat, less social iCat) concerned the same type of users with the same type of functionalities, the same robot and the same questionnaire. In order to test the relationships in our hypothetical model we combined the data of the two studies to perform a linear regression analysis on the hypothesized relationships, thus examining the probability of one construct (listed under 'Independent variables') determining the other (listed under ('Dependent variable').

The results are shown in table 4.

TABLE 4
LINEAR REGRESSION SCORES FOR THE HYPOTHESIZED RELATIONSHIPS

Hypothesis	Independent variables	Dependent variable	Beta	t	Sig.
H1	ITU	Minutes	,625	4,236	,000**
H2	PS	SP	,540	3,399	,002**
H3	SP	PENJ	,606	4,033	,000**
H4	PENJ	PEOU	,163	1,366	,176
H5	PENJ	ITU	,382	3,913	,000**
	PEOU		,347	3,329	,001**
	PU		,163	1,537	,129

For H1 only the results from the first experiment (n=30) were used, for the other hypotheses the results of both experiments were added (n=70)

The results confirm our first three hypotheses. The fourth hypotheses has to be rejected according to these results – a conclusion that could already be drawn from correlation analysis of our first experiment: Perceived Enjoyment does not have a predictive influence on Perceived Ease of Use. The fifth hypothesis can only partly be confirmed: Perceived Ease of Use and Perceived Enjoyment both have a predictive influence on Intention to Use. The influence of Perceived Usefulness on Intention to Use does not show despite its correlation in the first experiment. An explanation would be that the effect is captured by the determining influence of Perceived Enjoyment and Perceived Ease of Use.

V. Discussion and Conclusion

Our results show that a robot with more social abilities has a higher score on Social Presence and this results in a higher score on Perceived Enjoyment which again leads to a higher Intention to Use the system. Our first experiment showed that this Intention to Use predicts the actual use of the system.

Summarizing our findings:

H1 Prediction of Usage by Intention to Use for this type of technology and user group has been confirmed both by correlation and regression analysis on data of our first experiment.

H2a The data from our first experiment show a strong correlation between Perceived Sociability and Social Presence. Besides, regression analysis shows Perceived Sociability to be a determining influence on Social Presence.

H2b Responses to the robot in the more sociable condition in our second experiment clearly show a higher score on Social Presence.

H3 Regression analysis shows Social Presence is a determining influence on Perceived Enjoyment.

H4 We could not confirm Perceived Enjoyment to be a determining influence on Perceived Ease of Use.

H5 Intention to Use is indeed determined by Perceived Enjoyment and Perceived Ease of Use, but the influence of Perceived Usefulness does not show in a regression analysis due to the effect being captured by the other constructs.

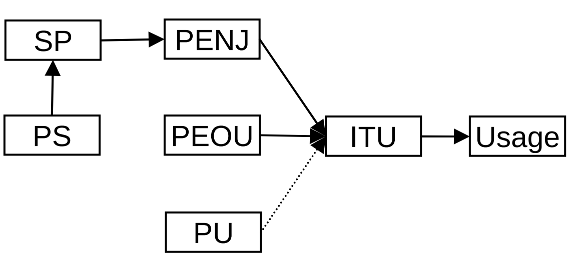

Fig. 4. Confirmed relationships

We may conclude that the sense of presence that people feel with a robot can be manipulated by changing its social abilities. This sense of presence has a positive impact on the enjoyment that is felt and this is influencing its acceptance.

Of course our experiments have been carried out with a specific non mobile robot and we have to be careful to generalize. Besides, usage data in our first experiment have been collected over a five day period: it would be interesting to see how usage patterns develop over weeks and months. Nevertheless, the significance of the data gives us some quantitative evidence that social abilities are very relevant in this context and that the 'social presence experience' is something beyond age. This means that it may be important to optimize this experience to make robots more acceptable to this specific user group.

It would be interesting to see if these conclusions are specific for this type of robot and for older users or can be generalized. Future research could focus on different robots (and perhaps screen agents) but also on different user groups and examine how the variable of user age relates to the impact of social presence. A problem that such research would face – and that we also face with our present research - is the relationship between generations and experience with advanced technology. As we found in earlier research, there is a strong relationship between experience with technology such as computers and the way new technologies are experienced and evaluated. This experience is different for every generation and findings based on research on the presently older population could very well differ from what we would find in the older population of the future.

Besides, the different ways this experience can be optimized can of course be object of study, especially for this user group.

Acknowledgment

We like to express our gratitude to the staff and caretakers as well as the test participants of "De Kiekendief", "De Overloop" in Almere, "De Uiterton" in Lelystad and "De Emtinckhof" in Loosdrecht for their cooperation. Also we like to thank Willem Notten, Bas Terwijn and Rogier Voors for their work on programming the system and Rick van Midde, Albert van Breemen and Martin Saerbeck for their support.

References

[1] M. Pollack, "Intelligent Technology for an Aging Population: The Use of AI to Assist Elders with Cognitive Impairment," AI Magazine, Summer, pp 9-24, 2005.

[2] R. Reddy, "Robotics and Intelligent Systems in Support of Society," IEEE Intelligent Systems, vol. 21, pp. 24-31, 2006.

[3] J. Forlizzi, "Robotic products to assist the aging population," Interactions, Volume 12 Issue 2, vol. march, pp. 16-18, 2005.

[4] J. Forlizzi, DiSalvo, C. and Gemperle, F., "Assistive Robotics and an Ecology of Elders Living Independently in Their Homes," Journal of HCI Special Issue on Human-Robot Interaction, January, pp. 25-59, 2004.

[5] E. D. Mynatt, I. Essa, and W. Rogers, "Increasing the opportunities for aging in place," Proceedings of the CUU 2000 Conference on Universal Usability, New York, 2000.

[6] W. Taggart, S. Turkle, and C. Kidd, "An interactive robot in a nursing home: Preliminary remarks," in Towards Social Mechanisms of Android Science, vol. July. Stresa, Italy: Cognitive Science Society, pp. 56-61, 2005.

[7] L. M. Camarinha-Matos and H. Afsarmanesh, "Design of a virtual community infrastructure for elderly care," presented at 3rd IFIP Working Conference on Infrastructures for Virtual Enterprises, 2002.

[8] M. Heerink, B. J. A. Kröse, B. J. Wielinga, and V. Evers, "Studying the acceptance of a robotic agent by elderly users," International Journal of Assistive Robotics and Mechatronics, vol. 7, pp. 33-43, 2006.

[9] M. Heerink, B. J. A. Kröse, B. J. Wielinga, and V. Evers, "Enjoyment, Intention to Use and Actual Use of a Conversational Robot by Elderly People," Proceedings of The third ACM/IEEE International Conference on Human-Robot Interaction, Amsterdam, March 2008.

[10] C. Breazeal, "Socially intelligent robots," interactions, vol. 12 (March), 2005.

[11] B. de Ruyter, P. Saini, P. Markopoulos, and A. J. N. van Breemen, "Assessing the Effects of Building Social Intelligence in a Robotic Interface for the Home," Special Issue of IwC: social impact of emerging technologies, 2005.

[12] B. R. Duffy, "Anthropomorphism and The Social Robot," Special Issue on Socially Interactive Robots, Robotics and Autonomous Systems, Robotics and Autonomous Systems, vol. 42, pp. 170-190, 2003.

[13] T. Shibata, K. Wada, and K. Tanie, "Statistical Analysis and Comparison of Questionnaire Results of Subjective Evaluations of Seal Robot in Japan and U.K," presented at Proceedings of the 2003 IEEE International Conference on Robotics & Automation, 2003.

[14] K. Wada and T. Shibata, "Robot Therapy in a Care House - Results of Case Studies," presented at Proceedings RO-MAN, 2006.

[15] W. D. Stiehl, J. Lieberman, C. Breazeal, L. Basel, R. Cooper, H. Knight, L. Lalla, A. Maymin, and S. Purchase, "The Huggable: A Therapeutic Robotic Companion for Relational, Affective Touch," presented at IEEE Consumer Communications and Networking Conference, 2006.

[16] J. Pineau, M. Montemerlo, M. Pollack, N. Roy, and S. Thrun, "Towards robotic assistants in nursing homes: Challenges and results," Robotics and Autonomous Systems, vol. 42, pp. 271-281, 2003.

[17] M. E. Pollack, L. Brown, D. Colbry, C. Orosz, B. Peintner, S. Ramakrishnan, S. Engberg, J. T. Matthews, J. Dunbar-Jacob, C. E. McCarthy, S. Thrun, M. Montemerlo, J. Pineau, and N. Roy, "Pearl: A Mobile Robotic Assistant for the Elderly," in AAAI Workshop on Automation as Eldercare, 2002.

[18] B. Graf, M. Hans, and R. D. Schraft, "Care-O-bot II—Development of a Next Generation Robotic Home Assistant," Autonomous Robots, vol. 16, pp. 193-205, 2004.

[19] C. Parlitz, W. Baum, U. Reiser, and M. Hagele, "Intuitive Human-Machine-Interaction and Implementation on a Household Robot Companion," Lecture notes in computer science, vol. 4557, pp. 922, 2007.

[20] A. Cesta, G. Cortellessa, F. Pecora, and R. Rasconi, "Supporting Interaction in the Robo Care Intelligent Assistive Environment," Proceedings of AAAI Spring Symposium on Interaction Challenges for Intelligent Assistants, pp. 18-25, 2007.

[21] F. D. Davis, "Technology Acceptance Model for Empirically Testing New End-user Information Systems Theory and Results," in Unpublished Doctoral Dissertation: MIT, 1986.

[22] F. D. Davis, "Perceived usefulness, perceived ease of use, and user acceptance of information technology," MIS Quarterly, vol. September, 1989.

[23] V. Venkatesh, M. G. Morris, G. B. Davis, and F. D. Davis, "User Acceptance of Information Technology: Toward a Unified View," MIS Quarterly, vol. 27, pp. 425-478, 2003.

[24] R. Looije, F. Cnossen, and M. A. Neerincx, "Incorporating Guidelines for Health Assistance into a Socially Intelligent Robot," presented at Proceedings RO-MAN, 2006.

[25] H. Van der Heijden, "User acceptance of hedonic information systems," MIS quarterly, vol. 28, 2004.

[26] H. Sun and P. Zhang, "Causal Relationships between Perceived Enjoyment and Perceived Ease of Use: An Alternative Approach," Journal of the Association for Information Systems, vol. 7, 2006.

[27] M. Y. Yi and Y. Hwang, "Predicting the use of web-based information systems: self-efficacy, enjoyment, learning goal orientation, and the technology acceptance model.," International Journal of Human-Computer Studies vol. 59, 2003.

[28] T. Chesney, "An Acceptance Model for Useful and Fun Information Systems," Human Technology, vol. 2, 2006.

[29] B. Reeves and C. Nash, The Media Equation: How People Treat Computers, Televisions, and New Media as Real People and Places. New York: Cambridge University Press, 1996.

[30] T. Bickmore and D. Schulman, "The Comforting Presence of Relational Agents," presented at Proceedings of CHI, 2006.

[31] C. F. DiSalvo, F. Gemperle, J. Forlizzi, and S. Kiesler, "All robots are not created equal: the design and perception of humanoid robot heads," presented at Proceedings of the conference on Designing interactive systems: processes, practices, methods, and techniques, 2002.

[32] K. M. Lee and C. Nass, "Designing social presence of social actors in human computer interaction," presented at Proceedings of the SIGCHI conference on Human factors in computing systems, 2003.

[33] B. G. Witmer and M. J. Singer, "Measuring Presence in Virtual environments: A Presence Questionnaire," Presence, vol. 7, pp. 225-240, 1998.

[34] M. Lombard and T. B. Ditton, "At the heart of it all: The concept of presence," Journal of Computer-Mediated-Communication, vol. 3, 1997.

[35] K. Dautenhahn, "Roles and functions of robots in human society: implications from research in autism therapy," Robotica, Volume 21 Issue 4, vol. august, 2003.

[36] T. Fong, I. Nourbakhsh, C. Kunz, L. Fluckiger, and J. Schreiner, "The Peer-to-Peer Human-Robot Interaction Project," presented at AIAA Space, 2005.

[37] T. Fong, I. Nourbakhsh, and K. Dautenhahn, "A survey of socially interactive robots," Robotics and Autonomous Systems, vol. 42, pp. 143-166, 2003.

[38] C. Breazeal, "Towards sociable robots," Robotics and Autonomous Systems, vol. 42, 2003.

[39] F. M. Gresham and S. N. Elliot, Social skills rating system: Circle Pines: American Guidance Service, 1990.

[40] K. Shinozawa, F. Naya, J. Yamato, and K. Kogure, "Differences in Effect of Robot and Screen Agent Recommendations on Human Decision-Making," IJHCS, vol. 62, pp. 267-279, 2005.

[41] T. W. Bickmore, L. Caruso, K. Clough-Gorr, and T. Heeren, "'It's just like you talk to a friend' relational agents for older adults," Interacting with Computers, vol. 17, pp. 711-735, 2005.

Inter-Classifier Feedback for Human-Robot Interaction in a Domestic Setting

Juhyun Lee, W. Bradley Knox, and Peter Stone

***Abstract*—For a mobile robot that interacts with humans such as a home assistant or a tour guide robot, activities such as locating objects, following specific people, and distinguishing among different people are fundamental, yet challenging robotic vision tasks. For complex object recognition and tracking tasks such as person recognition and tracking, we use the method of inter-classifier feedback to track both uniquely identifying characteristics of a person (e.g. face), and more frequently visible, but perhaps less uniquely identifying characteristics (e.g. shirt). The inter-classifier feedback enables merging multiple, heterogeneous sub-classifiers designed to track and associate different characteristics of a person being tracked. These heterogeneous sub-classifiers give feedback to each other by identifying additional online training data for one another, thus improving the performance of the overall tracking system. We implement the tracking system on a Segway base that successfully performed aforementioned activities to a second place finish in the RoboCup@Home 2007 competition. The main contribution is a complete description and analysis of the robot system and its implemented algorithms.**

***Index Terms*—Robotics vision, Human-robot interaction, RoboCup@Home**

I. Introduction

With the growing possibility of and demand for robots interacting in real-world environments, it is becoming increasingly important for robots to be able to interact with people. For robust human interaction in domestic environment, the robot must be able (1) to locate and identify common objects, (2) to follow people or guide people to places of interest, and (3) to distinguish the set of people with whom it commonly interacts while also successfully identifying strangers. In RoboCup@Home, an international competition designed to foster research on such interactive domestic robots, the robot has to show its performance in these tasks [1].

This paper presents the UT Austin Villa RoboCup@Home 2007 entry, a Segway-based robot and the second-place finisher in the competition. The robot demonstrated its ability to complete versions of all three of the tasks mentioned above. The main contribution of this paper is a complete description of the robot system and its implemented algorithms which enabled the robot's successful human-robot interaction in this broad, challenging, and relevant event, with an emphasis on the key component of our person recognition algorithm.

Detecting and/or tracking a particular person among multiple persons can be challenging for three reasons. The first reason is the noisy data. A person's most uniquely identifying visual feature is his or her face, which is not always present in a given video frame. Even if it is present, face detection algorithms may fail due to motion blur or bad lighting. The second reason is the demanding constraints of the task. Because a robot needs to operate in real-time with its limited processing power shared among all its tasks, the computational resources available for person tracking are constrained, thus limiting the algorithms that may be considered. The third reason is the mobile nature of the robot. The robot may only get to see a very limited view of a person under one lighting condition when it is trained. Worse, the trained characteristics of the person can change over space and time, due to pose and illumination changes. Then, the robot must be able to detect such changes autonomously and select new training data for its classifiers.

We use inter-classifier feedback for person tracking in a video stream that uses face recognition as a starting point, but augments it with tracking of more frequently visible, but perhaps less uniquely identifying characteristics such as the person's clothes. The main idea is that primary, uniquely identifying characteristics (e.g. faces) can be dynamically associated with secondary, ambiguous, possibly transient, but more easily computable characteristics (e.g. shirt colors). When primary characteristics are identifiable, they are re-associated with the secondary characteristics currently visible on the person. The secondary characteristics can then be used to track the person, even when the primary characteristics are not detected. We also show how each classifier helps the other classifiers to update their training data online to improve the overall performance of the system.

The main technical focus of our approach was on person tracking and recognition. As such, we focus in detail in this article on our algorithms for these tasks, including a novel method of combining classifiers of multiple characteristics of the person. The tasks our robot performed also required object tracking, for which we use the previously mentioned ARTags [14]. We summarize our use of ARTags in the context of their task-specific uses in Section VI.

The remainder of this paper is organized as follows. Section II describes the RoboCup@Home 2007 competition including its goals and format. After the motivation for a strong person tracking algorithm for mobile robots, we introduce the concept of heterogeneous inter-classifier feedback in domain-independent terms in III. We provide a proof-of-concept with a simple person tracker that we used in the competition in Section IV. Section V introduces the UT Austin Villa robot, including both hardware and software systems. Section VI

Juhyun Lee, W. Bradley Knox, and Peter Stone are with the Department of Computer Sciences, the University of Texas at Austin, Austin, TX 78712, the United States of America.

E-mail: {impjdi,bradknox,pstone}@cs.utexas.edu

Fig. 1. RoboCup@Home 2007 setting.

describes our specific solutions for each task and our respective performances in them. Section VII discusses related works and Section VIII concludes the paper.

II. RoboCup@Home 2007

RoboCup@Home is an international research initiative that aims "to foster the development of useful robotic applications that can assist humans in everyday life" [1]. The eventual goal is to create fully functional robots that can assist humans at home in a variety of ways, performing any function that humans are currently hired to do, including assisted living and nannying. The RoboCup@Home community created a compelling and challenging set of tasks for the first year of the event in 2006 and raised the bar in 2007 [33].

In the 2007 competition, robots in a living room and kitchen environment (Fig. 1) had to complete up to four of six specified tasks. These tasks can be considered fundamental building blocks toward the complex behavior and capabilities that would be required of a fully functional home assistant robot. The specific tasks are described in Fig. 2.

Within each task, there were two levels of difficulty. The easier level, called the first phase, existed as a proof of concept and often abstracted away part of the problem (e.g. object recognition or mapping and navigation). The second, more difficult phase of each task was structured similarly to how the task would need to be performed in a real domestic setting. During each phase, there was a ten minute time limit to complete the task objectives.

After the specific tasks, all teams performed a free-form demonstration in what was called the *Open Challenge*, during which they showed off their most impressive technical achievements to a panel of other team leaders. Each event was scored and five teams advanced to the Finals. In the Finals, the five finalists performed demonstrations for trustees of the RoboCup organization, who determined the final standings.

Our robot attempted three of the six possible RoboCup@Home tasks. These tasks were *Lost and Found*, *Follow and Guide a Human*, and *Who Is Who?*. Each task is described in the following subsections. Our specific approaches to the three tasks are detailed in Section VI.

Task	**Description**
Navigate	Navigate to a commanded location
Manipulate	Manipulate one of three chosen objects
Follow and Guide a Human	Follow a human around the room
Lost and Found	Search for and locate previously seen objects
Who Is Who?	Differentiate previously seen and unseen humans
Copy-cat	Copy a human's movement in a game-like setting

Fig. 2. List of RoboCup@Home 2007 tasks.

A. Lost and Found

This task tested a robot's ability to find an object that had been "lost" in the home environment. We competed in only the first phase of the *Lost and Found* task. In that phase, a team would hide a chosen object somewhere in the living environment at least five meters from their robot and out of its view. If the referees approved the location, the task began. The task ended successfully when the robot had moved within 50 cm of the item and had announced that it found it.

B. Follow and Guide a Human

In *Follow and Guide a Human*, a robot followed a designated human as he or she walked throughout the home and then, optionally, returned to the starting position (thus "guiding" the human).

1) First Phase: A team member led his or her robot across a path determined by the competition referees. The leader was permitted to wear any clothing or markers he chose. Once the leader and the robot reached the destination, an optional extension was to have the robot return back to the starting point with the human-following.

2) Second Phase: The rules were the same except that the human leader was a volunteer chosen from the audience. Therefore, the algorithm needed to robustly identify a person without markers or pre-planned clothing.

C. Who Is Who?

The *Who Is Who?* task tested person-recognition capabilities on a mobile robot. Both phases of the task involved the robot learning to recognize four people, the referees rearranging the people and adding one new person (a "stranger"), and the robot subsequently identifying the four known people and the stranger accurately.

1) First Phase: The four people lined up side-to-side while a robot moved among them and learned their appearances and names. Once the robot finished training, the four people and a stranger were arranged into a new order by the referees. Then, the robot again moved among the people, announcing their names as each was identified. One mistake was allowed.

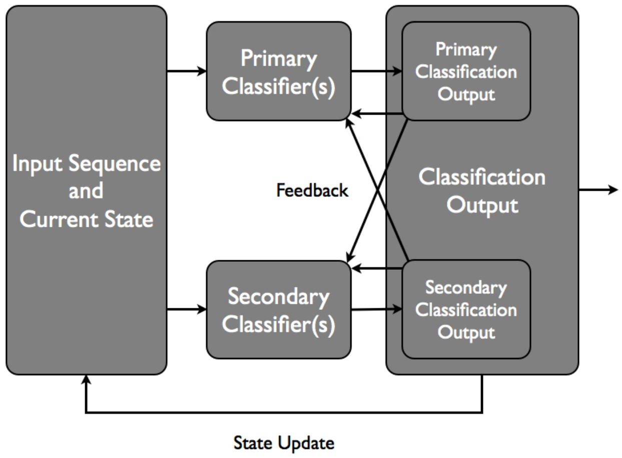

Fig. 3. Classification with heterogeneous inter-classifier feedback

2) Second Phase: The second phase was much like the first, but after the robot finishes training, the four known people and the stranger were placed by the referees in various locations around the entire living room and kitchen environment. The robot then had to search them out *and* correctly identify them.

III. Classification with Heterogeneous Inter-Classifier Feedback

A robust person tracking and recognition algorithm was essential in order to do well in two out of three tasks we decided to attempt. Before moving on to the implementation details of the person tracker we used, we describe the concept of heterogeneous inter-classifier feedback in domain-independent terms.

The overall system is a learning system which takes its current state and a part of the input sequence to compute its output and update its current state. During the output computation, an overall classifier is used which is built up from two or more heterogeneous sub-classifiers. Each sub-classifier solves its own classification problem by extracting different characteristics from the same input.

We divide the characteristics into two groups: primary and secondary. A primary characteristic must be a unique one that identifies a class. The classification problem of such primary characteristic may be computationally expensive, or susceptible to noisy input data. A secondary characteristic may be ambiguous, but computationally less expensive and more robust with respect to noise. Secondary characteristics can be introduced to leverage the shortcomings of a classification solely based on primary characteristics. This is also one of the main differences between our method and an ensemble. A secondary classifier is not used to vote for a better answer in case of an ambiguous classification result, but as a fall-back classifier for the times when the primary classifier returns no answer. There can be multiple characteristics in the same level, or more levels of characteristics may be introduced if the inter-characteristic relationship can be well-defined. Fig. 3 illustrates our scheme.

Algorithm 1 Classification with heterogeneous inter-classifier feedback (with 1 primary and 1 secondary classifier)

Require: $Input$: Input sequence, $State$: Current state
1: $SecChar \leftarrow ExtractSecChar(Input)$
2: $SecClass \leftarrow ClassifySecChar(SecChar)$
3: **if** $(IsPriCharRequired(State) =$ **true**$)$ **then**
4:     $PriChar \leftarrow ExtractPriChar(Input)$
5:     $PriClass \leftarrow ClassifyPriChar(PriChar)$
6: **else**
7:     $PriClass \leftarrow \emptyset$
8: $Class \leftarrow \emptyset$
9: **if** $(PriClass \neq \emptyset)$ **then**
10:     $Class \leftarrow PriClass$
11:     **if** $(SecClass \neq \emptyset)$ **then**
12:         **if** $(PriClass \neq SecClass)$ **then**
13:             **if** $(PriClass.Confidence > SecClass.Confidence)$ **then**
14:                 $TrainSecChar(SecChar, Class)$
15:             **else**
16:                 $Class \leftarrow SecClass$
17:                 $TrainPriChar(PriChar, Class)$
18:     **else**
19:         $TrainSecChar(SecChar, Class)$
20: **else if** $(SecClass \neq \emptyset)$ **then**
21:     $Class \leftarrow SecClass$
22: Update $State$
23: **return** $Class$

Alg. 1 shows the basic structure of the algorithm we propose. $ExtractPriChar$ and $ExtractSecChar$ extract and return primary and secondary characteristics, respectively, of a given raw input. The returned characteristics are fed into each characteristic's classifiers $ClassifyPriChar$ and $ClassifySecChar$, respectively, which return the class label of the input. $TrainPriChar$ and $TrainSecChar$ are procedures for training the primary and the secondary classifier, respectively, with the training data and the class label. Finally, $IsPriCharRequired$ is a simple helper function that determines whether the heavy primary classifier should be run in the given cycle for performance reasons.

The computationally cheap, and thus more frequently invocable, secondary classifier can be used as the default (lines 1–2), while the more expensive primary classifier is invoked whenever a more accurate classification is needed (lines 3–7). If the condition of taking the branch is carefully chosen, near real-time performance can be achieved by avoiding a large classification expense each cycle. In case of a mismatch of the class labels returned by each classifier (line 12), the algorithm picks the class label with higher confidence depending on each characteristic's classification accuracy and/or $State$. Lines 14, 17, and 19 comprise the inter-classifier feedback which improves the classification performance of each classifier by adding more training data to the other class. In case all sub-classifiers do not return an answer, the overall classifier does not return an answer either. Our scheme does not try to find an answer if an answer cannot be determined from its sub-classifiers. However, our scheme still performs better than a

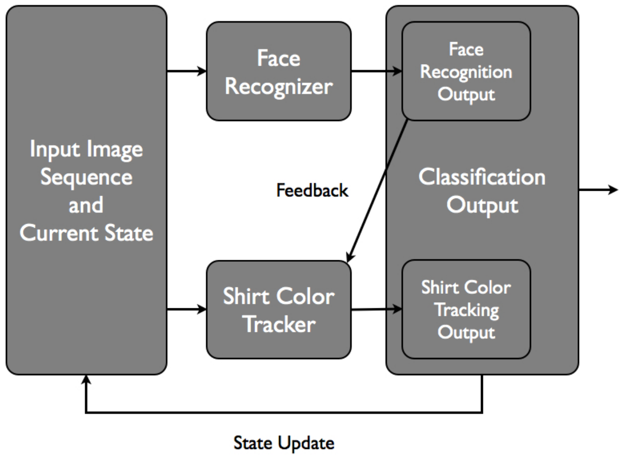

Fig. 4. Person tracking with 1 primary and 1 secondary classifier

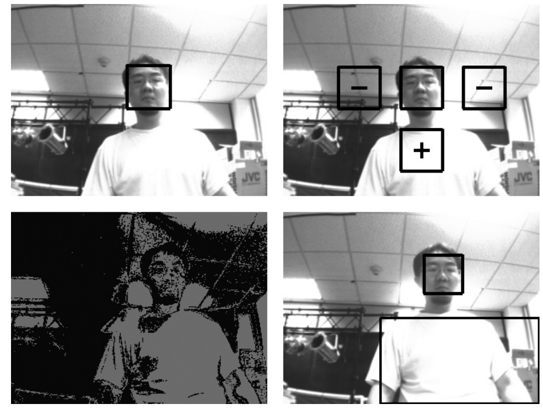

Fig. 5. Once the face is detected (a), the face's SIFT features are extracted to the face database and positive and negative regions of the shirt are sampled (b). The RGB-to-person mapping generated with the positive and negative histograms are shown in (c), and the shirt is detected in (d).

primary classifier alone.

IV. Person Tracking with Heterogeneous Inter-Classifier Feedback

Having discussed the general concept of heterogeneous inter-classifier feedback, next we apply the algorithm to a person tracking task. Since faces are unique, the primary characteristic for the person tracking task can be chosen to be the face. Since tracking the face alone is not sufficient to robustly track the person for previously mentioned reasons, a secondary characteristic of a person which is independent from the primary characteristic is chosen. Among different candidate characteristics, we choose the shirt of a person to be the secondary characteristic because it is easily visible, unless he or she is completely occluded by other objects. Fig. 3 is implemented for our domain as shown in Fig. 4.

A. Primary Characteristic Tracking

We divide the primary characteristic tracking task in two: the face detection and the face recognition. These correspond to $ExtractPriChar$ and $ClassifyPriChar$ in Algorithm 1, respectively. The face detection algorithm we use for the task is a boosted cascade of Haar-like features as discussed in [29]. It is implemented in the Intel Open Source Computer Vision Library, and shows a near-real-time performance (15 Hz) using limited resolution (160×120) images with our tablet PC. Extracting rectangular features from integral images as described in [29] does not suffer from a slight resolution decrease. The face recognition algorithm which extracts scale-invariant features (SIFT) [22] from cut-out face images suffers more from a resolution decrease. Rather than clipping the faces from the small 160×120 image used for the face detection, we extract the corresponding region in the original 640×480 image and extract the SIFT features of that region. These are used to distinguish among different faces by counting the number of matches during the recognition phase.

B. Secondary Characteristic Tracking

The secondary characteristic, a person's shirt, is trained when that person's face is successfully classified for several (e.g. 10) frames. Each person has his or her own positive and a negative histogram each with a size of $64 \times 64 \times 64$ RGB bins that contains the color information of the shirt the person is wearing. For example, a shirt with red and green stripes has high counts in $(63, 0, 0)$ and $(0, 63, 0)$. Fig. 5 shows which regions in an image are scanned for positive and negative samples of the shirt. Positive samples of the shirt colors are taken from a region as large as the face's bounding box, located 0.5 bounding boxes below the face. Negative samples are taken from two regions each as large as the face's bounding box, located 0.5 bounding boxes left and right of the face which should be the background or other objects in the scene. By maintaining positive and negative samples separately, a more accurate RGB-to-person mapping can be generated than by generating the mapping with positive samples alone. With this sampling scheme, we assume that the color of the shirt is relatively uniform in direction, i.e. we do not consider shirts having different colors in the front and in the back but we do not assume constant-colored shirts. We assume that each person has a distinctly colored shirt. In case there is more than one person having similarly colored shirts, the shirt of the latest person of interest is recorded, and the corresponding RGB values are mapped to that person.

To detect the shirt of a person in a given scene, we map each RGB pixel to a person ID with the mapping generated as described in the previous paragraph, and find the largest continuous blob containing only 1 ID. This approach is a modification of color-blob segmentation [26] where the colors of interest are assigned the same label. The blob detection and recognition algorithm is a lightweight operation that is carried out in real-time, 25 to 30 frames per second with a 320×240 resolution image. A more sophisticated algorithm such as edge detection may also be applied, but it requires additional object classification which needs a computation close to the face recognition itself (e.g. the Canny edge detector runs in 15 Hz) which is not desirable for tracking a weaker characteristic.

Another SIFT matching algorithm could have been chosen to distinguish shirts, but we found the color information of shirts yields better classification than the gray-scale SIFT features.

C. Adaptive Characteristic Tracking Algorithm Selection

Heavier vision processing is undesirable, since it results in lower frame rates which leads to less reactive robot behavior. We use an adaptive characteristic selection scheme for the robot's vision to achieve a higher frame rate. By the nature of human motion, the face is either constantly visible if facing the camera with limited movement, or constantly unrecognizable or occluded if not facing the camera or moving rapidly, although there can be a transition period between the two states. The face detection algorithm we use shows an average frame rate of 15 Hz. If the face detector can be skipped every other frame without decreasing the detection rate, the average frame rate would increase up to 22.5 Hz. Referring back to Algorithm 1, $IsPriCharRequired$ is defined as "every other frame". To avoid compromising the person detection rate, the secondary shirt detector has to show an equal or better detection rate than the face detector. We found this to be true in relatively steady lighting conditions.

D. Autonomous Real-Time Training Data Selection

Although we introduce the notion of primary and secondary characteristics indicating the different weights of each characteristic, there is no guarantee that a lower weighted characteristic will positively impact other characteristics, and vice versa. The primary tracking system can give feedback to the secondary tracking system to choose new training data for accurate classification. In our person tracking application, the face recognizing algorithm which computes scale-invariant features in normalized gray-scale images is more robust to color changes caused by ambient brightness changes. On the other hand, the RGB-to-person mapping used for shirt tracking is highly susceptible to such changes. If a person's face is correctly recognized, but the shirt is not detected, the RGB-to-person mapping can re-learn the shirt's colors, or update the RGB values for better classification under the changed lighting condition.

Since SIFT features are sensitive to directed lighting, a person moving in an indoor environment may be classified as a different person where there is more directed lighting than ambient lighting. However, the shirt's colors sampled with a Gaussian distribution has a slightly wider range in this case, and thus is still visible with directed lighting. Since the shirt is already known to belong to a certain person, the false-negative unknown face is then added to the training data of the primary classifier. Although conceptually possible, we decided not to integrate the re-training of the face recognizer on our laptop. The re-computation of the probability density function in our face recognizer takes more than 3 seconds on our robot-mounted laptop and less than 1 second on a 2 GHz dual-core laptop. We found that the robot operates more smoothly without the re-training, since it does not have to stop frequently for the PDF computation. The effect of autonomous real-time training data selection is shown in the Finals described in Section VI.

Fig. 6. UT Austin Villa home assistant robot.

V. PLATFORM

This section introduces the hardware and software systems of the UT Austin Villa RoboCup@Home 2007 entry, shown in Fig. 6. The robot consists of a Segway Robotic Mobility Platform (RMP) 100[1], supporting an on-board computer and various sensors. No other team used a Segway as its robotic platform. The Segway provides controlled power in a relatively small package. This suits a domestic environment well, for it is small enough to maneuver a living environment built for humans and powerful enough to reliably traverse varying indoor terrain including rugs, power cords, tile, and other uneven surfaces. The large wheels easily navigate small bumps that challenged other indoor robots during the competition.

The two-wheeled, self-balancing robot reaches speeds up to six mph, exerts two horsepower, and has a zero turning radius, freeing it from worry about getting out of tight corners and corridors. The Segway moves with two degrees of freedom, receiving motion commands in the form of forward velocity (m/sec) and angular velocity (radians/sec). It provides proprioceptive feedback in the form of measurements of odometry and pitch. With a payload capacity of 100–150 lbs., the Segway could easily carry several times the weight of its current load.

A 1 GHz Fujitsu tablet PC sits atop the Segway platform, performing all sensory processing, behavior generation, and the generation of motor commands on-board. It interfaces with the Segway via USB at 20 Hz.

[1]http://www.segway.com/rmp/

Two cameras and one laser range finder are available to sense the robot's environment. The Videre Design STOC camera[2] provides depth information, but is not used for the tasks and experiments described in this paper. Higher picture quality is obtained by the second camera, an inexpensive Intel webcam which sends 30 frames per second. The attached Hokuyo URG-04LX[3] is a short range, high resolution laser range finder that is well-suited for indoor environments. It collects 769 readings across 270° at 10 Hz. Also, a Logitech microphone and USB speakers are attached.

The Segway RMP 100 is based on the p-Series Segway line for human transport. Despite its power, the robot is quite safe, featuring safety mechanisms such as automatic shut-off, an emergency kill rope, and speed caps at both the hardware and software levels.

A multi-threaded program, written from scratch, operates the robot. The program's structure can be divided into five modules: the camera input processing, the laser range finder input processing, the motion input/output, speech output, the high-level behavior unit, and the GUI.

VI. Approach and Performance

This section describes the strategies and algorithms the Segway used in the tasks described in Section V. All tasks were performed in the same home environment (Fig. 1).

A. Lost and Found

In *Lost and Found*, a robot searched for a known object that had been placed in an unknown location in the home environment. The task setup is described in Section II. Our robot competed in the first phase of *Lost and Found*.

1) First Phase: We chose to use an ARTag marker as the target object [14]. ARTag is a system of 2D fiducial markers and vision-based detection. The markers are robustly detected from impressive distances (more than 5 m at 320 × 240 resolution in our lab with a 20 cm×20 cm marker) with varying light and even partial occlusion. Each marker is mapped to an integer by the provided software library. We did not observe any false positives from our ARTag system.

For the *Lost and Found* task, our robot searched the environment using a reflexive, model-free algorithm that relied on a fusion of range data and camera input. The Segway moved forward until its laser range finder detects an obstacle in its path. It would then look for free space, defined as an unoccupied rectangular section of the laser plane 75 cm deep and a few centimeters wider than the segway, to the left and right and turned until facing the free space. If both sides were free, the robot randomly chose a direction. If neither side was free, it turned to the right until it found free space. Algorithmically, free space was determined by a robustly tuned set of pie pieces in the laser data which overlapped to approximate a rectangle (see Fig. 7).

We placed the object on a table at the opposite end from where the Segway began. A straight line between the two would have passed through a television, shelves, and a kitchen table. The robot had neither prior knowledge of the object's location nor any model of the environment. The Segway successfully completed its search with more than three minutes to spare. Of the six teams that attempted *Lost and Found*, only three teams, including our team, completed it.

[2]http://www.videredesign.com/vision/stereo_products.htm
[3]http://www.hokuyo-aut.jp/02sensor/07scanner/urg.html

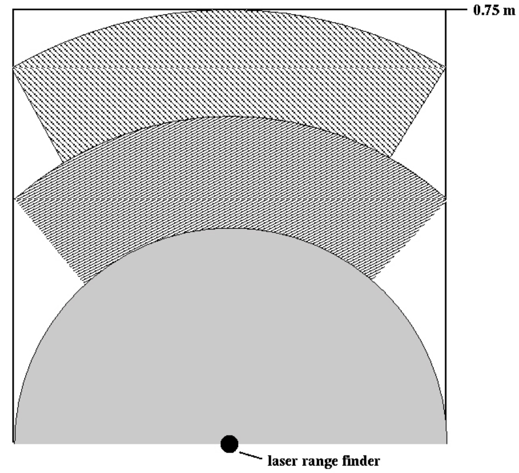

Fig. 7. The laser range finder data is checked for occupancy at three different ranges and angles to approximate a rectangle. The rectangle was a bit wider and deeper than the Segway for safety.

B. Follow and Guide a Human

In this task, a robot followed behind a human as he or she walked around the home environment, winding around the furniture. Its setup is described in Section II. The Segway attempted both the first and second phases of the *Follow and Guide a Human* task.

1) First Phase: We attempted only the following (not guiding) portion of this first phase. We did not attempt the extension because time constraints and technical difficulties left the Segway without functional mapping software. (No team finished the extension of returning back to the starting point.) Again, we used an ARTag marker on the shirt of the leading human. The robot flawlessly followed the human leader, without touching furniture or the human. Six of eight teams that attempted the first phase of *Follow and Guide a Human* completed this portion of the task.

2) Second Phase: Without the ARTags of the first phase, the robot instead trained and used a shirt classifier as described in Section IV (Fig. 8). Since we anticipated following a human with his back turned, and thus never return a positive classification result, the face recognition component of our person recognition algorithm was not used. This is an example of the secondary classifier acting as a fall-back classifier, when the primary classifier does not return any result (Fig. 9).

In the competition, the referees chose an African-American volunteer wearing a white shirt. This choice presented two problems that each were sufficient to break our algorithm.

The first problem was that the Viola and Jones' face detection algorithm was unable to detect the human's dark-skinned face. The face detector extracts contrast-based features

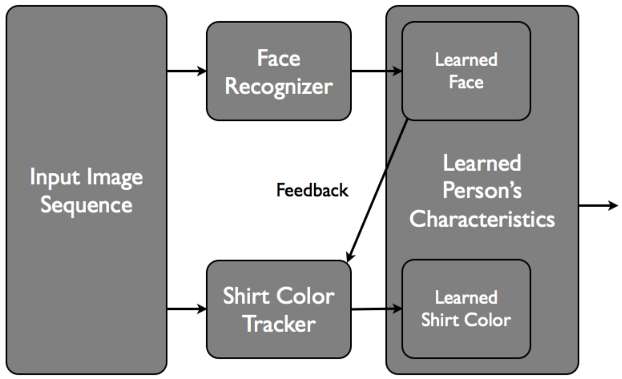

Fig. 8. Learning a person for the *Follow and Guide a Human* task and the *Who Is Who?* task. The face learner indicates the location of the face in the image, and thus the location of the person's shirt. The shirt color tracker can then learn the person's. Note that the state update arrow from Fig. 4 is removed since no motion cue is involved during training.

from a potential face location in the image and uses those features to classify the location as a face or not. However, the lighting was too dark for the detector to capture both bright and dark regions in a dark-skinned face. Without detecting a face, the robot merely waited to see one. We restarted the robot with a LED flashlight attached below the camera to add contrast to the volunteer's face. This time, the face detector managed to locate the face in the video images, and the shirt classifier learned the white shirt.

Tracking the volunteer's shirt was also problematic. His white shirt blended with the background, much of which was white as well. Collecting negative samples helps discriminate between similar colors to some extent, but the shirt and large elements of the background were too alike for the algorithm to handle. Instead of tracking the volunteer's shirt as intended, the robot classified a large portion of the wall as the person and was unable to follow the volunteer.

The choice of volunteer revealed weaknesses in our shirt-following algorithm. However, in the *Open Challenge* and Final rounds, we demonstrated that, given a human leader with light to moderately dark skin and a shirt color that is distinguishable from the background colors, the robot could follow a person for whom it has no a priori data.

C. *Who Is Who?*

The *Who Is Who?* task tested a mobile robot's ability to meet and later recognize humans. To learn the faces of multiple people, we train a face classifier for each person as described in Section IV. For *Who Is Who?*, the output of the multiple-face classifier is the set of identities which had a number of SIFT feature matches above an empirically determined threshold. If the output set is empty, then the threshold is lowered and the classifier is re-run.

Given the set of candidate identities, a shirt classifier takes over. This classifier gathers samples as described in Section IV, but otherwise the shirt classifier is different, having been modified to eliminate blob selection. Since the face is required for classification in this task, the shirt pixels are simply taken from below the face as in training. For each candidate identity, the Euclidean distance between the average RGB values of the pixels on the persons shirt (a 3-tuple) and the average RGB values of the specific identity's shirt samples is calculated. If at least one candidate's shirt distance is above a shirt threshold, then the candidate with the shortest distance is chosen as the identity of the person. If none are above the shirt threshold, the person is announced as a stranger. This is an example of the secondary classifier being a fall-back classifier in case the primary characteristic based classification result is not confident enough (Fig. 10).

1) First Phase: In the first phase, we chose the four people and their shirts. We gave them strongly distinguishable shirt colors – red, green, blue, and yellow. Our robot correctly identified four of the five people. The stranger was misidentified as one of the known people.

We believe this error occurred specifically on the stranger for two reasons. First, the volunteer's SIFT features matched many features of at least one of the known people. Second, the volunteer's shirt was colored similarly to the person whose SIFT features were similar. With both the primary characteristic (the face) and the secondary characteristic (the shirt) testing as false positives, the person tracker did not correctly classify the stranger.

Of seven teams that attempted this task, some of which used commercial software packages, only one other received points by identifying at least four of the five people.

2) Second Phase: The training of the second phase is the same as in phase one, except the persons were chosen randomly by the committee. The testing is especially more challenging in the second phase. The five people (four known and one stranger) are not standing in a line anymore, but are instead randomly distributed throughout the home environment.

As in the *Lost and Found* task, we used a stochastic search to look for candidate people as recognized by positive identification from the face detection module. During the allotted time, the robot found one of the people and correctly identified him. No other team identified a single person during the second phase.

D. Open Challenge

Once all teams had attempted their specific tasks, each competed in what was called the *Open Challenge*. This consisted of a presentation and free-form demonstration. Going into this event, after receiving scores from the specific tasks, UT Austin Villa ranked third of eleven. A jury of the other team's leaders ranked us second for the *Open Challenge*. The robot's demonstration was a simplified version of the one performed in the Finals, so it will not be described.

E. Finals

The top five teams competed in the Finals. Having ranked third in the specific tasks and second in the open challenge, UT Austin Villa advanced along with Pumas from UNAM in

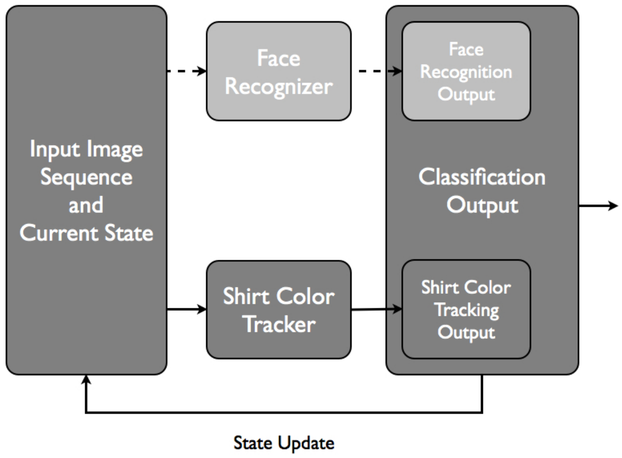

Fig. 9. Tracking a person for the *Follow and Guide a Human* task. The face recognizer is disabled, because the person will never show his face during this phase. The system relies on the shirt tracker only. Otherwise, the scheme is similar to Fig. 4.

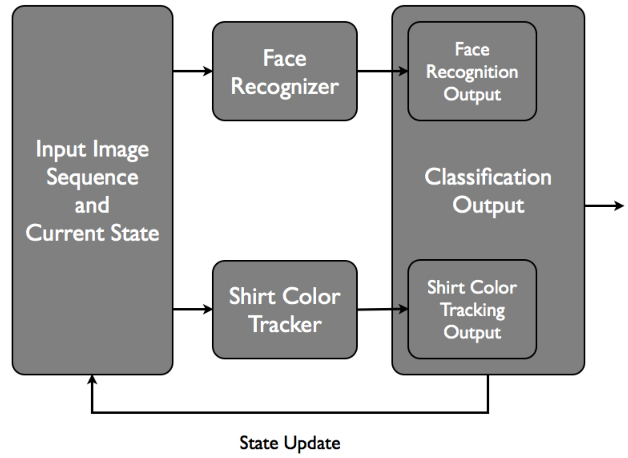

Fig. 10. Recognizing a person for the *Who Is Who?* task. Note that the feedback is disabled, since no re-training is desired. Otherwise, the scheme is similar to Fig. 4.

Mexico, AllemaniACs from RWTH Aachen in Germany, RH2-Y from iAi in Austria, and Robot Cognition Lab from NCRM in France. The Finals were judged by a panel of trustees of the RoboCup organization, all well-known robotics researchers.

Before describing the demonstration itself, we begin with some motivation for the scenario we focused on. Accurate person-recognition will be essential in any fully functional home assistant robot. For instance, if a person refers to himself or another by a pronoun (i.e. "Get *me* my medicine."), the robot needs to know who is being referenced. Asking for identification each time would be cumbersome and unnatural. Instead, the robot should identify the person by context as a human would. This context includes, among other things, visual data, which our algorithm uses.

Person recognition must be robust. Facial recognition alone is not enough, since humans will sometimes be facing away from the robot's camera. Similarly to our previously described algorithm for the *Who Is Who?* person recognition task, we again use shirt color as a secondary classifier. Whereas before it was used to differentiate people after comparing their faces, here we demonstrate using it to identify a person when he turns his back to the robot's camera.

Person recognition, in addition to being robust, must be flexible. Rigidly learning a person's exact appearance at one moment will likely not be sufficient to identify him or her after a significant change in appearance. Changes in human appearance can be roughly categorized into two types. One occurs quickly, like the changing of clothes every day or cutting one's hair. The other type of change occurs very gradually and includes growing older and losing or gaining weight. Although we created an algorithm to handle certain cases of both types, the five minute window of our demonstration limited us to creating a scenario that includes only quick changes.

Our scenario was designed to display our algorithm's robustness and adaptability. Specifically, it shows person identification using shirt color as a secondary classifier in the absence of the primary classifier, the face. It also mimics the daily (or so) occurrence of a human changing clothes, showing the robot adapt to this change in the secondary classifier. Lastly, it shows that the Segway robot can effectively follow a recently learned person without markers, as we unfortunately were unable to show during the second phase of the *Follow and Guide a Human* task. The only differences were that we used a lighter-skinned human and shirt colors which stood out from the colors of the background (as opposed to brown-skinned and white-shirted).

Before the demonstration, we again presented a short talk about the robot and our algorithms. A video of the presentation and demonstration can be found at our team web page[4].

The demonstration involved two people, one with whom the robot intended to interact and another who was unrelated to the robot's primary task (stranger). At the beginning, the robot trains classifiers for the intended person's face and shirt. It then follows the learned person based on only shirt color when face is not visible, first with a green shirt and later with a red shirt. The Segway twice gets "passed" to a stranger, whose back is turned (i.e. face invisible) and is wearing the same shirt color. Each time, it follows the stranger until it can see his face. At that point, the face classifier returns a negative classification and supercedes the shirt classifier, and the robot announces that it has lost the learned person and turns away to look for him. Upon finding the original person based on a positive facial classification, it retrains the person's shirt, subsequently stating whether the shirt color has changed.

In the demonstration, the interaction between the face and shirt classifiers was different than in the *Who Is Who?* task. In that task, the shirt classifier refined the results of the face classifier, choosing from possibly several candidate identities. In this demonstration, however, the shirt classifier worked when the robot did not detect a face in its vision input. Also when both classifiers were running (a face and a shirt are detected) but gave contradicting results, the shirt classifier would re-train using samples obtained from the face classifier.

[4]http://www.cs.utexas.edu/~AustinVilla/?p=athome

Team	Final Score
AllemaniACs	256
UT Austin Villa	**238**
Pumas	217
RH2-Y	199
Robot Cognition Lab	190

Fig. 11. RoboCup@Home 2007 final results

This demonstration shows a full implementation of our scheme as depicted in Fig. 4.

The panel of judges scored the presentations and demonstrations of each finalist, determining each team's final standing in RoboCup@Home 2007. We finished in second place (Fig. 11). Of the top three teams, we had a couple of unique characteristics. Our team size of three people was half that of the next smallest team. We were also the only team in the top three that was competing for the first time. We were very successful as well in the specific tasks in which we competed. We received more points than any other team in the person-recognition task of *Who Is Who?* and accomplished all tasks that we attempted in the first phases of *Lost and Found* and *Follow and Guide a Human*.

VII. Related Work

A variety of home assistant robots have been created in the past decade. Many exhibited impressive specific capabilities. Care-O-bot II [16] brought items to a human user and took them away in a domestic setting. It also functioned as a walking aid, with handles and an interactive motion system that could be controlled directly or given a destination. Earlier systems include HERMES [6] and MOVAID [10].

Person following specifically has received much attention from researchers. A recent laser-based person-tracking method was developed by Gockley et al. [15]. Their robot Grace combined effective following with social interaction. A vision-based approach similar to our own was created by Schlegel et al. [25]. In their system, the robot also tracked shirts using color blobs, but the shirts had to be manually labeled in the training images. Some more recent approaches have used stereo vision and color-based methods to track humans [12], [19].

Person tracking is an extensively researched area in computer vision. Several person tracking systems detecting the number of persons and their positions over time use a combination of foreground/background classification, clustering of novel points, and trajectory estimation [11], [17], [27], [32]. These systems focus on algorithms tracking persons using a stationary camera from a relatively distant, high viewpoint from which most of the people's bodies are consistently visible. In contrast, we consider a camera mounted on a mobile robot that may be moving in close proximity to and often at a lower vantage point than the people in question.

In this setting, the target person's unpredictable movement, the robot's inaccurate motion, obstacles occluding the target, and inconsistent lighting conditions can cause the robot to frequently lose sight of its target. To relocate its target after such out-of-sight situations, the robot must be capable of re-recognizing the person it was tracking. For such person recognition, faces are the most natural identifier, and various studies have been conducted on face recognition [29], [20], [4], [28]. Although these systems achieve reasonably high accuracy with well-aligned faces, they are infeasible for a real-time robotic platform due to heavy computation of face alignment or facial component extraction. Instead of recognition methods relying on careful alignment, we extract SIFT features [22] from faces similar to work proposed in [23], [5] and recognize faces by counting the number of matching SIFT features which is performed in near real-time.

To address the brittleness of tracking faces in light of changing poses and inconsistent lighting, we augment a face classifier with other classifiers, e.g. a shirt classifier. Previous work on integrating multiple classifiers has shown that integrating multiple weak learners ("ensemble methods") can improve classification accuracy [24], and the idea has been extended to multiple reinforcement learning agents giving feedback to each other [9], [18]. In [21], multiple visual detectors (e.g. Grey vs. BackSub) are co-trained [7] on each other to improve classification performance. These methods typically focus on merging classifiers that aim to classify the same target function, possibly using different input features. In contrast, the classifiers we merge are trained on different concepts (e.g. faces vs. shirts) and integrated primarily by associating their target classes with one another in order to provide redundant recognition, as well as to provide dynamically revised training labels to one another. Tracking faces and shirts is a known technique [13], [30], but we express the scheme in general terms and focus on the interaction of the classifiers.

There are various data fusion techniques for detecting objects in the environment. Multi-sensor fusion combines readings of multiple sensor devices to improve accuracy and confidence [8], [31]. In our method, we use one input from a single sensor device that is processed in multiple ways. Techniques such as MCOR combine multiple cues for object recognition in the environment [2]. Unlike their approach of adjusting the weight of each cue, we assign static weights to each classifier, but update the classifiers with additional training data using inter-classifier feedback.

VIII. Conclusion and Future Work

The main contribution of this paper was the complete description of our Segway-based platform that performed successfully in the RoboCup@Home 2007 competition. Leveraging our main technical innovation of using co-training classifiers for different characteristics of a person (face and shirt), it was able to follow a person, distinguish different people, identify them by name, and ultimately combine these abilities into a single robust behavior, adapting to a person changing his or her clothes.

The proposed vision algorithm makes use of the shirt or the face color as a fixed secondary characteristic. We have shown how the system adapts when a secondary classifier fails, if for example the background is similar to the shirt color. However, if people have similar shirts, other vision algorithms need to be considered adaptively. Switching the algorithm online would be another interesting application of the inter-classifier feedback.

Though the Segway is adept at identifying previously seen humans, it lacks general object recognition capabilities, instead relying on the ARTag system. Future work that gives the robot the ability to learn and later recognize objects other than people would greatly increase its ability to interact within the home environment.

Mapping capabilities will also be necessary on any fully functional domestic robot. One option is Kuipers' Hybrid Spatial Semantic Hierarchy, a system that provides simultaneous localization and mapping, path planning, and an abstraction from its occupancy grid to an idea of places and portals [3]. Other packages are available as well.

ACKNOWLEDGMENT

The authors would like to thank Youngin Shin for his real-time face recognizer and Mohan Sridharan for the robot's initial development. We are also grateful for Selim Erdogan's invaluable help with our robot's operating system and hardware integration. This research is supported in part by NSF CAREER award IIS-0237699 and ONR YIP award N00014-04-1-0545.

REFERENCES

[1] Robocup@home rules & regulations.
[2] S. Aboutalib and M. Veloso. Towards using multiple cues for robust object recognition. In *International Conference on Autonomous Agents and Multiagent Systems*, 2007.
[3] P. Beeson, M. MacMahon, J. Modayil, A. Murarka, B. Kuipers, and B. Stankiewicz. Integrating multiple representations of spatial knowledge for mapping, navigation, and communication. AAAI technical report ss-07-04, The University of Texas at Austin, 2004.
[4] P. Belhumeur, J. Hesphana, and D. Kriegman. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1997.
[5] M. Bicego, A. Logorio, E. Grosso, and M. Tistarelli. On the use of sift features for face authentication. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshop*, 2006.
[6] R. Bischoff. *Design Concept and Realization of the Humanoid Service Robot HERMES*. 1998.
[7] A. Blum and T. Mitchell. Combining labeled and unlabeled data with co-training. In *Conference on Learning Theory*, 1998.
[8] C. Chong, S. Mori, K. Chang, and W. Barker. Architectures and algorithms for track association and fusion. *IEEE Aerospace and Electronic Systems Magazine*, 2000.
[9] R. Crites and A. Barto. Elevator group control using multiple reinforcement learning agents. *Machine Learning*, 1998.
[10] P. Dario, E. Guglielmelli, C. Laschi, and G. Teti. Movaid: A mobile robotic system for residential care to disabled and elderly people. In *MobiNet Symposium*, 1997.
[11] T. Darrell, G. Gordon, M. Harville, and J. Woodfill. Integrated person tracking using stereo, color, and pattern detection. *International Journal of Computer Vision*, 2000.
[12] V. Enescu, G. Cubber, K. Cauwerts, H. Sahli, E. Demeester, D. Vanhooydonck, and M. Nuttin. Active stereo vision-based mobile robot navigation for person tracking. *Integrated Computer-Aided Engineering*, 2006.
[13] M. Everingham, J. Sivic, and A. Zisserman. Hello! my name is... buffy – automatic naming of characters in tv video. In *British Machine Vision Conference*, 2006.
[14] M. Fiala. Artag, a fiducial marker system using digital techniques. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2005.
[15] R. Gockley, J. Forlizzi, and R. Simmons. Natural person-following behavior or social robots. In *Human-Robot Interaction*, 2007.
[16] B. Graf, M. Hans, and R. Schraft. Care-o-bot ii–development of a next generation robotic home assistant. *Autonomous Robots*, 2004.
[17] I. Haritaoglu, D. Harwood, and L. Davis. W4: Real-time surveillance of people and their activities. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2000.
[18] S. Kumar and R. Miikkulainen. Dual reinforcement q-routing: An on-line adaptive routing algorithm. In *Artificial Neural Networks In Engineering*, 1997.
[19] H. Kwon, Y. Yoon, J. Park, and A. Kak. Human-following mobile robot in a distributed intelligent sensor network. In *IEEE International Conference on Robotics and Automation*, 2005.
[20] D. Lee and H. Seung. Learning the parts of objects by non-negative matrix factorization. *Nature*, 1999.
[21] A. Levin, P. Viola, and Y. Freund. Unsupervised improvement of visual detectors using co-training. In *IEEE International Conference on Computer Vision*, 2003.
[22] D. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 2004.
[23] J. Luo, Y. Ma, E. Takikawa, S. Lao, M. Kawade, and B. Lu. Person-specific sift features for face recognition. In *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2007.
[24] R. Schapire and Y. Singer. Improved boosting algorithms using confidence-rated predictors. *Machine Learning*, 1999.
[25] C. Schlegel, J. Illmann, H. Jaberg, M. Schuster, and R. Worz. Vision-based person tracking with a mobile robot. In *British Machine Vision Conference*, 1998.
[26] M. Sridharan and P. Stone. Real-time vision on a mobile robot platform. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2005.
[27] C. Stauffer and W. Grimson. Adaptive background mixture models for real-time tracking. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1999.
[28] M. Turk and A. Pentland. Face recognition using eigenfaces. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1991.
[29] P. Viola and M. Jones. Robust real-time face detection. *International Journal of Computer Vision*, 2004.
[30] S. Waldherr, R. Romero, and S. Thrun. A gesture based interface for human-robot interaction. *Autonomous Robots*, 2000.
[31] E. Waltz and J. Llinas. *Multisensor Data Fusion*. Artech House, 1990.
[32] C. Wren, A. Azarbayejani, T. Darrell, and A. Pentland. Pfinder: Real-time tracking of the human body. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1997.
[33] T. Zant and T. Wisspeintner. Robocup x: A proposal for a new league where robocup goes real world. In *RoboCup*, 2006.

GlowBots: Designing and Implementing Engaging Human-Robot Interaction

Mattias Jacobsson, Ylva Fernaeus and Lars Erik Holmquist

Abstract—**GlowBots are small tangible, communicating and interactive robots that show eye-catching visual patterns on a round LED display. This paper details the development of the GlowBots from the early user-oriented design phase, through hardware and software development and onto preliminary user studies. In the design phase we outlined a robot application based on a study of how owners relate with unusual pets, such as snakes and lizards. This led to an application concept of a set of ”hobby robots” which would communicate with each other and the user through dynamic patterns. Based on these requirements, we developed a LED display called see-Puck, which together with an open robot platform was used for the GlowBots application itself. One particular issue is dealing with energy consumption problems, as resources in embedded systems often limit the potential time for user interaction. We conclude with a report on early user experiences from demonstrating GlowBots and a preliminary user study in a home environment as well as remarks about future directions.**

Index Terms—**GlowBots, Human Robot Interaction, Tangible Interfaces, Ubiquitous Computing.**

I. Introduction

THE ROBOTS are coming, but are they here to stay? [1] Human-robot interaction is a rapidly expanding area, with many new journals and workshops appearing in recent years. However, in order for robots to truly become a part of our everyday life they should provide a meaningful and sustainable presence as a result of interaction with other robots, humans, pets or devices. Seen from this perspective, everyday robotics shares a strong synergy with the vision of ubiquitous computing [2], and tangible computing [3], where technology tends to become more and more intimate [4]. The main difference from these emerging interaction paradigms is that robots manifest themselves as mobile embodied units that can affect the world physically.

In the European project *ECAgents – Embodied Communicating Agents,* [5] we have been actively working to expand the boundaries of what interaction with robots might be like in the future. Mundane labor such as vacuuming, cleaning or other practically oriented chores are merely a subset of existing needs where robots could play a role [6]. From a design point of view, it is also important to use ”out of the box thinking”, as we might miss out on important areas and interaction modes that are difficult to imagine before they exist.

As a way to stimulate new ways of thinking about robots, we first gathered researchers in the field for a two days workshop called ”Designing Robots for Everyday Life” to brainstorm about innovative new robotic design spaces [7]. As a direct outcome from this we got a number of robotic mockup scenarios, for instance, robot plants that would re-arrange themselves in order to guide queues in a complicated environment such as an airport, or the listening psychologist, bean-bag shaped robot that would attach itself to a car’s rear mirror. But more interesting than these design suggestions was that we learned that designing robotic applications often results in far-fetched expectations and visions of problem oriented scenarios – even though we did all we could to be as open minded as possible.

To further explore how robotic appliances could be designed, we started to experiment with a new design method, *transfer scenarios* [8]. In this process we sought ways of grounding our designs in existing practices where relationship, autonomy and embodiment were essential [9]. We looked for an existing human practice that could be used as inspiration and guidance for the design of new forms of robots. Eventually, we decided to study owners of unusual pets, such as snakes, spiders and lizards as inspiration for designs. One of the results pointed towards an application where robots are engaging, but not overly personal, similar to a dynamic, mobile and visually appealing trading card game collection.

Fig. 1. Exhibition visitors playing with a swarm of GlowBots.

ted

Mattias Jacobsson, Ylva Fernaeus and Lars Erik Holmquist are with the Swedish Institute of Computer Science.
E-mail: majac@sics.se, ylva@sics.se, leh@sics.se

a round LED-display that could extend an open educational robot platform, the *e-Puck*. The result was a top mounted extension module that we swiftly named *see-Puck* [10]. We then also needed to make several energy-optimization changes to the e-Puck firmware in order to cope with sustainability and stability problems. The resulting application is called *GlowBots* and consists of the assembled hardware together with software for infrared communication and animated morphing patterns (Fig. 1).

Not only did this project begin with a study, but we will also conclude this report with a preliminary user study conducted in a real home environment. In the discussion we will also compare findings and comment on the overall design process.

II. Background and Related Work

What would be the design requirements for a more subtle robot technology, one which could be found in the intersection between robotics and ubiquitous computing - robots that quietly find their ways into our everyday life and eventually become an integral part of it? Today new robots are appearing almost every day, so first we would like to recapture some historical points and put our standpoint in contextual highlights.

People have an underlying assumption that robots are socially capable [11]; hence they are quite biased when it comes to their image of a robot. The word itself (although not the concept) originates from Czech ”robota”, and was popularized through the theatric play by Karel Ĉapek called ”Rossum's Universal Robots” (R.U.R) in 1921. The word at that time simply meant work or compulsive labor, but a general definition once given in the Merriam-Webster dictionary centuries later still reflected this common perception:

”An automatic apparatus or device that performs functions ordinarily ascribed to human being or operates with what appears to be almost human intelligence”.

It may be a pity that we did not catch up on the Japanese older and more humble notion similar to *automaton*. The profound cultural differences to western attitudes could be seen as in contrast to the Japanese compassion for robotic characters like the *Mighty Atom* (Astro Boy in the US) which is more emotionally oriented rather than labor oriented.

Today, the word robot still represents a governing descriptive purpose, but we also have a flora of words in the subsequent field of robotics that captures more fine tuned distinctions, e.g. android, humanoid, mecha, cyborg, but which all still inherits much of the original anthropomorphic connotations. Another example, the *Robot Fish* [12], is designed to be a copy of a common fish in terms of looks, properties and behavior. This approach is common, especially from a robotic toys perspective as anthropomorphic values are added to the designs as a mean to extend interaction. To mention just a few of these commercial examples of robotic pets we have Aibo, RoboPanda, Furby [13] and now also Pleo [14].

Masahiro Mori's *uncanny valley* is an example of what happens when more subtle expectations do not correspond to the perceived input in human-robot interaction [15]. Instead of getting relaxed and enjoying the anthropomorphic features the hypothesis states that we unconsciously start to focus on the dissimilarities which in the end results in an uneasy repulsive reaction. In relation to this theory we notice that anthropomorphism and zoomorphism play an important role in setting the levels of expectations, and by being aware, and taking control of these insights would be a key component in taming robotic design.

As a consequence we sometimes instead prefer to use the term ”embodied agents” to describe a more general and open view of robots that moves the focus away from traditionally biased anthropomorphic preconceptions [7]. Other researchers prefer the term ”robotic product” to denote mechanically based interactive applications [16]. Examples of robots intended for labor oriented work include the *Roomba* vacuum cleaner [17], the *Artemis* guard robot [18] or the *Minerva* museum tour guide robot [19].

In a sense our work is the opposite of the above approaches; we have absolutely no intention to make a new dog or cat, or replace work already performed by humans. Several researchers are also pursuing such alternative views of robots. For instance, *The Hug* [20] is an example of a robot that does not look like anything biological, but instead reminiscent of an artifact that can be found in an everyday setting, in this case a pillow. It does not have any sophisticated communication capabilities like speech, or complex behavior like walking. Instead it appeals to our most primitive need of affection. Yet another example of a design that expresses life-like qualities but also integrity is *Tabby* [21] – a simple interactive furniture demonstrator. Our work is thus similar in that we also move our focus away from the ordinary expectancies of robotics and at the same time avoid elevated expectations.

Another relevant study looked at peoples' relationships with everyday artifacts, such as computers, corkscrews and notebooks [22]. It points out that a notebook will increase in perceived value over time as it is filled with notes and sketches, while e.g. fashionable clothes value actually decreases as it becomes increasingly obsolete. We found such observations inspirational in regards to where we should position ourselves and think about future robot applications.

III. Design Method

One of the problems of designing novel forms of robots could be the lack of perspective outside that of the experts and scientist who are already designing robots. We have taken inspiration from the field of human-computer interaction to find methods that infer design implications either directly from studies of users or by extrapolating from known human needs and interests. One such method is to use fictive representative characters called *personas* [23].

We started by seeking out possible sources where established interaction and engagement are essential properties and autonomy plays a significant role. From earlier studies we knew that looking at practices that lies down the long tail of practices, so called *marginal practices*, tend to turn our minds away from the established discourse [8]. When looking for a suitable practice to engage with, we were interested in people who were interacting with living things – but not necessarily commonplace pets, such as cats and dogs, since such an interaction has already been proven hard to capture in a robot. We decided to study and interview owners of pets with fairly low cognitive capabilities and unusual affordances for interaction, e.g. spiders, lizards and snakes (Fig. 2). In total we conducted ten interviews with six male and four female participants. Three of the interviews were made face-to-face and the rest by phone due to logistic restrictions. Typical questions during the interviews would be about why they were interested in a particular pet, what the pet does, what they do together and how they could tell the mood of their pet.

Fig. 3. When designing the GlowBots we took inspiration from the relationship people develop with unusual pets such as spiders.

We then transcribed the answers from the interviews and cut up quotes and wrote them down onto Post-its. From the scrambled Post-its we then tinkered and linked together different properties and intrinsic characteristics in various constellations. After iterating this process several times four distinct clusters started to emerge representing the rough outline for the four personas (Fig. 3). In one of the affinity clusters we could then read several statements without any apparent contradiction e.g.:

- *He does not pet his pets, nor is he interested in different personalities of the pets.*
- *He is interested in breeding his pets in order to create nice patterns.*
- *He enjoys reading about his pets and often meets up with people that have similar pets, to look at or even exchange pets.*

The next step was to create the personas from these clusters, which are descriptive scenarios of imagined users. The complete scenario was then created by filling in general fictive ”glue data”, connecting such different quotes into meaningful coherent descriptions. In total we created four such personas [9] but in this case we will focus only on the persona that is relevant in the context of GlowBots.

We then named the persona, which is a powerful way of building a mental image around a common reference. This particular persona goes by the name Nadim. At this stage the

Fig. 2. a.) Selected data was taken out as notes from the transcribed data. b.) The notes where clustered, each being a starting-point for one persona.

scenario would still refer to a relationship with pets; however, by simply changing the word ”pet” to ”agent”, we *transferred* the scenarios to our target domain [8]:

Nadim is 32-years old and works as a network engineer, living alone in a two-bedroom flat in a small town. He has always had a great interest in collecting and exploring various things, and as he got older he became fascinated in having ***agents*** *as a hobby. Nadim finds it exciting to try to understand their behavior and sees them as a research area where there is always something more to learn. He enjoys watching them communicating to each other and changing their patterns. Every single* ***agent*** *has its own specific colour pattern, and when it is put close to another* ***agent*** *they both start to change their individual patterns. The surrounding light, sounds and movement etc, also affects their patterns. The changes are slow, and sometimes it takes several days until it Nadim can see how an* ***agent*** *is reacting. The challenge is to avoid results that are bland or unattractive. Nadim is quite good in developing* ***agents*** *with unique interesting patterns, and he puts pictures of the* ***agents*** *on his website. The number of* ***agents*** *Nadim owns varies, and he has never bothered to give them any names. He likes to read everything that crosses his path; Internet pages and magazines. He also frequently visits other sites to compare patterns and sometimes he writes in a forum for people with the same type of* ***agents****. They sometimes also meet to let their* ***agents*** *affect each other’s patterns.*

This scenario now expresses what a potential user of an autonomous agent would look like. The final step in this process includes matching technology with the scenarios to sketch out real designs:

The agents can evolve interesting patterns over time, but it is a lengthy process and might not always succeed. Agents will be equipped with a color display on their back and have one or more sensors for light, movement and sound. The sensing can be different for different agents. Each agent will have a unique color pattern, developed from meetings with other agents the environment it is in. By touching the agent in a particular way makes it possible to temporarily freeze a pattern. Achieving a nice pattern requires several agent-agent interactions and an attention to timing.

Based on this description we could then proceed with sketching and implement a rough first prototype.

IV. SeePuck Development

Our design pointed towards some kind of visual interface as one of the central components. We also decided to base the project on an educational robot platform, the e-Puck, developed by Ecole Polytechnique Fédérale de Lausanne [24]. We looked for an existing display but found that all currently available displays had a rectangular shape, often needed backlight to be visible from a distance and prioritized resolution and color depth over cost. We decided to design a new display that fit our needs, and in particular one that had a shape that would fit on the round, roughly coffee-cup sized e-Puck. When designing the new platform, much effort was put into hardware and software design, keeping it simple, modular, obvious, cheap, energy efficient and robust.

A. Hardware

The see-Puck is designed to fit on top of the *e-Puck* robot and connect through a serial interface. We use the version 2.0 of the e-Puck, which features a number of sensors and actuators including infrared (IR) proximity sensors, one camera, three microphones, a 3-axis accelerometer, loudspeaker, stepper motors, Bluetooth interface, a number of LEDs, a PIC microcontroller, and a twelve step mode-selector.

Fig. 4. The two printed circuit boards of the see-Puck module are mounted on top on an e-Puck.

The see-Puck display module (Fig. 4) consists of two printed circuit boards, one *controller board* and one *matrix board*, sandwiched together by two perpendicular connectors. This design ensures that the matrix board that holds all the LEDs can only be fitted in one way. The controller board (Fig. 5) holds its own microcontroller (Atmel ATmega8L) and firmware to handle higher level instructions from the e-Puck through a RS232 serial interface.

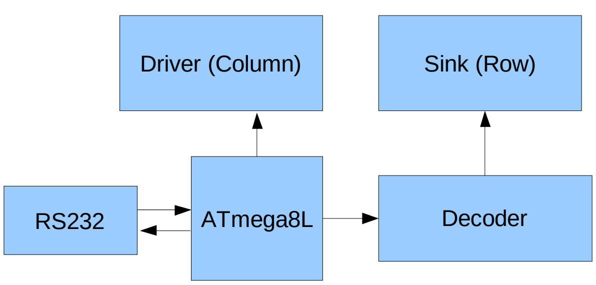

Fig. 5. Controller board overview with arrows indicating the flow of information. The driver sets a column high while the sink grounds one of the rows given by the decoder.

The matrix board holds 148 LEDs in a rounded 14 by 14 matrix. To keep the energy consumption down and also maximize light intensity we exploit a known, but often overlooked feature of the LEDs – the possibility to light them up using short rapid pulses of higher current. To the human eye the quickly flashing the LEDs will appear as a constant light. The gain is significantly lower total energy consumption, which is one of the most important factors when designing for devices that rely on batteries. Furthermore, flashing the LEDs is a perfect fit with the electronic design, since only one LED per column can be lit at a time.

B. Software

The software has two parts, one *library* for the e-Puck consisting of the higher level commands that are sent to the see-Puck module from the e-Puck and one *firmware* part for the microcontroller on the see-Puck controller board. The range of graphical commands available in the library represents the most basic ones e.g. *set a pixel, draw a line, draw a circle, shift screen,* etc. These commands often take arguments in form of coordinates and LED brightness. We also decided to make graphics double buffered, i.e. the actual drawing is done to one buffer while the other is shown, so that flickering in animations is kept at minimum.
The firmware consists of two interrupt-driven subsystems - the communication and the graphical subsystems, which run side by side parallel to a continuous main loop (Fig. 6**¡Error! No se encuentra el origen de la referencia.**).

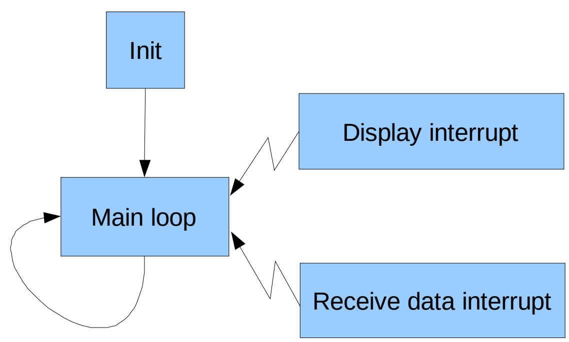

Fig. 6. See-Puck firmware schematic overview with two simple interrupts.

When the communication subsystem receives a byte over the UART (the serial interface on the microcontroller side), it calls the receive data interrupt. After checking the integrity of the message it gets stored into a ten level sized software implemented FIFO buffer shared with the graphics subsystem. At the end the receive interrupt is reset.
The graphics subsystem interrupt is timer-based and called about 60 times a second. When called it starts with getting a pointer to the current front buffer. It then cycles through each row sending a PWM (Pulse-Width Modulation) signal with a four bit resolution for each LED, representing the specified brightness.

The firmware starts with an initialization of the graphics subsystem. It then turns on all the LEDs for about a second before it initializes the communications subsystem and enters the main loop. The interruptable main loop then continuously checks the FIFO for new commands and executes them, preparing the back graphic buffer.

```
#include "e_see_puck_lib.h"

int main(void){
  int x = 7, y = 7, r = 4, c = BRIGHT;
  e_see_puck_init();
  e_see_puck_draw_circle(x,y,r,c);
  while(1){
    e_see_puck_swap_buffers();
    e_see_puck_copy_buffers(FRONT_TO_BACK);
    e_see_puck_hscroll(-1);
    //Waste some cycles here
  }
}
```

Fig. 7. An example program for the e-Puck using the standard graphics library developed for the see-Puck.

In the illustrative C-code example (Fig. 7) a circle is first drawn at the center of the back buffer. It then enters the main loop executing a buffer swap to make it visible. The front buffer is now copied onto the back buffer and scrolled one step horizontally (to the left) before next iteration. The final result is a circle that scrolls over the screen.

C. Energy Optimizations

When the see-Puck modules arrived from factory we measured the average power consumption for it to be in the range of 20 mA. During initial tests it all seemed fine until we started using more and more sensors and actuators. After deeper investigation we found out that the biggest issue was the stepper motors that at the time ran in the range of 200 mA. During power peaks such as sudden friction events e.g. running into an obstacle or another robot this would cause instability problems for the display or even the e-Puck. This forced us to soft-optimize some portions of the e-Puck libraries and to use PWM where possible. This trick worked out very efficiently for the stepper motors, which landed on an average of about 30 mA afterwards (no load). Similarly, all LED's on the e-Puck were also pulsed to save even more power.

V. GlowBots Development

Based on the see-Puck hardware, we then constructed an interactive application inspired by the Nadim persona. We had a total of 20 complete robots (e-Puck platform plus see-Puck display) which would allow for large groups of interacting robots. Here we will outline the steps involved creating the GlowBots demonstrator application from based on the design proposal and readied platform.

A. Visualizations

The idea with GlowBots was to let the users interact with an ever-changing set of robots, which would express themselves with dynamic patterns on the LED display. In the early proof-of-concept prototype we started out with Conway's Game of Life, a well known cellular automata example, to produce interesting dynamics on the display when the robots interact. Although it was relatively open-ended, it did not satisfactory convey the intended interaction. We then sought a way of displaying interesting shapes that could be semantically interpreted and that somehow would morph more intuitively as interaction took place. After a great deal of investigation we came to use analytical curves based on the super-formula equation [25], chosen for its richness of simple shapes. The resulting shapes can be anything from star, square, circle, egg, flower and any intermediate state in between (Fig. 8).

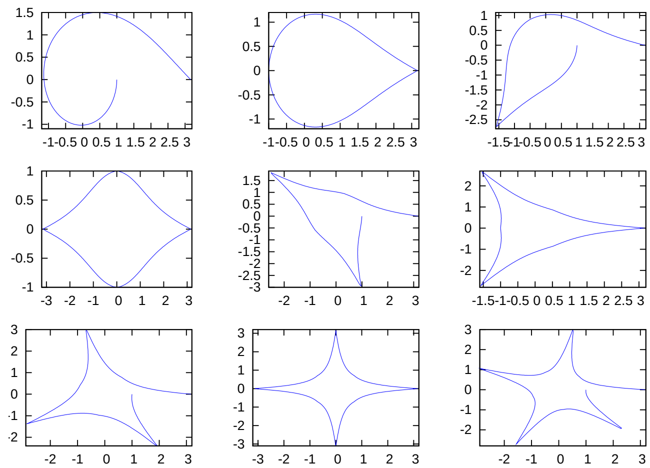

Fig. 8. Examples of shapes generated by the super-formula that would typically occur in the GlowBots application.

B. User Interaction

The user interaction stems from the developed persona description from the design step. Users interact with the robots directly, either by moving them around on the surface (to place a robot next to another with an interesting pattern) or by gently shaking them. If the user shakes the robot up and down this will encourage the pattern that the robot is currently displaying to become more dominant. If the user shakes the robot side to side, this will instead have the effect of making the robot more susceptible to be influenced by other patterns. Thus, while the users cannot directly create new patterns, they can indirectly influence the visuals by encouraging certain patterns and discouraging other. As two robots stand next to each other, they will start communicating and slowly converge to showing the same pattern, which will be a mix of both the original patterns. The effect is that of a slowly evolving, constantly surprising collection of a tangible autonomous robotic display.

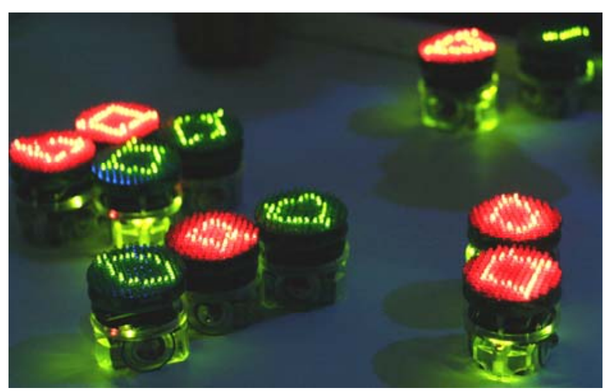

Fig. 9. A group of interacting robots that uses their patterns to attract users to play and interact with them.

From an application point of view, each GlowBot will communicate their respective parameterized internal states, including current motion and the shape visualized on the display. The robot-robot communication uses the infrared proximity sensors for broadcasting and receiving data. There are two important reasons for choosing IR over e.g. Bluetooth. First, since there are eight IR-sensors distributed around the robot, we can get a sense of directionality. Second and most important is the situatedness of the communication. The communication radius of IR is typically 10-15 cm, which means that only robots that are close to each other will communicate (Fig. 9).

VI. Demonstration Evaluation

GlowBots have been shown at several major venues such as SIGGRAPH Emerging Technologies [26] and WIRED NextFest, with a combined audience of over 60.000 people. At these settings, we had the opportunity to observe how the GlowBots demonstration was received by ordinary people. Rather than only having them for show, we encouraged people to come up and play with the GlowBots. Literally thousands of people have thus gotten hands-on experience, many of them school children from the Los Angeles area. Having this kind of demo was possible only due to the efforts made in energy consumption optimizations and carefully planned continuous maintenance during the exhibitions. It also helps to have a swarm of units so that the demo does not rely on a single unit.

As with big exhibitions like SIGGRAPH people come to see the latest news in technology, listen and ask questions about the presented material. In the end they usually end up with a flyer or brochure to take home and reflect upon. The one thing we really could expect in this type of setting is the brief experience based on very first contact that the visitors would have with our GlowBots as they stumbled into our presentation booth.

Based on informal observations of how users interact with the GlowBots in exhibition settings we noted that many users spontaneously thought that the display would react by touch, similar to a large press button. Since this had not yet been conceived of as a possible use, we soon realized that the robots were not robust enough for such treatment. We thus had to stabilize the robot platform so that even though the robots still did not work as push buttons, they would not break in case someone tried to use them as such.

At SIGGRAPH, we recorded several hours video of the demonstration as people stopped by and interacted with our GlowBots. When reviewing this material we saw that the complexity of the setting involving many moving glowing tangible artifacts, crude and developing use of speech and gesture made it an analysis problem. Also, the level of noise from surrounding demonstrators and the fact that we had used a hand camera resulted unfortunately in very poor sound coverage.

As a first step in the analysis we published a small video-clip, showing how a little girl, five to six years old, plays with GlowBots for several minutes, before her dad wants to leave [27]. From the look of her face and posture she is totally immersed with the interaction and very hesitant about leaving the newly found little friends. We will also use the video material as a testbed for applying analysis tools for video encoding. For example, only transcribing the user side of the interaction would result in encoding only part of the story, leaving out important aspects related to the multitude of interactions.

Our experience from the demonstrations suggests that even though the design was initially based on a scenario of an adult persona, in its current state there is even more potential of GlowBots as used by children.

VII. Preliminary User Evaluation

During demo sessions, most focus was on the hardware platform, and the actual and intended implementation of the software could be discussed with the presenters as part of the demonstration. More recently, we have also explored more long term use in a home environment. Leaving the robots ”on their own” with users, and allowing users to make their own interpretations of what the robots should do and what they should be good for could then potentially give much valuable input to the design process, apart from also being a more realistic case for testing the robustness of the hardware platform. In HCI research, this approach to user studies is sometimes referred to as *Technology Probes* [28], which is an increasingly popular method for user-inspired interaction design. The goal with our user studies were thereby not primarily to *evaluate* the technology, i.e., to say if it works or not, but to explore *how* the robots are used, what intended and unintended usages that may arise, and to feedback into what directions to further develop the designs.

With the initial development primarily focusing on the hardware and internal infrastructure of the robots, we needed to perform a pilot to investigate both the robustness of the hardware platform, as well as to gain more input to the details of how the software to run on the robots should be take shape. The interaction pace that had been developed was at this point geared towards exhibition settings where people typically only have a few minutes for every demonstration, which we assumed was rather different from the interaction in the intended home environment. To be able to learn how the robots could be used, and how to further develop the technology at the application level, we therefore complemented or experiences from demo and lab studies with a long-term study of how the robots were taken into use in a domestic setting.

A collection of 10 GlowBots were placed in the home of a 34 year old man with two children, (a girl of six and a boy of four years old), for a period of six weeks. The children were staying with their mother every other week, so their father was left alone with the GlowBots during half of the study period. From our previous demo sessions, we knew that the robots would need quite some maintenance with exchange of batteries, which was taken care of by the father in the family. Below we report on some initial findings based on video recordings and interviews.

The fact that the robots *glow*, meant that they became quite specifically experienced as to be used in darkness. This could be observed for instance in how a natural part in ‘staging’ the play session with the GlowBots was to switch off the ordinary light in the room. The displays then worked as decorative toys that could be played with in the dark, at the same time placing attention on themselves as the focus in the play activity.

The first spontaneous comments that we got on the GlowBots functionality were concerned with how the robots *moved*. It was repeatedly pointed out by the children that they moved too slowly, especially since after a period of active play, the robots usually stopped moving due to low battery levels. Moreover, the robots were at this stage programmed to keep going until they reached a wall, and then stayed there, which made them appear ’stupid’. Although it became part of the play to go collecting the robots that were on escape towards the wall, this soon became rather uninteresting as an activity on its own. The users suggested that instead of moving in a straight line in one direction, the robots should be able to wander about in a more complex manner. This would make the robots feel more unpredictable and interesting to play with. They were also interested in being able to in some way control how the robots moved, e.g. by waving or putting something in front of the robots.

As soon as the robots stopped moving, they did seem to get transformed into a kind of static mechanic sculptures, bringing back the glowing LED surface into attention. These were clearly attractive for the children and were used in a variety of ways in their play. Surprisingly though, the children did not initially seem to pay much attention towards the actual patterns that were on display (Fig.10).

Fig. 10. Using a GlowBot as a vehicle, podium or stage for other toys (left), and stacking the GlowBots into a tower (right).

The users were clearly attracted by the looks of the dynamic and glowing patterns, but they did not seem to reflect as much as one could expect on how the patterns arouse and how these were communicated between the robots. Instead, more focus was placed on the behaviour of how the robots moved, expressed for instance in discussions on what made them move in a certain direction and whether or not their movement could be controlled somehow. This suggests that physical robotic movement possibly overrides patterns on a visual display in terms of users’ direct experience. Although this needs to be further investigated, it could be valuable aspects to consider in the development of new interactive technologies that make use of a combination of motion and visual display technologies.

VIII. Discussion

Looking back onto the original design-proposal that came out from the Nadim persona, we believe that the governing idea is still on track, while minor changes have been introduced to allow for a more seamless interaction. For instance the software does not impose heterogeneous sensing capabilities, but instead small and big differences in hardware settings contribute to individuality. Infrared sensors are bent, batteries end up having different mileage causing robustness problems and there are different manufactures of IC-circuits between hardware revisions. All this contributes to making even the most mass-produced robot more individual and characteristic, something that in the end would benefit personalized, although subtle, interaction between man and robots.

It is interesting to observe how natural it looks when people interact with embodied, tangible and communicating digital artifacts, like the GlowBots. It not only becomes a bonding experience, but it also lets the users explore communication through observing cause and effect. It is also important to notice that embodiment and communication are closely entangled, which becomes very important when another type of embodied element, as for instance another user, enters the picture. We noticed that for humans to be a part of an ongoing communication we observed that the setting benefitted from being truly situated. For instance, if the range of GlowBots communication would have been in e.g. the Bluetooth reach, the perception of the swarm would have been very different and more resembling a simulation running on a computer.

Our design process illustrates how sensitivity to changes in the technology, and experiences of user interaction sometimes result in essentially new use settings which was not envisioned. For instance there became much more hands on and petting activity than envisioned in the original design descriptions.

IX. Conclusions and Future Work

We have detailed the work on a novel robotic prototype, GlowBots, that was the result of a design effort developed to open up new perspectives on the future role of everyday robots. We ended up creating a form of robot that would entice an aesthetic experience outside the domains of the zoomorphic pet robots previously seen in research and products. Although the initial design came from a specific scenario [9] the see-Puck platform is not limited to the GlowBots application. We hope that the detailed development of the see-Puck could work as inspiration for how to construct simple displays with rather unconventional shapes. All see-Puck hardware and software is released under a GPLcompatible license so that anyone can use, revise, extend upon and improve it.

Fig. 11. When demonstrating the GlowBots we encouraged people to come forward and get the hands-on experience.

At SIGGRAPH and NextFest, we demonstrated the GlowBots continuously for several days at a stretch to thousands of users (Fig. 11), but this required almost constant battery changes and continuous maintenance of the robots. We have observed that energy consumption in this type of setting can vary greatly, not only because not two robots are perfectly identical, but also because they are both autonomous and tangible. We would also like to note that our efforts in optimizing energy consumption resulted not only in improved mean runtime, but more importantly contributed to an overall increased robustness.

We wanted to encourage a long-term relationship with the robots, inspired by how people interact with artifacts and creatures in everyday settings. One aspect that crystallized in this process is the need of open ended play – an important factor to sustain interest over time. We believe this work shows that it is possible to change how we think about new robotic products and how we can rethink their roles in our everyday environment. By grounding the design in existing needs, they will have the potential to last considerably longer and have a much more rewarding interaction than what in most cases is being offered today.

In the future, see several possible improvements for the see-Puck. One example would be to make the display touch sensitive by also using the LEDs as sensors [29]. This would allow users to point at and directly influence what is seen on the display, for instance to ”paint” patterns directly on the display. Another important improvement would be to continue the work on software optimization on the e-Puck in order to increase battery life and overall robustness even further.

From the preliminary user evaluation we found several interesting observations that requires further investigation, but also implications guiding further development. The more immediate step will then be to tune our pilot application and once more place it in an everyday setting to study it in more depth. In this case we will slow down the interaction, which is currently geared towards exhibition settings where people typically only have a few minutes for every demonstration. We will also take more consideration to motion behavior due to the much larger spaces that home environments offers. For a truly long-lasting relationship to develop between robots and humans, it is necessary to sustain the interest level over weeks, months and hopefully years. Achieving this sustained level of interest is an important challenge for future human-robot interaction applications.

ACKNOWLEDGMENTS

First of all we would like to thank our intern student Johan Bodin for his great efforts on the hardware and software, as well as helping out with the demonstrations. Especially we would also like to acknowledge Sara Ljungblad for her work in the initial design phase and Katarina Walter who conducted the initial user studies, contributed to brainstorms and later also participated in the exhibitions. Further we would also like to thank Jeffery Knurek, Mattias Rost, Lorena Monsalve, Darby Thompson, Staffan Jonsson, Daniel Sundström and Henrik Berggren who all also participated in demonstrating GlowBots. Thanks also all of our other colleagues and fellow researchers for their valuable input and discussions during the project. This project is part of the ECAgents project and is sponsored by the Future and Emerging Technologies program (IST-FET) of the European Community under contract IST-2003-1940.

REFERENCES

[1] L.E. Holmquist, ”The robots are coming”, *Interactions* 12(3): 58-59, ACM 2005.
[2] M. Weiser, and J.S. Brown, ”Designing Calm Technology”, *PowerGrid Journal*, 1, 1, July 1996.
[3] H. Ishii and B. Ullmer, ”Tangible bits: towards seamless interfaces between people, bits and atoms”, *In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, New York, 234-241. 1997.
[4] G. Bell, T. Brooke, E. Churchill and E. Paulos, ”Intimate Ubiquitous Computing”, *Proc. UbiComp'03 Workshop*, 3-6, ACM, 2003.
[5] ECAgents project, http://ecagents.istc.cnr.it/
[6] J-Y. Sung, L. Guo, R. E. Grinter, and H. I. Christensen, ”My Roomba is Rambo”: Intimate Home Appliances, *In Proceedings of 9th International Conference on Ubiquitous Computing (UbiComp 07)*. Innsbrück, Austria, September 16-19. 2007.
[7] S. Ljungblad and L.E. Holmquist, ”Designing Robot Applications for Everyday Envionments”, *In Proceedings of Smart Objects & Ambient Intelligence 2005* (SOC-EUSAI), Grenoble, France, 2005.
[8] S. Ljungblad and L. E. Holmquist, ”Transfer scenarios: grounding innovation with marginal practices”. *In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, San Jose, California, April 28 - May 03, 2007.
[9] S. Ljungblad, K. Walter, M. Jacobsson and L.E. Holmquist, ”Designing Personal Embodied Agents with Personas”, *In Proceedings of RO-MAN 06*, Hatfield, United Kingdom, 2006.
[10] M. Jacobsson, J. Bodin, and L.E. Holmquist, ”The see-Puck: A Platform for Exploring Human-Robot Relationships”. To appear in Proceedings of CHI 2008, Florence, Italy, April 10-15, 2008.
[11] T. Fong, I. Nourbakhsh and K. Dautenhahn, A survey of Socially Interactive Robots, *Robotics and Autonomous Systems,* 42, 143-166, 2003.
[12] H. Liu and H. Hu, ”Biologically inspired behaviour design for autonomous robotic fish”, *International Journal of Automation and Computing*, Vol. V3, No. 4, pp. 336-347. October 2006.
[13] H. H. Lund, Adaptive Robotics in the Entertainment Industry, *In Proceedings of IEEE International Symposium on Computational Intelligence in Robotics and Automation* (CIRA2003), 16-20 July, Kobe, IEEE Press, 2003.
[14] Pleo, http://www.ugobe.com/pleo
[15] M. Mori, ”The uncanny valley” (K. F. MacDorman & T. Minato, Trans.). Energy, 7(4), 33, 1970.
[16] J. Forlizzi, C. DiSalvo and F. Gemperle, ”Assistive Robotics and an Ecology of Elders Living Independently in Their Homes”, *Journal of*

HCI Special Issue on Human-Robot Interaction, V19 N1/2, January, 2004.
[17] Roomba, http://www.irobot.com/roomba
[18] Artemis, http://www.tmsuk.co.jp/artemis/product.html
[19] Minerva, http://www.cs.cmu.edu/~minerva/
[20] C. DiSalvo, F. Gemperle, J. Forlizzi, E. Montgomery, W. Yonkers and J. Divine, ”The Hug: An Exploration of Robotic Form for Intimate Communication”, *In Proceedings of RO-MAN 03*, vol., no.pp. 403- 408, 2003.
[21] A. Ueki, M. Kamata and M. Inakage, ”Tabby: designing of coexisting entertainment content in everyday life by expanding the design of furniture”, *In Proceedings of the international Conference on Advances in Computer Entertainment Technology, ACE '07*, vol. 203. 2007.
[22] F. Kaplan, ”Everyday robotics: robots as everyday objects” *In Proceedings of Soc-Eusai 2005*, p.59-64, Grenoble, France, 2005.
[23] J.M. Carroll, ”Making Use: scenario-based design of human computer interactions”, *Cambridge, Mass, MIT Press*, 2000.
[24] e-Puck, http://www.e-puck.org/
[25] J. Gielis, ”A Generic Geometric Transformation That Unifies a Wide Range of Natural and Abstract Shapes”, *American Journal of Botany* 90(3): 333-338, 2003.
[26] M. Jacobsson, S. Ljungblad, J. Bodin, J. Knurek and L.E. Holmquist, ”GlowBots: Robots that Evolve Relationships”, *SIGGRAPH’07 Emerging Technologies*, San Diego, California, 2007.
[27] M. Jacobsson, ”GlowBots – A Love Story”, *In adjunct proceedings (video session) of Human Robot Interaction 2008*, Amsterdam, Netherlands, 2008.
[28] H. Hutchinson, W. Mackay, B. Westerlund, B. B. Bederson, A. Druin, C. Plaisant, M. Beaudouin-Lafon, S. Conversy, H. Evans, H. Hansen, N. Roussel, and B. Eiderbäck, ”Technology probes: inspiring design for and with families”, In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Ft. Lauderdale, Florida, USA, April 05 - 10, 2003.
[29] S.E. Hudson, ”Using Light Emitting Diode Arrays as Touch-Sensitive Input and Output Devices”, In *Proceedings of UIST’04*, Paris, France, 2004.