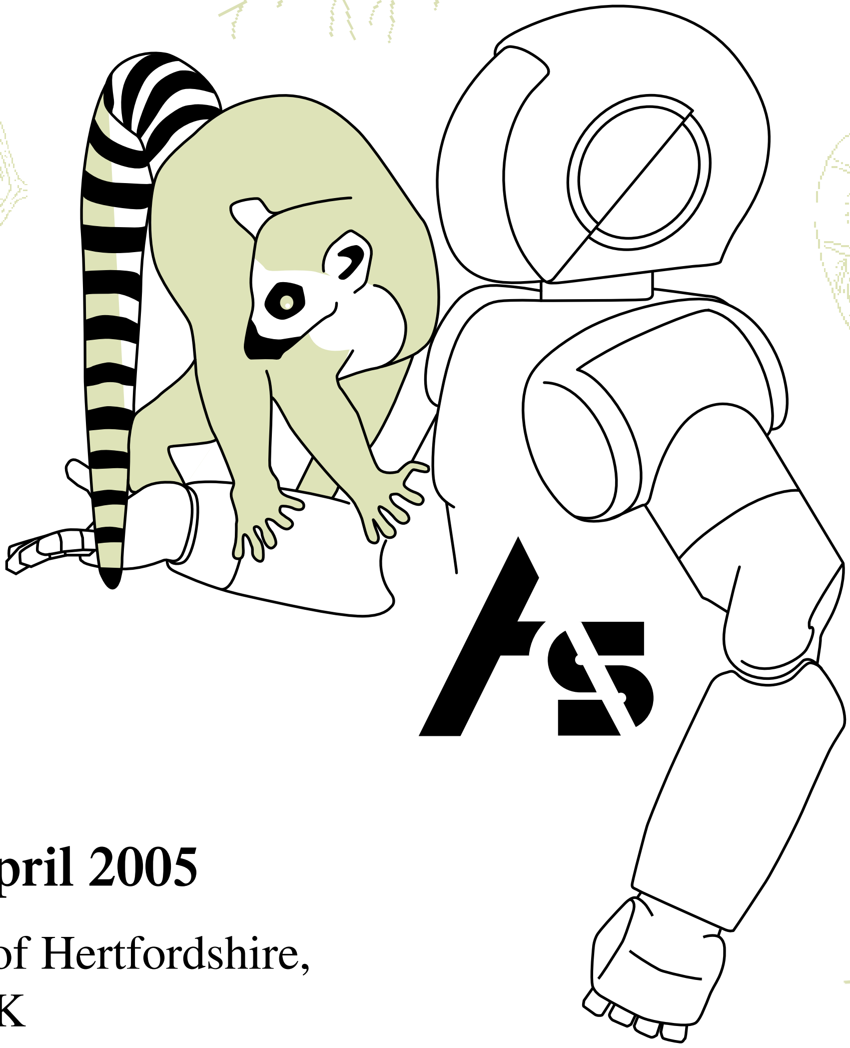
AISB'05: Social Intelligence and Interaction in Animals, Robots and Agents

Proceedings of the Symposium on Robot Companions: Hard Problems and Open Challenges in Robot-Human Interaction



12 - 15 April 2005

University of Hertfordshire,
Hatfield, UK

SSAISB 2005 Convention

AISB UH EPSRC Engineering and Physical Sciences Research Council

AISB'05 Convention

Social Intelligence and Interaction in Animals, Robots and Agents

12-15 April 2005
University of Hertfordshire, Hatfield, UK

Proceedings of the Symposium on

Robot Companions:
Hard Problems and Open Challenges in
Robot-Human Interaction

Published by

AISB

The Society for the Study of Artificial Intelligence and the Simulation of Behaviour
www.aisb.org.uk

Printed by

UH

The University of Hertfordshire, Hatfield, AL10 9AB UK
www.herts.ac.uk

Cover Design by Sue Attwood

ISBN 1 902956 44 1

AISB’05 Hosted by

The Adaptive Systems Research Group
adapsys.feis.herts.ac.uk

The AISB'05 Convention is partially supported by:

EPSRC Engineering and Physical Sciences Research Council

The proceedings of the ten symposia in the AISB’05 Convention are available from SSAISB:

Second International Symposium on the Emergence and Evolution of Linguistic Communication (EELC'05)
1 902956 40 9

Agents that Want and Like: Motivational and Emotional Roots of Cognition and Action
1 902956 41 7

Third International Symposium on Imitation in Animals and Artifacts
1 902956 42 5

Robotics, Mechatronics and Animatronics in the Creative and Entertainment Industries and Arts
1 902956 43 3

Robot Companions: Hard Problems and Open Challenges in Robot-Human Interaction
1 902956 44 1

Conversational Informatics for Supporting Social Intelligence and Interaction - Situational and Environmental Information Enforcing Involvement in Conversation
1 902956 45 X

Next Generation approaches to Machine Consciousness: Imagination, Development, Intersubjectivity, and Embodiment
1 902956 46 8

Normative Multi-Agent Systems
1 902956 47 6

Socially Inspired Computing Joint Symposium (Memetic theory in artificial systems & societies, Emerging Artificial Societies, and Engineering with Social Metaphors)
1 902956 48 4

Virtual Social Agents Joint Symposium (Social presence cues for virtual humanoids, Empathic Interaction with Synthetic Characters, Mind-minding Agents)
1 902956 49 2

Table of Contents

Poster presentations

The AISB'05 Convention
Social Intelligence and Interaction in Animals, Robots and Agents

Above all, the human animal is social. For an artificially intelligent system, how could it be otherwise?

We stated in our Call for Participation "The AISB'05 convention with the theme *Social Intelligence and Interaction in Animals, Robots and Agents* aims to facilitate the synthesis of new ideas, encourage new insights as well as novel applications, mediate new collaborations, and provide a context for lively and stimulating discussions in this exciting, truly interdisciplinary, and quickly growing research area that touches upon many deep issues regarding the nature of intelligence in human and other animals, and its potential application to robots and other artefacts".

Why is the theme of Social Intelligence and Interaction interesting to an Artificial Intelligence and Robotics community? We know that intelligence in humans and other animals has many facets and is expressed in a variety of ways in how the individual in its lifetime - or a population on an evolutionary timescale - deals with, adapts to, and co-evolves with the environment. Traditionally, social or emotional intelligence have been considered different from a more problem-solving, often called "rational", oriented view of human intelligence. However, more and more evidence from a variety of different research fields highlights the important role of social, emotional intelligence and interaction across all facets of intelligence in humans.

The Convention theme *Social Intelligence and Interaction in Animals, Robots and Agents* reflects a current trend towards increasingly interdisciplinary approaches that are pushing the boundaries of traditional science and are necessary in order to answer deep questions regarding the social nature of intelligence in humans and other animals, as well as to address the challenge of synthesizing computational agents or robotic artifacts that show aspects of biological social intelligence. Exciting new developments are emerging from collaborations among computer scientists, roboticists, psychologists, sociologists, cognitive scientists, primatologists, ethologists and researchers from other disciplines, e.g. leading to increasingly sophisticated simulation models of socially intelligent agents, or to a new generation of robots that are able to learn from and socially interact with each other or with people. Such interdisciplinary work advances our understanding of social intelligence in nature, and leads to new theories, models, architectures and designs in the domain of Artificial Intelligence and other sciences of the artificial.

New advancements in computer and robotic technology facilitate the emergence of multi-modal "natural" interfaces between computers or robots and people, including embodied conversational agents or robotic pets/assistants/companions that we are increasingly sharing our home and work space with. People tend to create certain relationships with such socially intelligent artifacts, and are even willing to accept them as helpers in healthcare, therapy or rehabilitation. Thus, socially intelligent artifacts are becoming part of our lives, including many desirable as well as possibly undesirable effects, and Artificial Intelligence and Cognitive Science research can play an important role in addressing many of the huge scientific challenges involved. Keeping an open mind towards other disciplines, embracing work from a variety of disciplines studying humans as well as non-human animals, might help us to create artifacts that might not only do their job, but that do their job right.

Thus, the convention hopes to provide a home for state-of-the-art research as well as a discussion forum for innovative ideas and approaches, pushing the frontiers of what is possible and/or desirable in this exciting, growing area.

The feedback to the initial Call for Symposia Proposals was overwhelming. Ten symposia were accepted (ranging from one-day to three-day events), organized by UK, European as well as international experts in the field of Social Intelligence and Interaction.

- Second International Symposium on the Emergence and Evolution of Linguistic Communication (EELC'05)
- Agents that Want and Like: Motivational and Emotional Roots of Cognition and Action
- Third International Symposium on Imitation in Animals and Artifacts
- Robotics, Mechatronics and Animatronics in the Creative and Entertainment Industries and Arts
- Robot Companions: Hard Problems and Open Challenges in Robot-Human Interaction
- Conversational Informatics for Supporting Social Intelligence and Interaction - Situational and Environmental Information Enforcing Involvement in Conversation
- Next Generation Approaches to Machine Consciousness: Imagination, Development, Intersubjectivity, and Embodiment
- Normative Multi-Agent Systems
- Socially Inspired Computing Joint Symposium (consisting of three themes: Memetic Theory in Artificial Systems & Societies, Emerging Artificial Societies, and Engineering with Social Metaphors)
- Virtual Social Agents Joint Symposium (consisting of three themes: Social Presence Cues for Virtual Humanoids, Empathic Interaction with Synthetic Characters, Mind-minding Agents)

I would like to thank the symposium organizers for their efforts in helping to put together an excellent scientific programme.

In order to complement the programme, five speakers known for pioneering work relevant to the convention theme accepted invitations to present plenary lectures at the convention: Prof. Nigel Gilbert (University of Surrey, UK), Prof. Hiroshi Ishiguro (Osaka University, Japan), Dr. Alison Jolly (University of Sussex, UK), Prof. Luc Steels (VUB, Belgium and Sony, France), and Prof. Jacqueline Nadel (National Centre of Scientific Research, France).

A number of people and groups helped to make this convention possible. First, I would like to thank SSAISB for the opportunity to host the convention under the special theme of *Social Intelligence and Interaction in Animals, Robots and Agents*. The AISB'05 convention is supported in part by a UK EPSRC grant to Prof. Kerstin Dautenhahn and Prof. C. L. Nehaniv. Further support was provided by Prof. Jill Hewitt and the School of Computer Science, as well as the Adaptive Systems Research Group at University of Hertfordshire. I would like to thank the Convention's Vice Chair Prof. Chrystopher L. Nehaniv for his invaluable continuous support during the planning and organization of the convention. Many thanks to the local organizing committee including Dr. René te Boekhorst, Dr. Lola Cañamero and Dr. Daniel Polani. I would like to single out two people who took over major roles in the local organization: Firstly, Johanna Hunt, Research Assistant in the School of Computer Science, who efficiently dealt primarily with the registration process, the AISB'05 website, and the coordination of ten proceedings. The number of convention registrants as well as different symposia by far exceeded our expectations and made this a major effort. Secondly, Bob Guscott, Research Administrator in the Adaptive Systems Research Group, competently and with great enthusiasm dealt with arrangements ranging from room bookings, catering, the organization of the banquet, and many other important elements in the convention. Thanks to Sue Attwood for the beautiful frontcover design. Also, a number of student helpers supported the convention. A great team made this convention possible!

I wish all participants of the AISB’05 convention an enjoyable and very productive time. On returning home, I hope you will take with you some new ideas or inspirations regarding our common goal of understanding social intelligence, and synthesizing artificially intelligent robots and agents. Progress in the field depends on scientific exchange, dialogue and critical evaluations by our peers and the research community, including senior members as well as students who bring in fresh viewpoints. For social animals such as humans, the construction of scientific knowledge can't be otherwise.

Beppu, Japan.

Dedication:

I am very confident that the future will bring us increasingly many instances of socially intelligent agents. I am similarly confident that we will see more and more socially intelligent robots sharing our lives. However, I would like to dedicate this convention to those people who fight for the survival of socially intelligent animals and their fellow creatures. What would 'life as it could be' be without 'life as we know it'?

Kerstin Dautenhahn

Professor of Artificial Intelligence,
General Chair, AISB'05 Convention *Social Intelligence and Interaction in Animals, Robots and Agents*

University of Hertfordshire
College Lane
Hatfield, Herts, AL10 9AB
United Kingdom

Symposium Preface
Robot Companions:
Hard Problems and Open Challenges in Robot-Human Interaction

SYMPOSIUM OVERVIEW

Human-Robot Interaction (HRI) is a growing and increasingly popular research area at the intersection of research fields such as robotics, psychology, ethology and cognitive science. Robots moving out of laboratory and manufacturing environments face hard problems of perception, action and cognition. Application areas that heavily involve human contact are a particularly challenging domain. Interaction and communication of embodied physical robots with humans is multimodal, and involves deep issues of social intelligence and interaction that have traditionally been studied e.g. in social sciences. The design of a robot's behaviour, appearance, and cognitive and social skills is highly challenging, and requires interdisciplinary collaborations across the traditional boundaries of established disciplines. It addresses deep issues into the nature of human social intelligence, as well as sensitive ethical issues in domains where robots are interacting with vulnerable people (e.g. children, elderly, people with special needs).

Assuming that the future will indeed give us a variety of different robots that inhabit our homes, it is at present not quite clear what roles the robots will adopt. Will they be effective machines performing tasks on our behalf, assistants, companions, or even friends? What social skills are desirable and necessary for such robots? People have often used technology very differently from what the designers originally envisaged, so the development of robots designed to interact with people requires a careful analysis and study of how people interact with robots and what roles a new generation of robot companions should adopt.

The symposium will present state-of-the-art in the field of HRI, focussing on hard problems and open challenges involved in studying 'robot companions'. The symposium will consist of invited talks as well as regular presentations. The invited speakers are: Christoph Bartneck (Eindhoven University of Technology, The Netherlands), Aude Billard (EPFL, Switzerland), Guido Bugmann (University of Plymouth, UK), Henrik I. Christensen (KTH, Sweden), Takayuki Kanda (ATR Intelligent Robotics - Communication Labs, Japan), Gerhard Sagerer (University of Bielefeld, Germany), and Takanori Shibata (AIST, Japan)

Topics relevant to the symposium are:

- Design of social robots for HRI research
- Requirement for socially interactive robots for HRI research
- Cognitive skills for robot companions
- Evaluation methods in HRI research
- Ethical issues in HRI research
- Creating relationships with social robots
- Developmental aspects of human-robot interaction
- Roles of robots in the home
- others

We would like to thank the Programme Committee for their assistance in reviewing the symposium submissions:

- Christoph Bartneck (Eindhoven University of Technology The Netherlands)
- René te Boekhorst (Adaptive Systems Research Group, University of Hertfordshire)
- Henrik I. Christensen (KTH, Sweden)
- Guido Bugmann (University of Plymouth, UK)
- Kerstin Dautenhahn (Adaptive Systems Research Group, University of Hertfordshire)
- Takayuki Kanda (ATR Intelligent Robotics - Communication Labs, Japan)
- Tatsuya Nomura (Ryukoku University, Japan)
- Gerhard Sagerer (University of Bielefeld, Germany)
- Takanori Shibata (AIST, Japan)

We intend that the symposium will contribute to the process of establishing a common understanding of important research directions, approaches, theories and methods in HRI. Last but not least, we hope that all presenters and participants will enjoy the symposium and interactions among its participants, as much as they enjoy working with social robots!

Kerstin Dautenhahn, René te Boekhorst (Symposium chairs)

Cultural Differences in Attitudes Towards Robots

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

Tatsuya Nomura
Department of Media Informatics
Ryukoku University
1-5, Yokotani, SetaOhe-Cho, Otsu
Shiga 520-2194, Japan
nomura@rins.ryukoku.ac.jp

Takayuki Kanda
ATR
IRC, Department 2,
2-2, Hikaridai, Seika-cho, Soraku-gun
Kyoto 619-0288, Japan
kanda@atr.jp

Tomohiro Suzuki
Graduate School of Sociology,
Toyo University
5-28-20, Hakusan, Bunkyo-ku,
Tokyo 112--8606, Japan
suzukirt@h9.dion.ne.jp

Kennsuke Kato
Graduate School of Human Science
Osaka University
1-2 Yamadaoka, Suita, Suita
Osaka 565-0871, Japan
DZL02550@nifty.ne.jp

Abstract

This study presents the result of a cross-cultural study of negative attitude towards robots. A questionnaire was presented to Dutch, Chinese, German, Mexican, American (USA) and Japanese participants based on the Negative Attitude towards Robots Scale (NARS). The American participants were least negative towards robots, while the Mexican were most negative. Against our expectation, the Japanese participants did not have a particularly positive attitude towards robots.

1 Introduction

The United Nations (UN), in a recent robotics survey, identified personal service robots as having the highest expected growth rate (UN, 2002). These robots help the elderly (Hirsch et al., 2000), support humans in the house (NEC, 2001), improve communication between distant partners (Gemperle, DiSalvo, Forlizzi, & Yonkers, 2003) and are research vehicles for the study on human-robot communication (Breazeal, 2003; Okada, 2001). A survey of relevant characters is available (Bartneck, 2002; Fong, Nourbakhsh, & Dautenhahn, 2003).

It appears that different cultures have a different exposure to robots through media or through personal experience. The number of humanoids robots, toy robots, games and TV shows give Japan the leading role in robotic development and culture. However, the typical "robots will take over the world" scenario that is so often used in western culture (Cameron, 1984; Wachowski & Wachowski, 2003) is less present in Japan. Yamamoto (1983) hypothesized that Confucianism might have had an influence on the positive development of robot culture in Japan. In the popular Japanese Manga movies good fights evil just like in the western world, but the role of the good and the evil is not mapped directly to humans as being the good against robots being the evil. In these movies the good and the evil are distributed. You might have a good robot that fights an evil human villain or a good robot fighting bad robots.

If we are to employ more and more robots in daily life it appears necessary to study what attitude the users have towards robots, which of course depend on culture.

Computer anxiety prevents users from using computers and educational psychologists have studied its effects in great detail (Raub, 1981). However, the effects of robot anxiety are still largely unknown. With an increasing number of robots, robot anxiety might become as important as computer anxiety is today.

2 Method

We therefore conducted a cross-cultural study that investigated the attitude towards robots. We presented 28 Dutch, 20 Chinese (living in the Netherlands), 69 German, 16 Mexican, 22 American (USA) and 53 Japanese participants a questionnaire based on the Negative Attitude towards Robots Scale (NARS). The original Japanese questionnaire was first translated to English and then to all other languages using the forth and back translation process.

Most of the participants were university students. The validity of the questionnaire has been previously assessed (Nomura, Kanda, & Suzuki, 2004). The questionnaire consisted of 14 items (5-point-scales) in three constructs:

1. attitude towards the interaction with robots *(interact)*
   (e.g. I would feel relaxed talking with robots)
2. attitude towards social influence of robots *(social)*
   (e.g. I am concerned that robots would have a bad influence on children)
3. attitude towards emotions in interaction with robots *(emotion)*
   (e.g. I would feel uneasy if robots really had emotions)

In the following text we will use the *italic style* to highlight the dependent variables.

3 Results

Table 1 presents the means and standard deviations of all measurements for all nationalities. An analysis of Covariance (ANCOVA) was performed in which nationality and gender were the independent variables. *Interact*, *social* and *emotion* were the dependent variables and age the covariant. Gender had no significant influence on the measurements. Nationality had a significant influence on *interact* ($F(5)=38.775$, $p<.001$), *social* ($F(5)=6.954$, $p<.001$) and *emotion* ($F(5)=5.004$, $p<.001$). Age had a significant ($F(1)=7.998$, $p=.005$) influence on *emotion*. Figure 1 presents the means of all conditions.

The Japanese participants (m=2.05) rated *interact* significantly ($t(73)=3.857$ $p<.001$) more negative than participants from the USA (m=1.49). Furthermore, Mexican participants (m=4.27) rated *interact* significantly ($t(79)=10.283$, $p<.001$) more negative than German participants (m=2.23). There was no significant difference in *interact* between German, Dutch, Chinese and Japanese participants.

		Mean	Std.Dev.
interact	CHN	2.22	0.55
	DEU	2.24	0.73
	JPN	2.05	0.61
	MEX	4.27	0.72
	NLD	2.10	0.68
	USA	1.45	0.50
social	CHN	2.71	0.62
	DEU	3.21	0.87
	JPN	3.17	0.69
	MEX	3.48	0.92
	NLD	2.69	0.60
	USA	2.40	0.79
emotion	CHN	2.77	0.88
	DEU	3.53	0.91
	JPN	3.06	0.79
	MEX	3.46	0.79
	NLD	2.99	0.96
	USA	2.62	0.72

Table 1: Mean and standard deviation of all measurements for all nationalities

For *social*, we could identify two groups that had no significant difference within them, but were significantly different from the other group. The group of German, Mexican and Japanese participants rated social significantly higher ($t(73)=3.807$, $p<.001$) than the group of Chinese, Dutch and American participants.

We found three groups of nationalities in the *emotion* measurement that were not significantly different within themselves, but different compared to the other groups. German (m=3.51) and Mexican participants rated *emotion* significantly ($t(116)=2.755$, $p=.007$) higher than Japanese (m=3.08) participants. The later rated *emotion* significantly ($t(73)=2.176$, $p=.033$) higher than American participants.

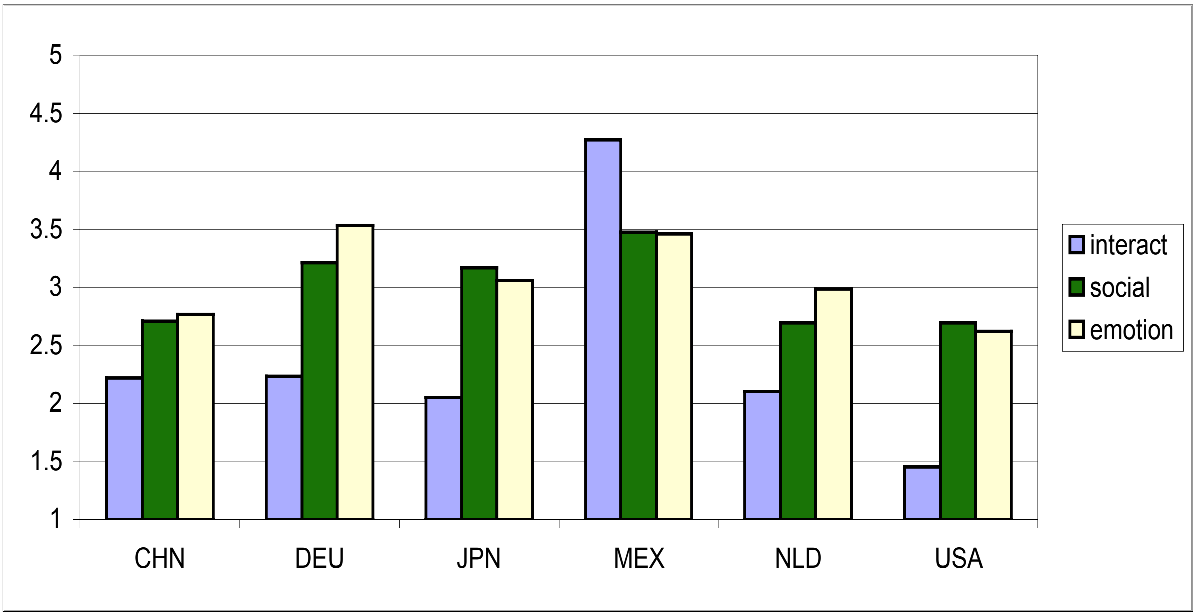


Figure 1: Means for all nationalities.

3 Conclusions

In contradiction to the popular believe that Japanese love robots our results show that the Japanese are concerned with the impact that robots might have on society. A possible explanation could be that through the high exposure to robots, the Japanese are more aware of robots abilities and also their lack of abilities.

Participants from the USA were least negative towards robots, in particular on the aspect of interacting with them. A possible reason could be that they are used to technology and at the same time easy going when it comes to talking to new people. Another striking difference can be found when looking at the ratings of the Mexican participants. They were most negative towards robots, in particular towards interacting with them. This is surprising, since they are a neighbor state of the USA which were least concerned.

The prior experience that the participants had with robots, such as a personal interaction with a robot, was not assessed by the NARS questionnaire. This experience might have an influence on the results and we are currently preparing to administer the questionnaire to owners of the Sony's robotic dog Aibo. In addition, we are planning to conduct the experiment in other eastern and western countries.

3 Acknowledgements

We would like to thank Jodi Forlizzi, Oscar Mayora Ibarra, Hu Jun and Juliane Reichenbach for their generous help in gathering the data. In addition, we would like to thank Chi Ho Chan, David Cournapeau, Nathalia Romero Herrera, Alice Jager, Roberto Lopez and Machi Takahachi for their efforts in translating the questionnaire.

References

Bartneck, C. (2002). *eMuu - an embodied emotional character for the ambient intelligent home*. Unpublished Ph.D. thesis, Eindhoven University of Technology, Eindhoven.

Breazeal, C. (2003). *Designing Sociable Robots*. Cambridge: MIT Press.

Cameron, J. (Writer), & J. Cameron (Director) (1984). The Terminator [DVD]. In J. Daly (Producer): MGM.

Fong, T., Nourbakhsh, I., & Dautenhahn, K. (2003). A survery of socially interactive robots. *Robotics and Autonomous Systems, 42*, 143-166.

Gemperle, F., DiSalvo, C., Forlizzi, J., & Yonkers, W. (2003). *The Hug: A new form for communication*. Paper presented at the Designing the User Experience (DUX2003), New York.

Hirsch, T., Forlizzi, J., Hyder, E., Goetz, J., Stroback, J., & Kurtz, C. (2000). *The ELDeR Project: Social and Emotional Factors in the Design of Eldercare Technologies*. Paper presented at the Conference on Universal Usability, Arlington.

NEC. (2001). *PaPeRo*, from http://www.incx.nec.co.jp/robot/PaPeRo/english/p_index.html

Nomura, T., Kanda, T., & Suzuki, T. (2004). *Experimental Investigationi into Influence of Negative Attitudes toward Robots on Human-Robot Interaction*. Paper presented at the 3rd Workshop on Social Intelligence Design (SID2004), Twente.

Okada, M. (2001, 2001). *Muu: Artificial Creatures as an Embodied Interface*. Paper presented at the ACM Siggraph 2001, New Orleans.

Raub, A. C. (1981). *Correlates of computer anxiety in college students*. Unpublished Ph.D. Thesis, University of Pennsylvania.

UN. (2002). *United Nations and the International Federation of Robotics*. Paper presented at the World Robotics 2002, New York.

Wachowski, A., & Wachowski, L. (Writer), & P. Chung & A. Jones (Director) (2003). Animatrix.

Yamamoto, S. (1983). Why the Japanese have no allergy towards robots. *L'esprit d'aujourd'hui (Gendai no Esupuri), 187*, 136-143.

Alternative model-building for the study of socially interactive robots

Meurig Beynon, Antony Harfield, Sunny Chang
*Department of Computer Science, University of Warwick, UK
{wmb,ant,csvmu}@dcs.warwick.ac.uk

Abstract

In this discussion paper, we consider the potential merits of applying an alternative approach to model-building (Empirical Modelling, also known as EM – see http://www.dcs.warwick.ac.uk/modelling) in studying social aspects of human-robot interaction (HRI). The first section of the paper considers issues in modelling for HRI. The second introduces EM principles, outlining their potential application to modelling for HRI and its implications. The final section examines the prospects for applying EM to HRI from a practical perspective with reference to a simple case study and to existing models.

Introduction

The difficulty of dealing effectively with issues relating to social intelligence in the design of robots is widely recognised. In discussing this challenge, Fong et al. (2002) identify two approaches to the design of socially intelligent agents, the "biological" and the "functional". The biological approach aims to draw on understanding of animals and their behaviour to design robots which exhibit similar properties to their biological counterparts. The functional approach only takes the functionality of such robots into account and is not concerned with the mechanisms by which this is achieved. Traditional AI generally takes a functional approach. The biological approach is favoured by those interested in the social sciences and biology.

Whatever the orientation of the robot design, there are major technical and conceptual issues to be addressed in developing robots that are socially responsive. It is implausible that such issues can be resolved by implementing behaviours that are comprehensively pre-specified by abstract analysis of the operating context. Dautenhahn (2004) proposes that the social personality of a robot should grow through a socialisation process similar to that observed in animals such as dogs. Adams et al. (2000) sees robotics as offering "a unique tool for testing models drawn from developmental psychology and cognitive science". His approach to building sophisticated robots incrementally, using the concept of a subsumption architecture (Adams et al., 2000; Brooks, 1991), indicates a possible way in which such a socialisation process might be supported. However, as Dautenhahn (1995) has observed, the role of 'the social factor' in the development of intelligence has been little explored in the 'sciences of the artificial', and we cannot necessarily expect that techniques for building intelligent robots will deal with social aspects. Adaptation to the social environment is likely to be a much more subtle process than adaptation to a physical context, and demands a more intimate interaction between human and automated activities than has been achieved hitherto.

This paper examines the prospects for deploying Empirical Modelling (EM) in HRI research. EM is an unconventional approach to modelling that reflects a fundamental shift in emphasis in the science of computing. In certain respects, this shift echoes Brooks's outlook. EM favours the construction of physical artefacts that in some sense 'embody knowledge', rather than abstract representations based on formal languages. It also promotes an evolutionary and incremental approach: models are initially very simple, but can eventually attain a high level of sophistication. For the present, EM research is not specifically concerned with how learning or other forms of adaptation might take place automatically. The focus of interest is rather on maintaining an intimate connection between the developing model and the modeller's perception and understanding, which grow in parallel throughout the model-building. The model development is sharply distinguished from other approaches by its emphasis on incrementally layering 'perceptions of relations' rather than 'functional behaviours'. In this way, the primary focus is enhancing the robot's capacity to respond to its current situation rather than on extending its current repertoire of behaviours.

1 Issues in modelling for Human Robot Interaction

Traditional techniques for modelling have problematic aspects in the context of robotics. Closed-world models of robot behaviour may appear to give useful insights in the abstract, but the vagaries of the physical world lead to serious discrepancies between real and virtual behaviours. It is such considerations that prompt Brooks to advocate '[using] the world as its own model' (Brooks, 1991). There is little doubt that problems of this nature will always be an issue, but EM is such a radical alternative to traditional modelling approaches (Beynon, 1999, 2003) that there is hope that it can offer new remedies or palliatives.

Our objective is to develop modelling techniques that can be used in direct and live conjunction with researches on actual robots in a laboratory. The aspiration is to make models that are sufficiently subtle to address social aspects of the interaction to some degree. There are many ways in which empirical study of HRI in the laboratory can potentially be combined with experiments in a virtual environment. We might wish to use virtual models to explore experimental scenarios for which empirical data has been derived, or to connect the behaviour of agents in the physical environment directly to that of their avatars in the virtual world. Possible goals might be formulating new hypotheses, making a virtual record of significant interactions observed in the laboratory, or identifying new patterns of robot behaviour to be programmed. For these purposes, models need to be sufficiently authentic that they can guide the programming of robots. Ideally, we would like to be able to direct the modelling activity freely in a variety of different ways, corresponding to different modes of observing the HRI, mixing modes of observation and experiment in real and virtual worlds.

In the context of modelling for HRI, we identify the following issues as particularly significant:

- having an approach to model development that is incremental, admits highly flexible adaptation through human intervention (because social conventions and physical interactions are so subtle and difficult to circumscribe), and is holistic (because, for instance, social conventions about personal space (Hall, 1966) and lower-level concerns such as navigation are inseparably linked).
- developing models that have explanatory power, so as to be able to trace the effects of robot action to their origins, attribute responses to stimuli appropriately and account for the fact that the robot does more than can be specified and represented in propositional terms.
- interrelating human and machine perspectives intimately so as to be able to exploit the qualities of both humans and robots, as is required to program robots to achieve the high degree of attentiveness to context that is demanded in social situations without compromising their capability to act obliviously with superhuman efficiency where appropriate.

Various kinds of relation have a significant impact upon social interaction. These include:

- Spatial relations - An agent's physical location and the surrounding space are likely to affect the behaviour of the agent. Actions in small confined spaces are usually different from those in large open spaces.
- Temporal relations - Time plays a significant role in human behaviour. When time is at a premium humans are likely to perform tasks differently from when they have plenty of time.
- Status relations - The status of human agents affects their interaction and expectations. Interaction with those with whom we are familiar differs from interaction with strangers. Interaction within the working environment, families and cultural contexts is likewise differentiated according to the status of the agents with whom we are interacting.

Taking account of such relations in interaction is something that humans typically learn from experience. On this basis, a most important characteristic in modelling for HRI is a capacity to accommodate learning in a dynamic fashion. This has particular relevance for the prospects of applying EM to HRI because EM proceeds by modelling relations as they are encountered in experience.

2 The Empirical Modelling Approach

The Empirical Modelling approach to HRI will be sketched with reference to the role played by the primary concepts – agents, observables and dependencies – and to the general characteristics of the development of a model as a *construal*.

2.1 Agents and Observables

Empirical Modelling (EM) approaches the construction of a model of a concurrent system from the perspective of an external observer trying to make sense of what is being observed (Beynon et al., 1990). If the task is to make a virtual representation of a physical system, the principles of EM can be seen as similar to identifying the situation within the context of familiar 'scientific' theory, complemented – where there is no such theory to hand – by the application of the 'scientific method'. In this context, the modeller identifies what they perceive to be the principal agents responsible for state change, and develops hypotheses about how their interaction is mediated by observables. This section will introduce EM as it might apply to the scenario of studying the social behaviour of robots, without particular concern for the technical challenges this might present for EM tools and other relevant technologies in their current state of development. Specific models that indicate the present state of EM development in key aspects will be discussed in the next section.

In the HRI laboratory, the most prominent agents are the robots and the humans who interact with them. Within the scope of the EM model, other elements of the situation are typically also construed as agents. For instance, an item of furniture, a door or a pet might be an agent in so far as its presence and current status influences the state change that takes place. If, moreover, there is some abstract activity in process, such as might be informally described as 'the robot is going to collect up the empty wine glasses', this too would naturally be represented by an agent of some kind. Relevant issues to be noted are:

- the concept of an observable is intended to embrace not only what the external observer can directly apprehend, but what the agents within the system are deemed to directly apprehend. For instance, the 'observables' relevant to the robot might include information from its distance sensor along a particular direction, and information about the status of the current task in hand.

- it is generally necessary to take account of the transient nature of observables, so as to reflect the presence or absence of agents in the situation. For instance, when the task of collecting empty wine glasses is accomplished or aborted, the related observables are no longer meaningful.

Because the model-building activity serves an explanatory function, it is appropriate to characterise an EM model as a 'construal' (cf. the extended discussion of this term in (Gooding, 1990)). Note that, in arriving at a construal, the external observer has to project agency that is human-like on to the non-human agents in the situation. For instance, to explain the behaviour of an automatic door, the modeller may postulate an observable by which the door 'perceives' itself as open, and consider the door to be responsible for manipulating its aperture accordingly.

2.2 Dependencies

Agents and observables are complemented by additional features of the situation that are most distinctive of EM – dependencies. A dependency is a relation between changes to observables that pertains in the view of an agent within the system. In effect, there are latent relationships between those things that an agent is deemed to observe, that are 'perceived' by the agent to be linked in change in an indivisible manner. This indivisibility is in general 'relative to the agent', and its status depends upon the nature of the construal. For instance, in some contexts, the activity of 'collecting an empty wine glass' might be viewed by the external observer as an atomic operation that indivisibly reduces the count of empty wine glasses so far accounted for. Where the robot is concerned, on the other hand, such an operation would necessarily involve a highly complex and intricate sequence of sensing and activation steps.

By their nature, the key concepts of EM are defined by experience. What is deemed to be an agent, an observable or a dependency is at all times subject to change through observation and experiment on the part of the modeller (cf. the way in which varieties of agency are seen to be socially constructed in (Dautenhahn, 1998)). The through-and-through empirical nature of these constituents is reflected in the character of the construal itself, which is conceived and developed quite differently from a traditional computer model.

In the first place, there is no notion of a static or comprehensive functional specification of the modeller's construal. The construal itself takes the form of a physical artefact, or set of artefacts, to represent a current situation and understanding on the part of the modeller; it embodies the patterns of agency, dependency and observation that are deemed to pertain in the situation. When a system has been - for certain practical purposes - comprehensively identified and understood, there will be a single unifying artefact that captures all the observables within the modeller's construal and represents the viewpoint and insight of the external observer. In so far as these observables have specific current values, the artefact itself will serve to represent the current state of the system to which it refers (cf. the way that a spreadsheet records the current status of a financial account). The atomic changes of state that are possible in this state will be represented by possible redefinitions to which appropriate observables are subject, whose impact is in general to change the values of several observables simultaneously, and perhaps change the pattern of dependencies itself. In the HRI laboratory scenario, such an atomic change might typically reflect an 'infinitesimal' movement or sensory update on the part of the robot, or a primitive action on the part of a human agent, such as pressing the television remote control. Note that - because of the dependencies - a single action on the part of an agent may update several observables simultaneously (as when pressing the remote switches the television on). There is also the possibility for independent changes of state to occur simultaneously (as when the robot moves, and the human agent presses the remote control at the same time). The modeller can make use of such a construal to trace characteristic system behaviours, though the effect is quite unlike the exercising of statically pre-specified behaviours in a closed-world that is commonplace in conventional computer programming. Suppose for example that the robot is programmed to collect the empty wine glasses, but that at some point during this collection process one of the wine glasses is accidentally smashed into pieces. It then becomes necessary to adapt the parameters of the collection activity to take account of the new situation - something which the modeller should be able to cope with dynamically when exercising a good construal of the situation, but would have had to have been within the explicit scope of a programmed behaviour.

2.3 Developing a construal

As the above discussion highlights, the development of an EM construal is concerned with something less specific than representing any particular set of functionalities. For any complex reactive system, the goal of developing a single unifying artefact to reflect the modeller's comprehensive understanding is a pipe dream. The quality of a construal is contingent upon the degree of familiarity and understanding that the modeller has built up through observation and experiment, typically over an extended period of interaction. The true potential and limitations of EM in concurrent systems modelling are best appreciated by viewing the construal not in some purported final perfected form, but as it evolves in conjunction with the development of the modeller's understanding. In applications such as HRI modelling, it is plausible that this development should ideally accompany the construction of the real environment from its inception, so that the model grows in subtlety and scope in counterpoint with the understanding of the laboratory technicians and experimenters. To conclude this brief overview of EM principles, it will be helpful to outline informally how such an incremental process of construal might take place.

Throughout the development process, the representation of the construal has two aspects: the physical artefact as it is realised on a computer, or more precisely using an appropriate suite of computer-based peripherals (cf. the distinction between a musical instrument and an orchestra), and documentation in the form of a textual description of the agents, observables and dependencies and their interrelationship within the modeller's construal. As will be illustrated below, in our implementation framework, these two ingredients of the construal are respectively associated with a set of *definitive scripts*, and a set of *LSD accounts* of the agents, to be referred to as 'scripts' and 'accounts' in what follows. An LSD account classifies the observables deemed to shape the behaviour of an agent, with reference to how it perceives, acts upon and responds to its environment. To put these ingredients in context, it is quite plausible that, in the HRI scenario, we might have a good grasp of the determinants of the robot behaviour in the abstract, and reasonable models for its behaviour in certain idealised scenarios (e.g. robot motion where the floor is level and the coefficient of friction is uniform, and the lighting conditions are favourable). We may also have reliable knowledge of the characteristics of the physical environment where issues such as the location of furniture and the operation of doors and light switches are concerned. Such information provides the basis for developing several ingredients that contribute to a useful construal. These might include:

- scripts to represent the principal features of the environment in which the robots and human agents interact.
- an account of a robot's behaviour with reference to the observables that are intrinsically associated with it (such as the current status of its sensors, its location and velocity), together with the external observables to which it responds.
- a script to represent a test environment within which idealised motion of a robot can be inves-

tigated experimentally, and interactively adapted through intervention by the modeller.

In this scenario, many more difficult issues remain to be addressed, such as understanding the relationship between what the robot sensors record (e.g. the distance from the nearest object in some direction) and how this needs to be interpreted in context (as in 'the robot is approaching the table on which there is a wine glass'): these will typically require extensive empirical investigation.

By its nature, an EM construal can accommodate partial understanding and support the modeller in gaining further insight. Though there is not typically one unifying script to represent the entire system comprehensively from an objective external observer's perspective, there will be a collection of subscripts associated with those scenarios for which the modeller has sufficiently detailed understanding. As explained in the above discussion, the behaviours that can be exercised using these scripts are open for the modeller to explore and extend in an experimental fashion. What is more, the behavioural interpretation of the construal can be modified by the modeller 'in-the-stream-of-thought'. This is in sharp contrast to modifying the behaviour of a conventional program, which entails terminating execution, changing the specification and attempting to reconstruct what – taking the changed specification into account – can only be an approximation to the original situation of use. It is also conceptually easy to exercise scripts representing independent aspects of the same situation in combination, as is appropriate where understanding of a situation is too partial to support a conventional implementation of behaviour, but significant behaviours can be explored subject to intervention by the modeller. Taking in conjunction, scripts and accounts also serve as a powerful way of communicating understanding between experimenters.

3 Practical Aspects of Empirical Modelling

This section illustrates how EM techniques can be applied in practice. The scenarios considered relate to interactions between humans and robots that might arise in a house environment. They help to indicate how EM might be used to support the development of a robot that exhibits some degree of social awareness. Our illustrative examples draw upon pre-existing EM models of a house environment, and of various activities that give insight into the potential for effective modelling of human and robot interaction.

3.1 Agent-oriented modelling

Though the term is widely and frequently used, the Artificial Intelligence (AI) community has great difficulty in agreeing on a definition for 'agent'. As Wooldridge and Jennings (1994) point out: "This need not necessarily be a problem: after all, if many people are successfully developing interesting and useful applications, then it hardly matters that they do not agree on potentially trivial terminological details.". This point of view is strongly endorsed by EM, where the implementation and interpretation of a specific pattern of activity that is conceptually associated with one and the same agent evolves with the model. In a typical pattern of model evolution, a pattern of behaviour that is initially carried out by a human agent can be progressively more comprehensively executed automatically, so that eventually it can be exercised without – or more precisely, in what seem to be the only realistic circumstances, without – the intervention of the human agent. What adds particular force to Wooldridge's observation in this context is that it is not appropriate in EM to conceive the evolution of a model in terms of a discrete succession of progressively more expressive models, each with its own distinctive functionality. In so far as it makes sense to speak of the identity of an EM model, it is more appropriate to think of this identity as unchanging throughout its development, in the same spirit in which we say that 'the child is father to the man'.

By way of illustration, consider the situation where a robot has to negotiate a corridor in which there is a person walking towards it. This situation is encountered by millions of people everyday as they walk down corridors, paths and streets. Because avoiding someone while walking is something we do with relative ease, it is easy to take it for granted. However, the factors affecting this behaviour are quite complex and reproducing this behaviour in a model is a non-trivial task. In applying EM in this context, it is initially appropriate to think about the robot's actions with reference to how a human agent with the same capacity to sense and react to its environment as the robot might respond. As the modeller's understanding of the issues involved matures, it will become possible to automate the more routine aspects of this response. For instance, the forward motion of the robot along the corridor could be automated, and only its lateral movement could be under the control of the human developer. Typically, successful negotiation of the corridor may be automatable subject to reasonable assumptions about the behaviour of the approaching person, or 'opponent'. There may be no satisfactory strategy if the opponent is malicious and sets out to obstruct the robot's passage. Even where the opponent is benign, there may still be exceptional circumstances in which the familiar parallel side-stepping behaviour arises, when the robot's forward motion may need to be suspended. To overcome this problem, which arises at a rather advanced stage in the modelling, it is in general necessary to combine automation of routine aspects of the robot behaviour with mechanisms for open-ended human intervention when singular scenarios arise. Only when these singular scenarios are understood in sufficient detail does full automation become possible. In the transition from an initial model in which the state change for collision avoidance is predominantly supplied by the modeller to a final model in which this state change can be carried out autonomously by a programmed agent, the nature of the agent effecting the state change evolves in ways that are liable to subvert any but the weakest notion of agency. This is in keeping with the observation by Lind (2000) that, in agent-oriented software engineering, "the conceptual integrity that is achieved by viewing every intentional entity in the system as an agent leads to a much clearer system design".

Our illustrative example can be further elaborated with reference to specific practical tools that support EM. To enable the developer to act in the role of the robot, it is first helpful to give an LSD account of the robot's relationship to its environment (cf. section 2.3). This involves classifying the observables that affect the behaviour of the robot as an agent. Projecting ourselves into the role of the agent, there are some observations that the agent can make about the environment – these determine the observables that are *oracles* to the agent. We might assume, for instance, that the robot agent has sufficient 'visual' capability to be able to identify other agents or static objects, to locate the positions of the other agents that are within the field of vision, and to determine in which direction the other agents are moving (the *state* observables of these agents). We can further suppose that the robot agent has conditionally control over certain observables (its *handles*), and that there are certain dependencies between observables that can be deemed to apply in the view of the agent (its *derivates*). It is then possible to describe simple strategies that a robot might employ with reference to the LSD classification of observables. For instance, one simple avoidance strategy is: *if an agent is in the direction that one is walking then take a step sideways*. This might be captured in an LSD account as shown in Figure 1.

As discussed in section 2.3, there are two aspects

```
agent SimpleAvoidingAgent {
   states
   //observables belonging to the agent
      position_x, position_y,
      direction, potential_collision
   handles
   //observables that the agent controls
      position_x, position_y, direction
   oracles
   //external observables the agent responds to
      opponent_position_x, opponent_position_y,
      opponent_direction
   derivates
   //dependency between observables
      potential_collision =
         (position_x == opponent_position_x) &&
         (direction != opponent_direction)
   protocol
   //trigger actions
      potential_collision -> position_x++,
      ! potential_collision ->
         position_y = position_y + direction
}
```

Figure 1: A simple example of an LSD account. The derivate *potential_collision* highlights the situation where a collision may occur and the protocol specifies a change in *position_x* aimed at avoiding a collision.

to the development of a construal in EM: the construction of a physical artefact on the computer, and the associated documentation of the modeller's construal. The physical artefact is a source of experience for the modeller that metaphorically represents perceptions of the environment by a whole range of agents. Figure 2 for example, is a snapshot from an EM model of collision avoidance developed by Warwick student Chris Martin in his final year project in 2003-4 (see Figure 2). The geometric elements of the figure are lines and circles that represent the paths traced by two agents, their fields of vision and current locations and headings. The perspective associated with the model is that of an external observer monitoring the behaviour of two people passing each other in a corridor, as if viewed from above. Our EM tools are such that this model could in principle be run in a distributed fashion in conjunction with variants of the model that represent the corridor interaction from the perspectives of the agents themselves. This allows the modeller to investigate through experiment how the roles of agents can be played by human agents or automated.

Martin's model embodies a construal of collision avoidance more sophisticated than that documented in Figure 1. The model was developed to explore how human agents manage collision avoidance, and hence involves a richer construal of visual activity, taking account of the idea that it is only possible to look in one direction at once, and that the eye is only sensi-

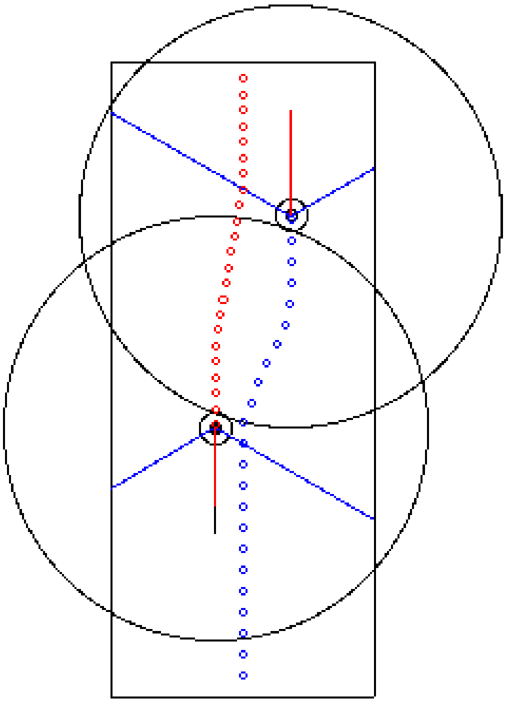


Figure 2: Two agents successfully avoiding a collision in a corridor.

tive within 80 degrees of the direction of looking. Because the modeller's construal is itself to some degree tacit in interaction with the model (cf. Gooding's observation that a construal must be viewed in conjunction with its associated body of ostensive practices (Gooding, 1990)), it is difficult to appreciate Martin's model fully without being able to consult him as the model-builder, or to have a dialogue with him about his interaction with the model. An LSD account is a useful adjunct to the computer model that helps to expose the most prominent meaningful interactions with the model. In practice, there is typically much interesting and significant interaction with and within a model that cannot be explicitly captured in an LSD account. For instance, the collision avoidance strategies used in the most advanced variants of Martin's model were never explicitly described by an LSD account, and involve spatial and temporal considerations that are too subtle to be conveniently specified in an abstract protocol in isolation from the model.

The above discussion illuminates the context for the development of EM artefacts and LSD accounts in HRI. Model construction and the elaboration of LSD accounts are symbiotic processes that do not follow a preconceived pattern, but are mutually supportive. Models and accounts can relate to many different perspectives on agency and modes of observation and construal. Artefact and documentation develop together, and serve complementary purposes both private to the modeller and in relation to the communication of construals.

The first objective in applying EM to HRI would be to better understand how human capabilities and behaviours and robot capabilities and behaviours can be most effectively concurrently elaborated and integrated. As has been illustrated, EM can help us to explore the factors that are significant in determining human behaviour in relation to such tasks as collision avoidance. It can also enable us to construct idealised prototype behaviours that are expressed in terms of high-level abstract observables that serve as useful reference models for devising and analysing robot behaviour. A more ambitious goal involves demonstrating that EM can be used in programming robots. A key aspect of this might involve implementing the `SimpleAvoidingAgent` model with reference to a more primitive and explicit account of the vision capability of an actual robot, through progressively elaborating its states, oracles, handles and protocol. It is in this connection that the usefulness of models and accounts that are intimately related and synchronised is most evident.

It is through developing and experimenting with models based on such construals that the modeller will be able to recognise and address more subtle features of problems of HRI. For instance, by playing out the role of a robot agent in collision avoidance, the modeller will be able to highlight the impact of spatial, temporal and status relations in the interaction. If the person walking towards you is elderly or infirm then it is appropriate to move out of their way so that they are inconvenienced as little as possible. If time is critical (as when there is a fire in the building) then observing social distances will be less of a priority than getting to the fire exit as quickly as possible. Our prior experience suggests that, provided our underlying construals of the more prosaic aspects of avoidance behaviour have been developed with due regard for EM principles and concepts, it will be possible to adapt models to reflect more sophisticated behavioural issues in social interaction. A key factor in this is the well-conceived application of modelling with dependency.

3.2 Modelling using dependency

Dependency is one of the main concepts underlying model-building using EM. Dependencies reflect relationships between characteristics and perceptions of objects that are always maintained. Dependency arises commonly in mechanical systems, where a change to one component directly affects another component in a predictable and indivisible manner. There is no context in which the state of one component is not synchronised with that of a related component.

Dependency maintenance is one of the central characteristics of the software tools that we have developed for EM. Our primary modelling tool supplies notations within which scripts of definitions can be formulated to express dependencies between the many different kinds of observables that determine the various aspects of the state of an EM artefact (see, for instance, the discussion of modelling situated, explicit, mental and internal aspects of state in (Beynon et al., 2001)). The simple illustrative example used in this section makes use of elements from one such model, originally developed by the third author in her final year project in modelling an intelligent house environment. An important feature of EM, to be elaborated in the next section, is the scope it offers for models to be re-used for different purposes, and for relatively complex models to be built up incrementally through assembling and combining simpler components.

Dependency plays a key role in all forms of human-robot interaction. With reference to each agent, there is a dependency between what is observed and what is inferred. With reference to an agent in its environment, there is a dependency between what exists and what is observed. In EM, models of environments are built up from observables and dependency. In modelling a house, for instance, the position of a lamp on a table is dependent on the position of the table: if a person moves the table then the lamp also moves, but not *vice versa*. The illumination of the room is dependent on the position of the lamp and also the position of other objects in the room. If a person or robot is obstructing the lamp then it will affect the illumination of the room with potentially undesirable effects. A socially sensitive robot will need to take account of these dependencies.

By way of illustration, consider the dependency involved in a living room, where there are likely to be people watching television. Clearly, it would be undesirable for a robot to obstruct someone while they are watching television. As in modelling a potential collision in the corridor (cf. the derivate in Figure 1), we can represent a potential obstruction by devising a system of dependencies. If we work with a 2D model of the living room such as is depicted in Figure 3 then we can identify certain areas of the room where the presence of a robot agent will cause an obstruction. Using dependency, these areas can be defined in terms of the position of other agents and the television, so that they change dynamically as agents move around. Other issues might also effect whether there is an obstruction. If the television is switched off then the robot can be fairly sure that it is not being watched. The obstruction is then dependent on: the robot being inside the area between the people and the television, and the television being switched on.

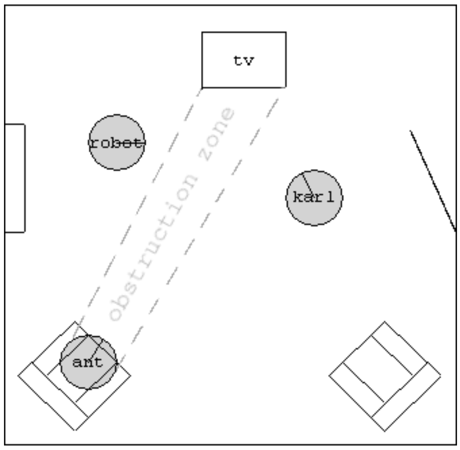


Figure 3: The living room model. Whether the robot causes an obstruction is dependent on the position of the people and the television.

The way in which these dependencies can be directly modelled using EM models is further illustrated in Figure 4, which comprises some key definitions drawn from the underlying model depicted in Figure 3.

When model building with dependency, we can explore the effects of altering observables which may have meaning in the environment. For instance, different people might have different sensitivities about how much space is unoccupied in the visual field around a television. This would mean that the possible obstruction areas would differ according to who was watching. The dependency in the model would make it possible to adapt the model without making any changes to our models of the living room environment or the robot.

The use of dependency in EM is much more significant than the mere introduction of an additional programming construct to a traditional programming language. Appropriately used, dependency serves to support the development of models that stand in a very different relation to interaction in the external

```
tv_is_on = true
tv_position = {16,28}
tv_left = tv_position - {3,0}
tv_right = tv_position + {3,0}

chair1_position = {5,5}
chair1_occupied = true

chair2_position = {27,5}
chair2_occupied = false

robot_position = {10,20}

robot_is_obstructing_chair1 =
        tv_is_on &&
        chair1_occupied &&
        insidetriangle( robot_position,
        chair1_position, tv_left, tv_right)

robot_is_obstructing =
        robot_is_obstructing_chair1 ||
        robot_is_obstruction_chair2
```

Figure 4: An extract of definitive script showing that an obstruction is dependent on the positions and status of other agents in the model.

world from traditional programs. The notion of ‘construal’ is categorically different in character from the idea of a program that is based on a functional specification and optimised to meet its specific functional objectives. This has significant implications for the way in which EM artefacts can be developed and combined.

3.3 Evolving the model

In conventional software development methods it is common for a specification to be formalised before any design or implementation has begun. EM in contrast is of its essence concerned with development that is incremental, at least where the modeller's conception of the artefact is concerned. That is to say, even if the modeller incorporates some pre-existing EM artefact in the development, as has been quite common in practice, the comprehensive understanding of the artefact that may be required to exploit it fully in the new context normally involves a corroborative process of interaction with the artefact and the external world that is similar in nature if not in scale to the interactions involved in its original construction. This corroborative activity is not an all-or-nothing exercise in comprehension such as is typically demanded of the programmer confronted with a formal specification for a requirement or a design, but an active process of becoming familiar through changing the state of the EM artefact in an experimental fashion. This is because a construal only serves its function subject to the engagement of the human interpreter, whether or not the interpreter was also responsible for its development.

In building an EM artefact from scratch, the model-builder takes experimental interaction a step further than simply experiment-for-confirmation. The model-building is exploratory: it is an exploration in the creation of a model, where there is a place for blind experiment-for-discovery. The model-building can begin with little knowledge of what a final model might embody. It is the job of the modeller to develop understanding through exploration of the model; at all times acquiring knowledge and insight in constant connection with the model. This activity of model-building establishes an intimate relation between the artefact itself and the mental model of the modeller, as expressed in terms of expectations about possible states of the artefact, and reliable patterns of interaction with it.

The EM environment goes some way to providing the exploratory power needed to bring the model into close alignment with the modeller's construal of a situation. The interactive nature of the environment enables the modeller to incrementally build artefacts and observe their effects on-the-fly. Some characteristic features of EM can be described with reference to illustrative examples.

Consider a possible development of the living room environment discussed previously. Suppose that we introduce more agents, including one intending to move from one side of the living room to the other – perhaps to reach the cocktail cabinet on the far side of the room. The agent will have to observe the avoidance zones in the living room by exploiting dependency, and also avoid oncoming agents that may be moving across the room. One way of building a model to represent this situation is to combine the living room model and the corridor model, and explore the effects of this conjunction. The result of combining two small models with relatively simple actions is a model with a more complex behaviour.

By evolving a model in this way, incrementally building new artefacts and combining them with existing artefacts, is becomes possible to observe new phenomena and gain insight into more complex behaviours.

The use of dependencies also enables other forms of direct extension of models. Since the EM environment provides a notation for 3D graphics, the modeller might consider extending the 2D model into a 3D model of the living room. This involves writing dependencies to link the positions of objects in the 3D model to their point locations in the 2D model. This kind of model extension can be developed on-the-fly in a exploratory manner.

It is important to note that EM models never reach a "final" state where the implementation is complete: they are continually refined and exercised in new ways through interaction on the part of many different agents (e.g. in the role of designer, observer, or interaction partner). That modelling social interaction should have this open-ended nature is completely plausible. As we do not fully understand the nature of social conventions (Gilbert, 1995) – even our own – it is unlikely that we will ever want to finalise (or completely formalise) a behavioural model.

It is natural for readers unfamiliar with EM thinking and practice to question whether our discussion of applying EM principles to HRI engages sufficiently with the hard problems that are the primary focus of the call for papers. The modest content and conservative themes that are represented in our illustrative examples may suggest a lack of ambition that is out of keeping with our pretensions to an 'alternative model-building' approach. Whilst it is true that our research on applying EM principles to HRI is as yet in its earliest stages, and that far more investment is required to evaluate its true potential, we are optimistic about the prospects of fruitful results in the long term. The same cultural influences that associate computation so strongly with specification and optimisation also often lead us to think of difficulty primarily in terms of problems that can be explicitly framed and whose solution we hope to address by dedicated directed effort that is informed – and in some respects limited – by specific goals. In this way, we come to attach great value to targeted specific techniques and solutions that take us beyond the commonplace territory of a problem domain, whether or not they can be integrated with other solutions of a similar nature, or usefully related to the more mundane regions of the problem space. This is not a concept of difficulty that is well-suited to interpreting our aspirations for EM.

To put the ambitions of EM in perspective, it is useful to contrast *having powerful algorithms to solve specific technical problems in a domain*, and *having a powerful construal of the key phenomena in a domain*. Gaining the latter is invariably a matter of acquiring a large body of experience – even when this experience is guided (as in an established science) by an advanced and comprehensive theory. Since EM is primarily concerned with using the computer to support the development of construals, rather than to implement sophisticated algorithms, it is unsurprising that EM has found broad application to many fields, but has yet to contribute conspicuous specific applications to any one. Similar considerations apply at a different level of abstraction when considering the relationship between ingenious solutions to specific problems in HRI and ways of thinking about the domain that can promote a general understanding and an integration of what may appear to be separate concerns.

The above discussion informs our orientation towards applying EM to problem-solving in HRI. Hard problems often come into being because our solutions to the easier problems are too tightly constrained. This is frequently the result of making the simplifications in our models that are necessary to generate solutions that are sufficiently efficient in execution or ingenious in conception to attract attention. Addressing social interaction will inevitably involve a complex model of activity. Exploratory model-building is a means by which we can start our model-building on a small scale and incrementally extend the model to ever increasing complexity. In this context, the challenge is to integrate the solutions to relatively easy problems without losing conceptual control. This is intimately connected with what this paper highlights as one of the most significant issues in modelling for HRI: supporting the exploratory activity needed to identify problems and learn about their nature, interrelationship and relative difficulty.

Acknowledgements

We are much indebted to Kerstin Dautenhahn and Chrystopher Nehaniv for introducing us to their research on socially interactive robots, and to Steve Russ and Chris Martin for their work on applying EM to crowd behaviour.

References

B. Adams, C. Breazeal, R. Brooks, and B. Scassellati. Humanoid robots: A new kind of tool. *IEEE Intelligent Systems*, 15(4):25–31, 2000.

W.M. Beynon. Empirical Modelling and the foundations of AI. *Computation for Metaphors, Analogy and Agents, Lecture Notes in Artificial Intelligence*, 1562:322–364, 1999.

W.M. Beynon. Radical Empiricism, Empirical Modelling and the nature of knowing. *In Proceedings of the WM 2003 Workshop on Knowledge Management and Philosophy, Luzern, April 3-4*, 2003.

W.M. Beynon, M.T. Norris, R.A. Orr, and M.D. Slade. Definitive specification of concurrent systems. *Proc UKIT'90 IEE Conference Publications*, 316:52–57, 1990.

W.M. Beynon, C. Roe, A.T. Ward, and K.T.A. Wong. Interactive situation models for cognitive aspects of user-artefact interaction. *In Proceedings of Cognitive Technology: Instruments of Mind (CT2001)*, pages 356–372, 2001.

R. Brooks. Intelligence without representation. *Artificial Intelligence*, 47:139–159, 1991.

K. Dautenhahn. Getting to know each other—artificial social intelligence for autonomous robots. *Robotics and Autonomous Systems*, 16:333–356, 1995.

K. Dautenhahn. The art of designing socially intelligent agents: science, fiction and the human in the loop. *Applied Artificial Intelligence Journal*, 12(7-8):573–617, 1998.

K. Dautenhahn. Robots we like to live with?! - a developmental perspective on a personalized, life-long robot companion. *in Proc IEEE Ro-man 2004, Kurashiki, Okayama, Japan (invited paper)*, 2004.

T. Fong, I. Nourbakhsh, and K. Dautenhahn. A survey of socially interactive robots. *Robotics and Autonomous Systems*, 42, 2002.

N. Gilbert. Emergence in social simulation. *in N. Gilbert, and R. Conte (eds), Artificial Societies, London: UCL Press*, pages 144–156, 1995.

D. Gooding. *Experiment and the Making of Meaning*. Kluwer, 1990.

E.T. Hall. *The Hidden Dimension*. Doubleday, 1966.

J. Lind. Issues in agent-oriented software engineering. In *Agent-Oriented Software Engineering*, pages 45–58, 2000.

M. Wooldridge and N.R. Jennings. Intelligent agents: Theory and practice. *Knowledge Engineering Review*, 10(2):115–152, 1994.

Challenges in designing the body and the mind of an interactive robot

Aude Billard
Autonomous Systems Laboratory, School of Engineering
EPFL, Swiss Federal Institute of Technology,
Lausanne CH 1015, Switzerland
`Aude.Billard@epfl.ch`

Abstract

In this talk, I will discuss two key challenges we face when designing both the body and the mind of the interactive humanoid robots, giving as example the Robota project. I will first spend some time on the design issues of the body, stressing the importance of the aesthetic of the robot. I will then discuss the importance of keeping a balanced interaction between the body and the mind of the robot. Below is a brief apercu of the key ideas I will stress.

Body and brain must match:
The term humanoid is, usually, associated to the "human-like" physical appearance of the robot, rather than to its human-like capabilities. It is, however, fundamental that the robot's cognitive capabilities match its physical appearance. For instance, if one is to interact with a robot showing a physical appearance close to that of a human adult, i.e. matching the body proportions and features of an adult, then, one will expect the robot to produce adult-like capabilities, such as, for instance, an understanding of speech and complex manipulation capabilities. Conversely, if one interacts with a baby-like robot, one will probably have lower expectations on the robot's speech and manipulation capabilities.

The Aesthetic of the Body:
There are several key issues in the design of interactive robots, such as, for instance, recognizing human faces, interpreting gestures or keeping interpersonal space. While other speakers in this symposium will discuss these issues, I will, here, stress an often-neglected issue; namely, that concerned with the aesthetic of the robot.
It is a truism that people will be more inclined to interact with "attractive" faces than with "unattractive" ones. Monsters faces, such as those displayed at Halloween, aim at discouraging interactions by frightening people. Dolls, in contrast, are designed to display cuteness and appealing features. Typical appealing features are large eyes, symmetric and round faces, pink cheeks and big eyelashes. Surprisingly, however, many of the humanoid robots developed so far have more in common with monsters than with dolls. It is highly likely that this has a negative effect on the acceptability of humanoid robots in the European society; a society already little inclined to accept robots in its everyday life.

Design Issues in Building Robota:
For the past 8 years, my group has been involved in the design of cute mini-humanoid robots, the Robota robots. Each Robota robot is a mini-humanoid robot, 60cm tall, whose face is that of a commercial doll. Over the years, we have provided Robota with more and more capabilities.
Crucial constraints when designing Robota's body are: cuteness, human-likeness, i.e. respecting the body proportion of a young child (between 16 and 20 months old), and naturalness of the motions, i.e. the robot's motions should be human-like (hence the attachment of the joint should be close to that of the human ones and the kinematics of motion must produce the major characteristics of the human motion).

Robota is provided with 7 degrees of freedom (DOF) articulated arm, including a gripper that allows it to manipulate objects using either a power grasp or a thumb index pinching. It has a 3 DOF neck and 3 DOFs pair of mobile eyes provided with 2 color cameras. A 3-joint spinal cord directs its torso.
Consequently, when designing Robota's brain, we ensure that it is provided with capabilities for interactions that a child of this age would display. These are the ability to recognize human faces and direct its gaze towards the user, the ability to understand and learn a restricted vocabulary and the ability for simple imitation of the user's motion. Note that Robota lacks locomotion capabilities. However, developmental studies do not show that these are necessary for the development of the child's major cognitive capabilities.

At the end of the talk, I will show a number of applications of our control algorithms for human-robot interaction, such as gesture recognition and imitation learning, applied to other humanoid robots than Robota, and, in particular, with the Fujitsu HOAP-2 robot.

Effective Spoken interfaces to service robots: Open problems.

Guido Bugmann
School of Computing, Communications & Electronics
University of Plymouth, Plymouth PL4 8AA, United Kingdom
gbugmann@plymouth.ac.uk

Abstract

This paper discusses some of the open problems in spoken man-machine interfaces that were highlighted during the development of a human-robot interaction (HRI) system enabling humans to give route instruction to a robot. i) Naïve users only know how to explain tasks to other humans, using task decomposition consistent with human execution capabilities. Robots able to understand such instructions need similar (high-level) execution capabilities. Consequently, the current lack of knowledge in some areas of artificial perception, motor control, etc. is a limiting factor in HRI and in the development of service robot applications. ii) Human language is full of inaccuracies and errors, yet communication is effective because of the use of error-repair strategies. Future HRI systems may need human-like repair mechanisms. iii) At the sensory level, the inability to deal with noisy environments limits the range of possible applications. It is suggested that the analysis, not of the user's needs, but of the user's ways of expressing these needs should drive research in robotics and HRI.

1 Introduction

This paper discusses a number of hard (yet unsolved) problems encountered during the development of a NL interface for the instruction of service robots. The development of such interfaces is based on following working assumptions:

1. A service robot cannot be pre-programmed by the manufacturer for all possible tasks and users will need to give some form of "instruction" to their new robot. E.g. specifying which pieces of furniture can be moved during cleaning, and which ones should not be moved, or how to prepare a given variety of tea.
2. To give instruction is in general a multimodal process including verbal description of rules, pointing movements to define objects and demonstration of complex movement sequences. Such process, although akin to programming, is not tractable with conventional programming tools.
3. Human instructors are familiar with methods for instructing other humans, but are unskilled in the art of robot programming. Few have the ability or inclination to learn formal programming languages.

It appears therefore that service robots need to be programmed in a novel way, by users who only know how to explain a task to another human. A solution to that problem is to give robots the ability to understand human-to-human instructions.

Such a solution was explored in a project on Instruction-Based Learning (IBL) which focused on verbal instructions in a direction-giving task in a miniature town. In one way, this project achieved its objectives in that it demonstrated an effective method for generating robust robot programs from spoken instructions (Bugmann et al., 2004). For that purpose, a corpus-based method was developed for building into the linguistic and functional domain of competence of the robot all expressions and action concepts natural to unskilled users, through the analysis of a corpus of utterances representative of the domain of application. However, the work also revealed a number of hard problems that need to be solved before effective commercial robot instruction systems can be developed. Interestingly, these are never pure robotics or natural language processing problems, but involve both areas to various degrees. Following sections detail these problems.

2 Spoken interfaces constrain the robot's design

In the area of computer software development, it is a recognized practice to specify the user interface early in the design process and then to design the software around the interface. In robotics, this is a new concept, as spoken interfaces were very much seen as the last component to be added to a robot. This traditional approach then automatically requires the user to learn the specific language and keywords prepared by the robot designer. However, if one expects the robot to understand unconstrained spoken language, then the question of interface needs to be considered prior to robot design.

To illustrate this, let us assume that a user of a domestic robot-cook needs to give an instruction involving the expression "a pinch of salt". This will clearly exert constraints on how the robot's manipulators are to be designed. Similarly, if a mobile robot needs to understand the command "turn right at the blue sign", it will need to be provided with colour vision.

Figure 1. A subject instructing the robot during corpus collection. Inset: remote-brained mini-robot.

In the IBL project, work started by collecting a corpus of route instructions from subjects explaining to human how to drive a robot in a miniature town towards a destination (fig 1). Their analysis revealed 13 primitives functions, some which where navigation procedures such as "take the n^{th} turn right/left" some where just informative statements such as "you pass the post-office to your left" (for more details see Bugmann et al., 2004). Only after this analysis did work start on designing the vision and control system, to build all robot functions required by HRI (Kyriacou et al., 2005).

Note that a command such as "turn right" is highly under-specified for a robot, with no details on what actuators must do. Hence service robots must gather missing information from the environment and make autonomous decisions, e.g. recognize the layout and plan a trajectory. To understand natural language, a robot needs a high level of functionality. In fact, utterance like "clean this window" or "hang up the washing" make demands on robot design and control that are beyond current knowledge. Given that these are expressions that future users are likely to use, it is of concern that relatively little research is devoted to the corresponding robot skills.

There are also examples where particularities of human language, e.g. references to combinations of procedures, exert more subtle constraints on various aspects of robot design, such as its computational architecture (see e.g. next section).

3 Spatial-language-specific problems.

Hereafter are examples of difficulties that natural language in the domain of route instruction creates in both the NLP and robotics domains.

D*etecting references to previous routes.* During corpus collection, subjects were encouraged to refer to previously taught routes whenever possible, rather than re-describing every step of a route. It turned out that such references are very difficult to detect in instructions. In one third of the cases, subjects referred to previous route implicitly, e.g. via a landmark that was part of a previous route. For instance, when a subject said "go to the roundabout", it was impossible to tell if this referred to a roundabout that is just in front of the robot or a roundabout further away that can be reached using parts of a route previously instructed. In two third of the cases, the destination of a previous route was explicitly mentioned "start as if you were going to the post-office" but in half of these cases, the sentences had structures that could not be properly translated by our NLP system.

Interestingly, experiments with human subjects listening to the same instructions showed that only 55% of references to previous routes were detected in the instruction. Only when subjects started to drive the robot (by remote control) did they notice that there was a problem.

Using references to previous routes when creating program codes. Almost all references to previous routes required only a partial use of the instruction sequence: e.g. "take the route to the station, but after the bridge turn left". One of the problems is that the bridge may not even be mentioned in the instruction of the route to the station. No definite solution has been found to that problem. One proposal was to implement a multi-threaded concurrent processing scheme where the robot would "follow the road to the station" and at the same time "try to find the left turn after the bridge". The second process would remain the sole active as soon as the turn is found (Lauria et al., 2002). It remains to be seen if this solution is general enough, but it is interesting to note that the way users express themselves could end up dictating the computational architecture of the robot controller.

Programming the final instruction. The final instruction of a route instruction is often a statement like "and you will see it there on your left". The final instruction is especially interesting as it is the one requiring the most autonomy from the robot. It is highly under-specified and the robot needs to visually locate the destination and then plan a path towards it. In our miniature town, we have not undertaken the difficult task of detecting the building, identifying it from its sign and locating its entrance. Instead, a coloured strip was placed at the foot of the building to signal its position. In a real urban environment the final instruction would pose vision and control challenges that are at the limits of current technical capabilities.

4 Handling misunderstandings

Robots are designed with a limited vocabulary corresponding to their action capabilities. In principle this simplifies the design of NLP components and improves the performance of speech recognition. However, users do not know the limits of the domain of competence of the robot and often address the robot with utterances that it cannot understand.

The standard approach to solving this problem is increasing the grammar, e.g. by collecting a larger corpus of utterances natural to users in that domain, then tuning the grammar to that corpus. However, this approach improves the grammar only modestly for a large effort in corpus collection (Bugmann et al., 2001). Another approach is to adding to the grammar a sample of potential out-of corpus expressions (Hockey et al., 2003). However, no matter how large the coverage of the grammar, a robot always has a limited domain of linguistic and functional competence. When the user steps out of this domain, communication brakes down.

Another approach is to accept the domain limitation and work with it. Somehow, the robot should be able to help the user naturally discover its domain of competence. An impractical alternative would be to ask the user to undergo long and detailed training sessions on the robot's capabilities. Both approaches can also be seen as two stages of dialogue system development (fig. 2)

A dialogue system that informs the user about the robot's competences is not straightforward to design. First, it requires that the out-of-domain error

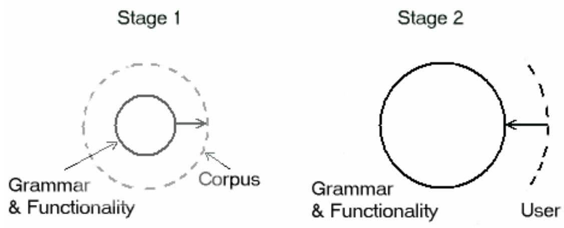


Figure 2. Two stages of spoken user-interface development. In stage one, the system is adapted to the user represented by the corpus. In stage 2, the user is informed of the capabilities of the robot.

is detected. Unfortunately, speech recognition systems are not good at detecting that their interpretation of the speech sounds is incorrect. They tend to generate a translation that is consistent with the grammar of the domain of competence. For instance, if the user asks the robot "go to Tescos", and the word "Tesco" is unknown to the system, the translation may be "go to the school", a perfectly legal request. It is possible that the user detects the error at some point in the dialogue and attempts to engage in a correction dialogue with the robot. However, research on such dialogues is still in its infancy. If the system detects the error, e.g. through a low speech recognition confidence score, how will it inform the user that it does not know the word "Tesco" if it is not in its vocabulary? Some errors may not be speech recognition errors, but requests incompatible with the robot's capabilities. In general, to generate a helpful message, the speech generation sub-system must be aware of what the robot can and cannot do. However, this is a manufacturer-specific knowledge. How should it be represented? In conversations between humans, all these comprehension problems also occur but, after a few clarifications, speakers usually manage to align their utterance to the domain of common ground (Garrod and Pickering, 2004), or learn new concepts if necessary. This is an area where findings from life sciences could help develop more effective human-robot dialogue systems. We are currently planning work in this area.

5 Speech recognition in noisy backgrounds

Speech recognition has made significant progress in recent years, as evidenced by a number of effective commercial software packages, and does not constitute anymore the bottleneck in natural language interfaces. However, this has given other problems more prominence. In the IBL project, we used a microphone placed near the mouth and switched it on only for the duration of the speech. This enabled effective speech recognition even in a noisy background such as an exhibition. However, if the microphone were always on, the system would start interpreting the background noise. A possible solution could be to establish a directional window using an array of microphones (e.g. eight microphones used on the JiJO-2 office robot (Asoh et al. 2001) or two "ears" used on the SIG active head by Nakadai et al. (2003)). How much of the problem is solved by such systems remains to be seen. Biological systems are also able to track an individual voice from its features, and ultimately hold the solution to noisy speech recognition. Until then, speech-enabled devices will require the user to wear a microphone. In practice, this eliminates all applications where an unknown user addresses a machine in a noisy environment.

6 Multimodal integration

This section is a brief reminder that verbal communication alone is insufficient for HRI. Natural language is a powerful tool for expression rules and sequences of operations. However, it is less expressive for shapes, locations and movements. Natural spoken communication is usually supported by gestures such as pointing to an object or a direction. Many tasks cannot be explained and are best demonstrated. This has long been recognized and research in speech interfaces must be considered as a part of the wider area of multi-modal communication. Some good examples are the GRAVIS system developed in Bielefeld (Steil et al., 2004), and systems developed by Imai et al. (1999) and Ono et al., (2001).

Given the functional consequences of accepting unconstrained spoken input (noted above), it may be interesting to investigate a corpus-based approach to unconstrained multimodal input. This should be done in the context of the instruction of tasks relevant for future users. It is possible that new aspects of verbal communication and its interaction with other forms of communication would then be highlighted.

7 Conclusion

For a robot to understand everyday language, it also needs to be able to execute tasks referred to in everyday language. At present, the problem of designing smart sensory-motor functions is much more difficult than speech recognition. How to recognize a dirty window, a wet piece of cloth? Realizing such difficult tasks could benefit from biological inspiration, especially in the area of vision.

Dialogues are full of misunderstandings and the ability to overcome these makes human-human communication so effective. In this respect, human-robot communication is very poor. A large number of problems remain to be solved, such as error detection, error repair, learning new words and actions, informative dialogues, etc. Such research is very much guided by findings and methods in psychology.

The human auditory system shows capabilities of filtering out background noise and can adapt to the speakers pitch and accent. Speech recognition systems do not process effectively voices in noisy environments or with unusual characteristics. Here, findings in the area of neuroscience of sensory systems could accelerate the solution of these problems.

Overall, speech interfaces require a high level of functional competence from the robot, as humans refer to high-level functions in their everyday language. What these functions should be is still speculative for most applications. The handling of misunderstandings requires from robots a high level of cognitive competence, mimicking many characteristics of human listeners.

Acknowledgements

The author is grateful for comments and references provided by anonymous referees.

References

Asoh H., Motmura Y., Asano F., Hara I., Hayamizu S., Itou K., Kurita T., Matsui T., Vlassis N., Bunschoten R., Krose B. (2001) Jijo-2: An office robot that communicates and learns. *IEEE Intelligent Systems*, 16:5, 46-55.

Bugmann G., Stanislao Lauria, Theocharis Kyriacou, Ewan Klein, Johan Bos, Kenny Coventry (2001) Using Verbal Instructions for Route Learning: Instruction Analysis . *Proc. TIMR 01 – Towards Intelligent Mobile Robots, Manchester 2001. Technical Report Series, Department of Computer Science, Manchester University,* ISSN 1361 – 6161. Report number UMC-01-4-1

Bugmann G., Klein E., Lauria S. and Kyriacou T. (2004) Corpus-Based Robotics: A Route Instruction Example. *Proceedings of IAS-8*, 10-13 March 2004, Amsterdam, pp. 96-103.

Garrod S, Pickering MJ (2004) Why is conversation so easy? Trends in Cognitive Sciences. 8 (1): 8-11.

Hockey B.A., Lemon O., Campana E., Hiatt L., Aist G., Hieronymus J., Gruenstein A. and Dowding J. (2003) Targeted help for spoken dialogue systems: Intelligent feedback improves naïve user's performance. *Proc. 10th Conf. of the European Chapter of the Association for Computational Linguistic* (EACL'03), Budapest, Hungary.

Imai M., Kazuo Hiraki, Tsutomu Miyasato. Physical Constraints on Human Roboto Interaction, *Proceedings of 16th International Joint Conference on Artificial Intelligence* (IJCAI99), PP.1124--1130 (1999).

Kyriacou T., Bugmann G. and E., Lauria S. (2005) Vision-Based Urban Navigation Procedures for Verbally Instructed Robots. To appear in *Robotics and Autonomous Systems*

Lauria S., Kyriacou T. Bugmann G., Bos J and Klein E. (2002) Converting Natural Language Route Instructions into Robot-Executable Procedures. *Proceedings of the 2002 IEEE Int. Workshop on Robot and Human Interactive Communication* (Roman'02), Berlin, Germany, pp. 223-228.

Nakadai K, Hiroshi G. Okuno, Hiroaki Kitano: Robot Recognizes Three Simultaneous Speech by Active Audition. *Proceedings of IEEE-RAS International Conference on Robots and Automation* (ICRA-2003), 398-403, IEEE, Sep. 2003

Ono T., Michita Imai, Hiroshi Ishiguro (2001). A Model of Embodied Communications with Gestures between Humans and Robots. *Proceedings of Twenty-third Annual Meeting of the Cognitive Science Society* (CogSci2001), pp. 732--737.

Steil J.J, Rothling F., Haschke R, and Ritter H. (2004) Situated Robot Learning for Multimodal Instruction and Imitation of Grasping. Robotics and Autonomous Systems, 47, 129-141.

Evaluation Criteria for Human Robot Interaction

Catherina R. Burghart*
*Institute of Process Control and Robotics
University of Karlsruhe
burghart@ira.uka.de

Roger Haeussling†
†Institute of Sociology
University of Karlsruhe
roger.haeussling@ sozio.geist-soz.uni-karlsruhe.de

Abstract

Human robot co-operation is an upcoming topic in robotics combining the characteristics of both human and robot in order to be able to perform co-operative tasks in a human-robot team. Normally, engineers start developing a co-operative robotic system with the aim of creating an intuitive interface for the user to interact with the robot, yet important social parameters for the intuitive interaction between robot and human partner are seldom contemplated. As sociology offers various methods to describe and evaluate the interaction between humans and between humans and machines, we have joined forces to apply these methods to human robot co-operation. Combining both sociological and technical parameters we have created a classification scheme for the analysis of human robot co-operation which is presented in this paper. This classification scheme was then applied in a field study evaluating four different methods to co-operatively carry a wooden bar with a robot; the results are also discussed in this paper.

1 Introduction

Human-robot teams performing co-operative tasks are a new field of application and research in service robotics, surgical robotics and industrial robotics. The specific abilities of both the human partner and the robotic system are combined in order to achieve flexible robotic systems which can be used in unstructured and dynamic environments performing difficult tasks, which neither can perform alone in the same manner. Humans are characterised by their flexibility, great experience, wide knowledge, ability to abstract and ability to recognize situations and react adequately. In contrast robots have a high accuracy, strength, dependability and endurance. The most important aspect of co-operative human robot teams is an intuitive interaction of the human with the robotic partner: this actually determines the quality of the co-operation and the obtained result of a co-operatively performed task. Therefore, psychological and sociological aspects have to be considered when developing co-operative robotic systems in order to recognize and respect the main social parameters of human-robot-interaction.

Sociology observes human robot co-operation as a network of social interactions between human and non-human actors. The co-operation rules and the roles for both actors grow during the progress. The aim to achieve a co-operation as intuitive as possible between the human and the robotic partner requires the recognition and consideration of the main social parameters of a co-operative task between human and robot. From the sociological perspective it is necessary that social interaction rules of co-operating robot systems become binding. However, this is only possible, if a methodical concept for the recording of new ways of co-operation between human being and robot is established.

Therefore, we have devised an extensive classification scheme respecting both technical details and conditions of the robotic system as well as social parameters. Using these a close analysis of several persons performing and testing a co-operative task shows critical points in the co-operation., which can serve as input for a system re-design.

This paper is organised as follows: Section 2 describes the devised classification scheme in detail. The results of a field experiment applying the classification scheme are illustrated in Section 3. Section 4 intensively discusses the results.

2 Classification Scheme

The detailed contemplation of technical and sociological views on human robot co-operation and interaction shows that on the one hand there is still missing a new sociological methodology which respects both human actors and object technologies and their reciprocal effect on each other in all aspects. On the other hand social parameters still are not considered in all facets by engineers working on co-operating robotic systems. But in order to achieve an optimal co-operation between human and robot social parameters have to be respected by the robot control. Therefore, an optimal intuitive interaction requires a new theoretical concept describing both human actors and machines as well as their operational processes, effects on each other and parameters.

We have devised a new theoretical concept which awards a direct social effectiveness to the operational processes of machines. Our theoretical framework comprises a sociological multi level model for comprehensive analysis of interaction sequences based on a network-concept (not only usable for human-robot-co-operation, but also for any "face-to-face" constellation). The following four levels are defined with respect to the sociological and technical view:

- **Level A: Interaction Context** Every interaction is concerned with its context and can only be interpreted correctly through its contents.
- **Level B: Interaction / Co-operation** The co-operation is considered as an interacting network with emerging interaction pathways, dynamics and allocations of positions.
- **Level C: Activity of Actors** The intentions and activities of both actors, robot and human, are essential for the actual performance of an interaction. Intentions (goals, roles, intended actions) can differ up to a great extent from their actual realization within an interaction.
- **Level D: Non-verbal Actions and Emotions** Non-verbal actions (gestures, postures, mimics) and emotions are constantly associated with steering, signalling and evaluating human intention and interaction. They play a fundamental role in the affirmation or declination of concrete interaction frames.

Two topics can be identified using this concept: the typical optimal co-operation under laboratory conditions (nearly the vision of the engineers) and a social assimilation of the robot system based on everyday-life including the re-definition of the system based on worth and functionality and the connection to human actors.

All forms of human-robot-co-operation can now be classified according to the following categories.

2.1 Level of Interaction Context

2.1.1 Interaction patterns

Interactions are influenced by a set of two different conditions: on the one hand necessary conditions like the participating actors, the sensors or robot programs used determine interaction patterns. On the other hand coincidentally existing conditions i.e. audience, sounds, coincidental events influence interaction patterns. Especially humans seek help or signals of an audience watching a co-operatively performed task by human and robot.

2.1.2 Rules of interaction

They are established and stabilized during the co-operation and give the process a dynamic orientation (corridor of interaction)

The general form of rules of interaction can be described as follows: if event A occurs then the chain of interactions B is performed as reaction upon A.

A rule of interaction can on the one hand be defined as an interaction pattern which the human partner always uses as reaction to a specific robot action and which has become routine in a timeline. On the other hand a robot rule of interaction is the selection of the robot interaction pattern according to a recognized and interpreted human behaviour and according to a grammar defining the resulting actions of the robot to take.

2.1.3 Roles of interaction

In order to perform a co-operative task both human and robot can have different roles i.e. who guides and who follows when carrying an object together. Possible roles can be described by a script listing an ideal sequence of chronological interactions and corresponding roles. Roles can change during an interaction; an efficient co-operation requires a specific role assigned to an actor at a given time. Interesting aspects are whether the role of each partner is transparent to the other and whether human and robot are free to choose their roles.

2.1.4 Co-operation manner

Both partners can be coupled in a different manner during a co-operation. This can be the context, i.e. the task to be performed, or an intention of mutual consent (determination of the same goal to be achieved). A coupling can also be achieved on a lower level using haptic or tactile, visual or acoustic sensors.

2.1.5 Degree of freedom of human and robot action

Interaction with a robotic system often requires of the human partner specific knowledge to handle the robot, constraints in motion during the interaction due to the rules established by the robot program and a restricted set of senses applicable to the task. In comparison, the same interaction between two humans including the complete scale of possible ideal human behaviour serves as reference for the co-operatively performed task between human and robot. Vice versa the robot's actions are restricted by the system hard- and software design.

2.2 Level of Interaction / Co-operation

2.2.1 Co-operation level

The abilities of a robotic system depend on the intelligence implemented into the robot control: the ability for complex interaction with the human partner enhances with increasing artificial intelligence of the robotic system. The aim of human robot interaction is an intuitive handling of the robotic system which is based on using the natural senses of the human (visual, acoustic and tactile) and the human ways of thinking and planning. Three different levels can be defined for human-robot-co-operation:

- reactive / sensor-motor level,
- rule based level and
- knowledge based level.

On the reactive level sensor data form the input of a reaction to be performed by the robot, which is initiated by comparing the sensor data to an evaluation function. In contrast, a set of rules stored in a rule base describes different co-operative tasks on the rule based level. Using these rules the robotic system can plan and execute the next step expected of him within a co-operative task. Finally, on the most sophisticated level, the robotic control system plans and performs a task or a subsequent working step by combining sensor data of various integrated sensors and data of a knowledge base storing information about different tasks, characteristics of the human partner and the environment the robot moves in.

2.2.2 Co-operation intensity

The intensity of a co-operation can be described by the number of single interactions observed during a specified time interval. An intense co-operation therefore is defined as a series of dense interactions. If robot or human partner does not know what the other partner is doing, the number of performed interactions decreases.

2.2.3 Congruity of interaction

If the offers of interaction between robotic system and human partner are tailored towards each other and interlocked, a co-operation can be defined as congruent. In all other cases, the goals of the co-operative task cannot be achieved.

2.2.4 Synergy of interaction

An optimal co-operation can be achieved, if the sequences of interactions between robot and human partner develop a dynamic flow throughout the co-operation amplifying each other and being subsequent steps of each other.

2.2.5 Co-operation efficiency

Different criteria are responsible for the efficiency of a co-operation. Here, especially the flow of interactions is contemplated: inconsistencies, complete stops or changes in direction lead to less efficiently performed co-operative tasks. Different criteria of efficiency are: duration, usage of resources (also cognition: reflection, physical compartment), learning needs (are there any learning procedures recognisable) and freedom of redundancies (i.e. does the interaction pass already reached stages).

2.3 Level of Activity of Actors

2.3.1 Goal orientation

Both actors, robot and human, each follow at least one goal during an interaction. In this context goal is considered as the result to be achieved by the interaction, not a goal in a psychological sense (personal intention). A goal of the human partner is the desired result to be obtained by interacting with the robot. In a simple robotic system, a goal can be just to follow a given impulse i.e. to transform a measured sensor signal into a movement. Research and work on more "intelligent" robotic systems intends these robots to see the common result to be obtained by the interaction, just as the humans do. Both actors can also follow different goals or several goals simultaneously.

2.3.2 Transparency of activities

For both actors it is important to know what the other partner in the interaction is doing. A human can only obtain a successful result when interacting with a robot, if all actions of the robot can be correctly interpreted. Vice versa a robotic system can only co-operate successfully with a human partner, if it can distinctly interpret and predict the human partner's actions. Especially, when only a small selection or just one of the possible communication channels between human and robot are implemented, transparency of the actions cannot be achieved by using other means of understanding i.e. if a robot cannot ask a human partner about his or her intention.

2.3.3 Transparency of roles

A successful interaction between robot and human can only be guaranteed if the roles both actors possess are transparent to the other side. I.e. when robot and human are carrying an object together, it has to be well defined who is in charge of determining the direction and orientation of the object (who guides and who follows). If the roles are not transparent, each actor cannot deduce the appropriate actions to be adopted in order to reach the goal of the interaction. Additionally, non-transparent roles can lead to severe safety risks for the human.

2.4 Level of Non-verbal Actions and Emotions

2.4.1 Human senses and robot sensors

Interacting with each other human beings use various senses to communicate and signal. The same applies, if a human partner interacts with a robot, even if the robot is equipped with only one type of sensor system and thus ignores all other signals by the human. For the analysis of an interaction it is important to know which senses are involved in the interaction (i.e. the human partner constantly stares at the robot hand) and to which extent the senses are used.

Vice versa the senses applied by the robot are well defined as the correlate with the hardware and according software methods used in the robotic system.

2.4.2 Non-verbal, emotional signalling (mimic, gesticulation, posture)

Using the sociological methods from conversation analysis (Sacks), from non-verbal communication analysis (Leventhal/Sharp, Exline/ Winters, Birdwhistell, Milgram, Jourard, Hall, Condon/ Ogston) and from interaction analysis (Bales, Borgatta) a flow chart of mimics and positional articulation responses escorting the human actions can be generated. Comparative analysis allows to specify to which extent non verbal activity accompanies the interactions and to which extent complementing explanations reinforce interactions of the human partner. Within this category it has to be examined whether mimic, posture and gesticulation go hand in hand with the progress of interactions (i.e. if a timid trial is complemented by a interrogating look of the actor).

The “normal” robotic system does not use any kind of emotional signalling; a small number of international research groups try to equip their robotic systems with the ability to show emotions (i.e. KISMET, ISAC, Leonardo).

2.4.3 Affects

In this context, affects are „actions committed under the influence of eruptive, not cognitively controlled emotions” i.e. the human partner stubbornly pulls at the endeffector to move the robot arm although the robot control has already been switched off. These affects go beyond the scope of staging a role of action. In order to achieve an intuitive interface between human partner and robot it is important to know in which phase of the interaction process an affective discontinuity takes place and under which circumstances.

3 Field Study

The afore presented criteria were used to evaluate four different methods to co-operatively carry a wooden bar with a robot in a field study. In this section the set-up of the performed field study, the performance itself and the evaluation methods used as well as the results are presented.

3.1 Experimental Set-up

Our experimental system comprises an anthropomorphic robot arm with 7 dof, which is equipped with different rigid and flexible tactile sensor arrays, a force-torque sensor and a gripper. Instead of the gripper an 11 dof anthropomorphic hand can also be used. Additionally, a stereo camera system is mounted onto a neck (pan-tilt unit) which is also attached to the torso (Fig. 1). All in all we have to control a complex system with 20 dof and co-ordinate arm, hand, neck, camera system and haptic sensors.

In order to perform different co-operative tasks a set of different control modes has been generated (Yigit et al., 2003). These control modes are based on various combinations of position control, force control, zero-force-control and contact control. A modified impedance control with additional constraints is used for our methods to co-operatively carry objects.

Four different methods for the co-operative carrying of objects are implemented in the robot control:

1. a method using wheelbarrow-like constraints (Takubo et al., 2000),
2. a method of mapping torques to rotations and then substituting the rotation by a translation and an inverse rotation (Yigit et al., 2003),
3. a simple mapping of torques to translations, and
4. a pumping method, analogous to a manual water pump.

As the two simpler ones of these methods provide no free combination of translating and rotating the carried object, an operator has to change the restrictions in these cases. Thus, if the human partner wants to change the orientation these modes are switched from allowing only translations to allowing only rotations.

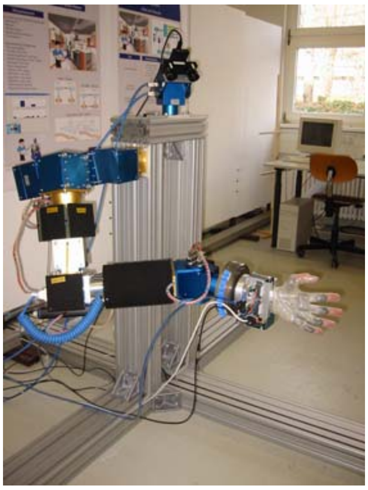


Figure 1. Experimental set-up

3.2 Methods

Eight students with no experience in the handling of robots were chosen to perform the experiment. Four of them had to test all four different methods to co-operatively carry an object; the other four just knew, which control modes were possible and how they worked, but they did not know, which of the methods was actually activated when they to their turn to perform the experiment.

As object to be carried by robot and human partner a wooden bar of 67 cm length was chosen. At one end the bar was rigidly gripped by the robot's gripper and additionally fixed with screws. The human partner gripped the other end of the bar.

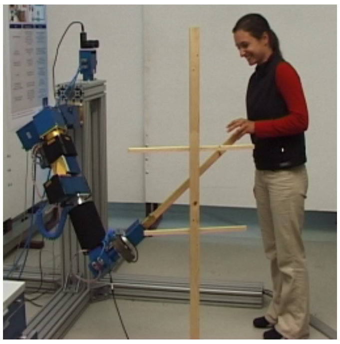


Figure 2. Run in field study

A course was set up which the students had to run through. First the bar had to be lifted up till it was level with the upper bar of an easel used as reference. Then the position of the easel was changed requiring the test person to move the bar sideward on the same level (requesting a change of the orientation). Finally, the easel was repositioned requesting the student to change the position and orientation of the bar to a position that lower and closer to the robot until the bar was level with the lower reference bar of the easel. All reference positions of the easel were marked on the floor thus guaranteeing that all test persons had to achieve the same reference positions.

All experimental runs were recorded by two different digital camera systems using different perspectives. One camera recorded the complete scene from the front (Fig. 2), the second camera just recorded the test persons by monitoring their faces.

After the students had completed their runs, they filled our individual protocols. Later all digital video sequences were analysed using the categories presented in Section II.

The results of each run were marked in tables and time charts as described in the following subsection.

Additionally, a co-operative carrying of a wooden bar between two human partners was performed and recorded as comparison.

3.3 Results

All tested methods to co-operatively carry a wooden bar with a robot are rather simple co-operative tasks based on a physical coupling between the human partner and the robot, as the only sensor system used by the robot is a force-torque sensor. Due to the simplicity of the co-operation the actual co-operation level concerned in the robot control is the reactive level. Although the goal of the human partner is to obtain the required position and orientation of the object, the goal of the robotic system is the mere reaction to a detected impulse (measured forces and torques). The roles of both partners within the co-operation are predefined due to the implemented methods. The human partner guides the robot; the robot follows the given impulse.

In the opinion of the robotic researchers the implementation of such a co-operation might be simple, but in the opinion of the human co-operative partners, the tested methods are rather inadequate. The field study has shown, that the actions of the robot are not very transparent to the inexperienced human user. Only the accurate knowledge and understanding of the actual methods for the co-operative carrying of an object leads to an effective performance of the co-operative task. All students declared that the method copying the human movements is the most intuitive. The pumping method is easier to perform than the wheelbarrow method, but both of them are not intuitive to the inexperienced user. Additionally, the wheelbarrow methods requires the human to use a lot of room in order to guide the robot.

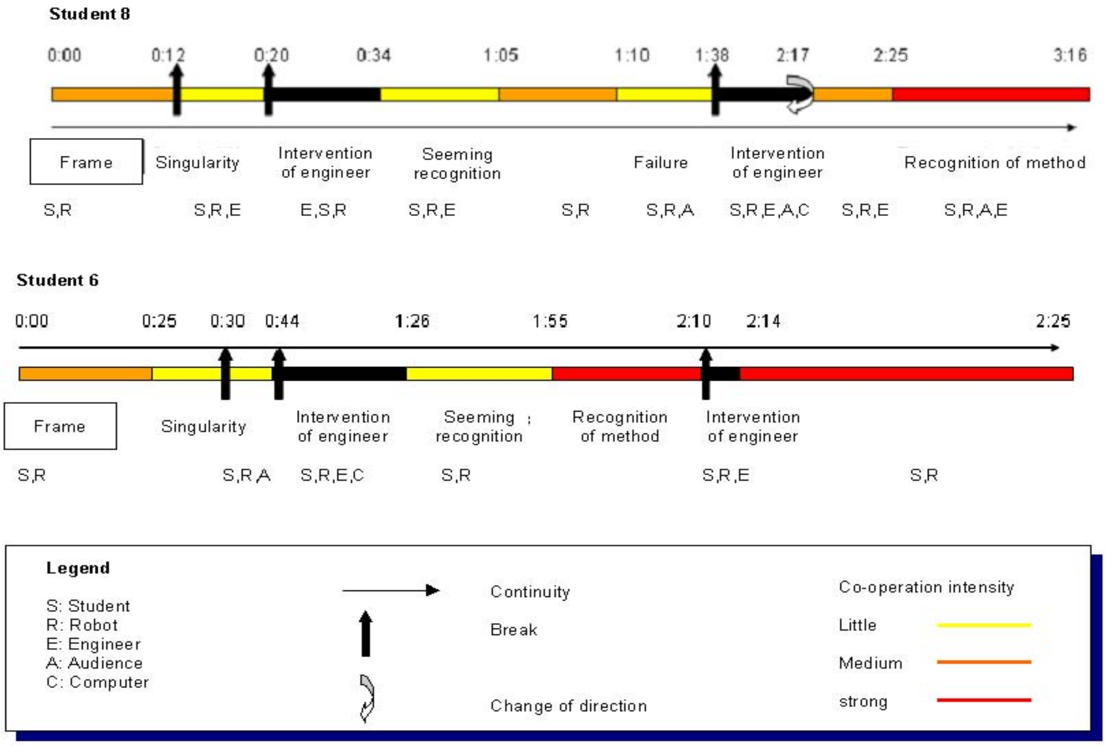


Figure 3. Comparison of experimental runs

The simplicity of the system and the methods used is the great disadvantage at the same time. The students tried to communicate with the robot by their mimic or even verbally. Being desperate they turned to the engineer in command or to the audience. Figure 3 depicts two runs of two students who did not know which methods was used. A detailed description of the analysed parameters can be found in Table 2 and Table 3.

Table 1: Abbreviations used for description of analysis

Abbreviation	Meaning
ae	at ease
c	concentrated
ce	contraction of eyebrows
cf	contraction of face
cm	contraction of mouth
co	communication
le	lifting of eyebrows
lo	letting off object
mm	movement of mouth
mo	mouth open
pl	pressing lips together
qg	questioning gaze
sl	Smiling / laughing
t	tense
-	Little / feable
o	medium
+	strong
/	not constrained
0	constrained
(0)	very constrained
Y	Yes
N	No
?	Not determinable

Both of them tried and watched the reaction of the robotic system, both of them were convinced to recognize the actual methods, although they did not. When the robot arm was guided into a singularity, they both desperately tried to move the robot arm (affect), although the robot control had already switched off and the engineer in command had to intervene. Student 8 took a long time to recognize the method used as the pumping method, till she finally ended in an intensive and effective performance of the co-operation. The interaction process swayed between breaks with no interaction and phases of little or medium interaction intensity until the student finally hit the correct method. Student 6 watched the reaction of the robotic system to the inputs more closely and thus sooner achieved an intense and effective co-operation after recognizing the virtual wheelbarrow method.

Table 2: Analysis of run of Student 8

Time	Goal orientation of actors	Dof of human action	Transparency of activities	Transparency of roles	Signaling (mimic, gestures, posture)	Affects
0:00	+	0	N	Y	c,le,sl,cm,t	N
0:15	-	0	N	N	sl,mm,c,ae	N
0:35	+,o	0	N	N	c,le,pl,sl,mm,mo	Y
0:50	-	(0)	N	N	c,ce,cm	N
1:05	-	(0)	N	N	c,cf,ce,mm	N
1:20	-	0	N	N	c,sl	N
2:15	+	0	N	N	c,le,sl	N
2:30	+	0	Y	Y	c,ae	N
2:45	+	0	Y	Y	c,pl,sl	N
3:00	+	0	y	y	c,sl,ae	n

In order to contrast the field study with the interaction of two humans carrying a bar another experiment was performed and analysed. In this case, a person gave two test persons different tasks to be performed: both persons were to grip one end of a wooden bar and carry it together with the other person, but the two test persons had different, contradictory goals. The bar was to be lifted over a small pyramid of chairs. For each test person the intended goal position of the bar was perpendicular to the intended goal position of the other test person. At the same time the test persons were not aware of the fact.

The results of the analysis are depicted in Table 4 and Figure 4. The analysis shows, that in this case the interaction at once becomes very intense and all communication channels are used to come to a fast understanding.

4 Discussion

The classification pattern allows us to register the complete social multidimensionality of the course of interactions, thus making it possible to identify the interrelations between events on different levels of interaction and to formulate indicators for typical interaction sequences.

Thus interaction-patterns, which have a specific portfolio of intensities of categories in all 4 levels, can be interpreted further on the time axis.

Table 3: Analysis of run of Student 6

Time	Goal orientation of actors	Dof of human action	Transparency of activities	Transparency of roles	Signaling (mimic, gestures, posture)	Affects
0:00	+	0	N	Y	c	N
0:15	o	0	N	N	c,ae	N
0:30	-	0	N	N	c,sl	Y
0:45	-	(0)	N	N	c	N
1:00	-	(0)	N	N	c,t	N
1:25	o	0	N	N	c,t	N
1:40	o	0	N	N	c,mm,lo	N
1:55	+	0	Y	Y	c,ae	N
2:10	+	0	Y	Y	qg,c,ae	N

Table 4: Analysis of human-human-interaction

Time	Test person	Goal orientation of actor	Dof of human action	Transparency of activities	Transparency of roles	Signaling (mimic, ges-tures, pos-ture)	Affects
0:43-0:48	1	o	0	N	N	sl,qg,ae	N
	2	o	0	N	N	sl,qg,ae	N
0:48-0:53	1	+	0	N	Y	sl,qg,ae	N
	2	+	0	N	Y	sl,qg,ae	N
0:53-0:58	1	+	0	N	Y	sl,qg,ae	N
	2	+	0	N	Y	sl,qg,ae	N
0:58-1:03	1	+	0	N	Y	sl,qg,ae	N
	2	+	0	N	Y	sl,qg,ae,co	N

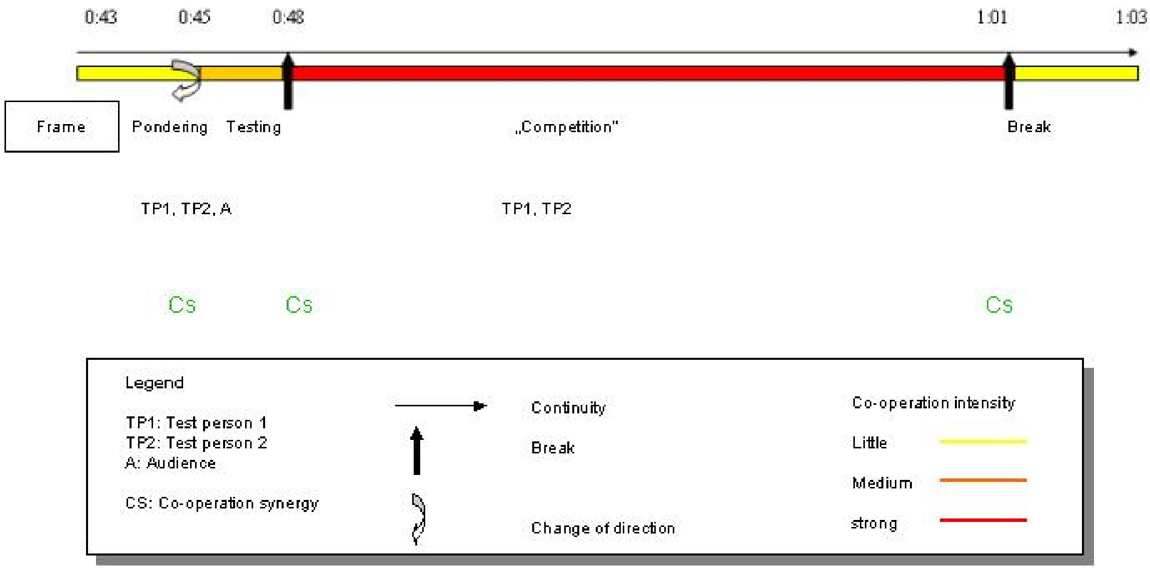


Figure 4. Results of human-human-interaction

Meanwhile some of the types of the categories act in a "synchronic" way to each other, which could be seen as a correlation. According to this, specific degrees of cooperation intensity and congruity of interaction could be given to the identified roles of interaction. The determined orientation itself has a strong effect on the intensity of co-operation, gesture and facial plays of the test persons.

Even specific non-verbal signals, such as a frown, lifting an eyebrow, pursed lips, a nervous smile or a thoughtful look, always occur in comparable instances throughout the course.

Therefore, a similar "dramaturgy" of human behaviour was recognized for all test persons producing a singularity. This "dramaturgy" of human behaviour can be divided into the following stages:

1. Stage of testing: This stage is quite short and allows the test person to interpret the procedure used by the robot. The test person is extremely concentrated (with corresponding facial play and gestures).
2. Stage of stabilization: The human part maintains the interpretation made in a previous stage (stage 1) without any irritation and starts the application of the corresponding behaviour. The test-person shows confidence, sovereignty and a relaxed posture.
3. Stage of irritation: After a longer period of time during which the desired results are not achieved, the test person finally becomes irritated and - as a result - acts aimlessly. Quick changes of gestures and facial play can be observed.
4. Stage of contact to the outside world: The test person now tries to contact others (spectators, experiment leader) in order to gather the necessary information to solve the problem (also the results of the well known workplace-studies). At the same time, his concentration is reduced to mere cooperation with the robot. During this stage, the test person uses non-verbal means such as attempting to establish an eye-contact or twisting the body in order to get in touch with the outside world.
5. Stage of adherence: As neither the robot, nor the spectators have delivered any information about potential misinterpretation, the test person, despite his irritation, maintains his first interpretation and provokes a state of singularity. Accordingly, affective reactions to the occurrence of the

singularity can be observed, despite previous indications.

When human-human-interaction is concerned, non-verbal signals not only have the function of clarifying verbal statements but also help to ensure that no permanent tasks of correction have to be carried out during focal interaction. Correspondingly, the test-series of failed cooperation can be interpreted as follows:

Cooperation during which the human actors misinterpret the procedure in current use at the beginning, inevitably results in a state of singularity, as the fourth dimension of interaction (non-verbal signals) is missing in human-robot relation.

On one hand, the robot does not deliver any signals which show the human actor that he has misinterpreted the procedure currently used by the robot. According to comparable research on human-human cooperation, weak signals on a tactical level already suffice to steer interactions into the right direction. (In order to determine this, hearing and seeing of the test persons were suppressed.)

On the other hand, the robot-system is not able to record non-verbal signals of irritation and analyse them accordingly, in order to provide assistance to the human-part of the relation. Thus a big breakthough in human robot-co-operation can be achieved, if a robotic system can automatically recognise human interaction patterns by additionally interpreting non-verbal communication. However, a compromise between complex hardware and software used and intuitive handling has to be found for each co-operative robot system.

Acknowledgment

The presented classification scheme for human-robot-interaction and the performed field study are part of the research of the centre of excellence SFB 588, “Humanoid Robots – Multi-modal learning and Co-operating System”, funded by the German Research Foundation. This research has been performed at the Institute of Process Control and Robotics headed by Prof. H. Woern and at the Institute of Sociology headed by Prof. B. Schaefers. Both institutes are members of the University of Karlsruhe.

References

S. Yigit, D. Osswald, C. Burghart, H. Woern. “Concept of combined control mechanisms for human-robot-co-operation”, *Proc. CCCT 2003,* Orlando, Florida, July 2003.

T. Takubo, H. Arai, K. Tanie, “Virtual Non-holonomic Constraint for Human-Robot Cooperation in 3-D space”, *Proc. 2000 IEEE Int. Conf. on Intelligent Robots and Systems (IROS'00)*, 2000.

S. Yigit, C. Burghart, H. Wörn. Co-operative Carrying Using Pump-like Constraints. *Proc. Int. Conf. on Intelligent Robots and Systems (IROS 2004)*, Sendai, Japan, 2004.

"Robotically Rich" Environments for Supporting Elderly People at Home: the RoboCare Experience

Amedeo Cesta⋆, Alessandro Farinelli†, Luca Iocchi†, Riccardo Leone†, Daniele Nardi†, Federico Pecora⋆, Riccardo Rasconi⋆

⋆Planning and Scheduling Team
Institute for Cognitive Science and Technology
Italian National Resource Council
Viale Marx 15, I-00137 Rome, Italy
name.surname@istc.cnr.it

†Dipartimento di Informatica e Sistemistica
Università di Roma "La Sapienza"
Via Salaria 113, I-00198 Rome, Italy
surname@dis.uniroma1.it

Abstract

The aim of the ROBOCARE project is to develop an intelligent domestic environment which allows elderly people to lead an independent lifestyle in their own homes. This paper describes a testbed environment which simulates the home of an elderly person whose daily routines need to be monitored by human caregivers such as physicians or family members. We focus on the issue of how to enhance the robotic, sensory and supervising components of the system in order to achieve an environment which is at the same time pro-active and non-invasive.

1 Introduction

The long term goal of the research developed by the ROBOCARE project[1] is to contribute to raising the quality of life of elderly persons. In particular, we are pursuing the idea of developing support technology which can play a role in allowing vulnerable elderly people to lead an independent lifestyle in their own homes. This paper describes a testbed environment (Robocare Domestic Environment — RDE) aimed at re-creating the home environment of an elderly person whose daily routines need to be loosely monitored by human supervisors such as physicians or family members. The assisted person's home is equipped with some fixed and mobile environmental sensors, consisting in embedded domotic components as well as mobile robots endowed with rich interactive capabilties. All components of the system interact by means of a service-oriented infrastructure [Bahadori et al., 2004], and are coordinated by a supervision framework.

The goal of the proposed supervision infrastrucure is to preserve the independent lifestyle of a cognitively and/or physically impaired elderly person while committing to the least possible level of invasiveness. The environment must therefore adapt to the assisted person's needs: the level of pervasiveness of the supervision framework in the assisted person's daily routine must be directly proportional to the level of handicap. The aim of this paper is to describe the components, algorithms and methodologies we have developed in order to achieve such a highly customizable supervision framework.

Our main objective is to develop an intelligent environment which is at the same time "active" (in the sense that it can effectively monitor the assisted person) and also not invasive. With the term non-invasiveness, we express that the actions performed by the system as a whole on the environment should occur pro-actively and only when they are beneficial to the assisted person[2]. Given the diverse nature of the technology involved in the RDE, implementing a non-invasive system implies a rich array of design issues, which we begin to address in this paper. After giving a brief system description in the following section, we proceed in a bottom-up fashion: section 3 describes the key features of the robotic components, addressing first the aspects related to their mobility, and then the user-interaction schemes that have been adopted; section 4 describes the mechanism by which the caregivers model the behavioural constraints which are mapped against the sensor-derived information by the supervision system; we conclude with a discussion on possible future developments.

[1] http://robocare.istc.cnr.it.

[2] Recent pshychological studies [Giuliani et al., 2005] address issues related to the acceptability of technology by elderly people.

2 System Description

The overall system architecture is described in figure 1. The central component is the supervision framework, whose goal is to survey the daily routines of the assisted person and to coordinate the behavior of the embedded technological components (sensors and robots) accordingly. As shown in the figure, it consists in two fundamental modules: a Constraint Manager (CM) and an Event Manager (EM). The CM maintains a set of tasks and complex time constraints which represent the assisted person's nominal daily routine, and are cast as a scheduling problem. The tasks and constraints which compose the nominal schedule are defined by the caregivers (doctor and family member in the figure). Moreover, the CM matches the prescriptions represented by the nominal schedule to the actual behaviours of the assisted person as they are perceived by the sensors. The execution monitoring technology [Cesta and Rasconi, 2003] built into the CM propagates the sensor-derived information and detects any deviations in the assisted person's behavior from the nominal schedule. The key feature of the CM is its capability of recognizing the degree to which the assisted person's real behaviour adheres to the caregivers' prescriptions.

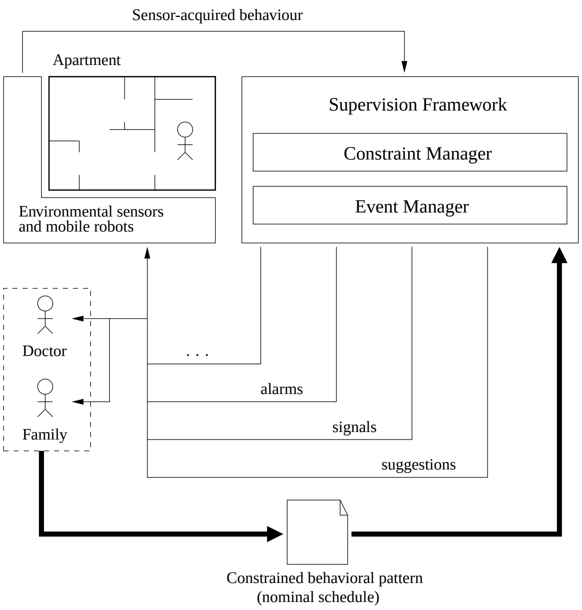


Figure 1: Overall system architecture of the ROBOCARE Domestic Environment.

The diagnosis performed by the CM during propagation is processed by the EM. It is the responsibility of the EM to trigger the appropriate event according to the specific behavioral constraint which is violated. For instance, the system should set off an alarm if, based on sensor-derived information, the CM detects that the assisted person's current behaviour compromises the successful completion of another important task. The EM defines how the system reacts to the contingencies in the nominal schedule by triggering events such as robot service invocations, alarms, suggestions, or simple logging of events.

The robotic subsystem which enhances the assisted person's domestic environment is composed of fixed and mobile components. Also these components have been engineered to reflect our main objective of low invasiveness. To this end, we have equipped our robots with localization and path-planning strategies which are oriented towards maintaining high levels of safety while ensuring adequate mobility. Moreover, human-robot interfaces have been developed using simple graphical schemes of interaction based on strong ergonomy and usability requirements. Solutions such as the use of clearly distinguishable buttons, high-contrast color schemes and input/output redundancy have been employed in an attempt to minimize the impact of high technology on the end-user.

3 Ergonomic Embedded Technology

The introduction of robots in domestic environments is a complex issue both from the technological point of view (houses are highly de-structured) and from the typology of the end-user (elderly people do not like to change their habits or to have their spaces reduced). An elderly person may have reduced physical and/or cognitive capabilities which can represent a barrier for the use of high-tech instrumentation.

Psychological studies [Scopelliti et al., 2004] show that in order to be successful in this project it is necessary that the elderly people perceive the robots as "friendly creatures" which are of some help in their every day life. The cohabitation with another beings, even though artificial, has beneficial effects on the individual, in the same way as with pets.

Hence the need to endow the robots with the capability to interact with people according to natural communication schemes: oral dialogues, facial expressions, prossemic and kinesic signals.

3.1 Robotic and sensory system

At the present stage of development, the RDE hosts three types of embedded technological components:

- stereo color camera based sensor, located in fixed positions of the environment;

- Pioneer 3 AT mobile robots, equipped with a ring of sonars, a Sick laser range finder device and a color omni-directional camera;
- palm devices for user interaction.

These three components are able to share information through a wireless network which covers the whole environment, and interact according to a service-oriented paradigm [Bahadori et al., 2004]. Our work focuses on monitoring-specific services, namely *People tracking* and *People localization* services provided by the fixed stereo camera, a *Objects Delivery* service provided by the mobile base, and a *Visualize* service provided by a Personal Digital Assistance, which allows a human operator to visualize the current state of the mobile robot through the palm device.

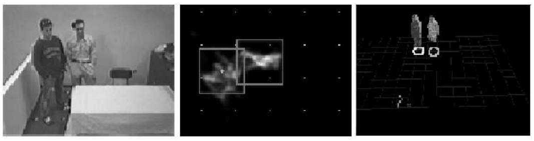


Figure 2: The different phases for people localization: original image, planar view, and 3D view. The two subjects are correctly mapped also in the presence of occlusions.

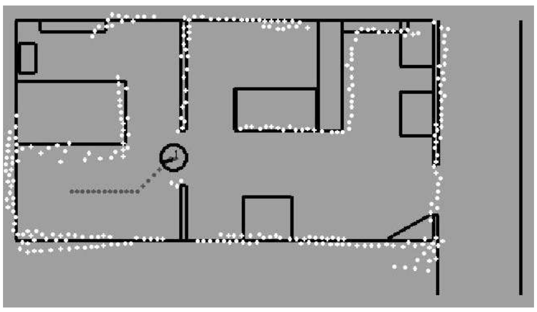
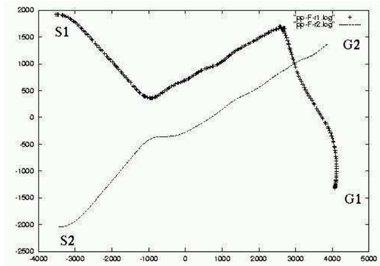
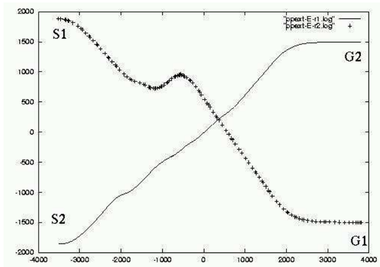


Figure 3: The robot autonomously navigating the RoboCare environment.

The *People localization* service is invoked to recognize a human being who is present in the environment, and to compute his/her coordinates with respect to the camera (see figure 2). The *People tracking* service is able to track a person in the environment following its movements. Moreover, the stereo camera is capable of correctly mapping partially occluded elements of the scene. The *Object Delivery* service allows the mobile base to safely navigate the environment bringing a light-weight object in a desired position. In particular, the robot is able to localize itself inside the environment, compute the best path to reach the desired position and follow the path while



(a)



(b)

Figure 4: Behavior of the robot without considering the obstacle dynamics (a), and behavior of the robot when the obstacle dynamics are taken into account (b).

avoiding possible unexpected obstacles such as moving person (see figure 3). The *Visualize* service exports the internal state of the robot to a palm device, in particular the service provides the robot's position in the environment, the current action the robot is performing (e.g., following a path), the current sensor readings (e.g., the obstacle detected through the sonars), or an image of the environment obtained with the on board camera (see next section for a detailed description).

A main requisite for the design of the embedded technology is to provide flexible solutions which can be easily integrated inside the environment. A necessary condition for minimizing the level of invasiveness of the technology is that it should not require re-engineering the environment. Flexibility and adaptation to the environment are crucial issues for the embedded technology, because the deployed devices are often intended to interact physically with the target user, and thus can interfere with everyday activities.

In order to satisfy these specifications, we have adopted a series of design choices aimed at adapting robots and hardware devices to the environment where they should operate. A first, fundamental issue is robot mobility. The navigation capabilities of robots are achieved without any changes to the domestic environment; no artificial markers are needed to localize a robot, and its path planning capabilities are designed to achieve safe navigation in cluttered environments with object of any shapes. In this way, the target user is not required to adapt the furniture or the colors of his or her living environment. Moreover, the path planning method (as described in [Farinelli and Iocchi, 2003]) explicitly takes into account the possibility of having persons moving in the environment. The method is able to take into account the movements of other persons in the environment, yielding in order to allow them to pass first. The people localization service does not rely on any device or particular cloth the target users should ware, rather, it automatically detects a person based on a fore-ground extraction method [Bahadori et al., 2004].

Figure 5: The Palm PDA interface after issuing the *WhatRUDoing* command.

The control of the robot is based on a high level representation of the world and on cognitive capabilities. For example, the robot is able to represent and recognize objects in the surrounding environment and to localize itself inside the environment. Since all the components are connected via the wireless network, in the execution of such behaviors the robot can use the information acquired by the stereo camera to map a person inside its own representation of the world. Using such high level information, the robot is able to execute complex plans which comprise the execution of several atomic actions. In this way the robot can perform a set of high level behaviors making it much easier for humans to interact with it.

All the components previously described have been tested and evaluated both in specific experiments and in coordinated demos. The interested reader can find more details on the specific methods in [Farinelli and Iocchi, 2003] and [Bahadori et al., 2004]. In particular for the path-planning method, specific experiments show how the behavior of the robot has improved, considering in the path-planning process the dynamics of the obstacles. Figure 4(a) and figure 4(b) represent the paths followed by the robot when a moving obstacle crosses its way. The robot's initial position is S_1 and its final destination is G_1. In figure 4(a) the path planning method does not take into account the obstacle dynamics (i.e., the velocity vector of the obstacle), while in figure 4(b) such information is conveniently exploited and the robot decides to pass behind the obstacle generating a path which is not only more convenient, but also safer.

3.2 Human Robot Interfaces

The main interaction between the assisted person and the system occurs through the use of a PDA (Personal Digital Assistant). The key idea is based on the fact that the PDA constitutes a sort of remote control as it represents the means by which the user can ask for service activation. The PDA is an instrument characterized by an extremely light weight and this makes it suitable to be easily carried by the assisted person; as a downside, its small size reduces the possibility of using its touch-screen as a full-functionality interface. For this reason it is necessary to implement some input/output features on the PDA's audio channel. The communication between the PDA and the rest of the system occurs in wireless mode.

The exported services are organized in two main categories depending on which event triggers them: i) the occurrence of a specific event, and ii) a user request. Services belonging to the first set are triggered either in presence of some kind of errors (for instance, *unrecognized vocal command*) or on occurrence of scheduled activities (e.g., *it's time to take the medicine* or *the news will start in five minutes*).

The services triggered by a user's request are tasks which are obviously not present in the original schedule.

Let us give an example of interaction with a single robotic agent. The services provided by the agent can be summarized as follows:

ComeHere instructs the robot to reach the user (which is equivalent to reaching the PDA)

WhatRUDoing allows to visualize the activity performed by the robot through the use of the on-board camera and receive some oral information related to the same activity

Go(*where*) instructs the robot to go to the place specified by the user in the parameter *where*

Stop instructs the robot to interrupt all the activities requested by the user (the activities belonging to the original schedule obviously continue their execution)

The interface main screen provides four buttons one for each of the previous services. Such functionalities are also associated to the four programmable buttons of the PDA. In case the user pushes the *Go* button, the *where* parameter can be specified by selecting the destination room directly from the environment map that appears on the screen. When the user selects the *WhatRUDoing* command, the PDA will reproduce both the instant image coming from the on-board camera, as well as the position of the robot in the house (see figure 5); clicking on the previous image returns a full screen picture, for better visualization. Another click leads back to the initial menu.

4 Monitoring Daily Routines

Now that we have described some aspects related to the acceptability of the sensors and robotic components embedded in the RDE, we address some issues related to the form of interaction between the supervision framework and the caregivers. In this section we describe the nature of the behavioral constraint specifications which are defined by the caregivers for the supervision framework to monitor. In particular, we show a modeling framework which allows the caregivers to harness the full expressiveness of the underlying category of scheduling problems.

As mentioned, the assisted person's daily behaviour is modeled as a set of activities and complex temporal constraints. The core technology we deploy consists in a CSP-based scheduler [Cesta et al., 2001] equipped with execution monitoring capabilities [Cesta and Rasconi, 2003], which is able to deal with rather complex scheduling problemsThis high complexity is supported by a highly expressive scheduling formalism which allows, among other things, for the definition of complex temporal relationships among tasks, such as minimum and maximum time lags[3]. The need for a highly expressive scheduling formalism[4] for the purpose of specifying the assisted person's behavioral constraints can be appreciated in the fact that often the constraints consist of complex time relationships between the daily tasks of the assisted person. Also, given the high degree of uncertainty in the exact timing of task execution (a person never has lunch at the same time every day, etc.), it is necessary to model flexible constraints among the tasks, while admitting the possibility of hard deadlines or fixed time-points. Overall, the aim is not to control task execution, nor to impose rigid routines, rather it is to monitor the extent to which the assisted person adheres to a predefined routine, defined together with a physician or family member.

Operator	*Semantics*
`create_task(t,min,max)`	Creates a task named `t` whose minimum and maximum durations are `min` and `max`.
`create_res(r,cap)`	Creates a resource named `r` whose capacity is `cap`.
`set_res_usage(t,r,use)`	Imposes that task `t` uses `use` units of resource `r`.
`create_pc(t1,p1,t2,p2,x)`	Imposes a precedence constraint of `x` time-units between time-point `p1` of task `t1` and time-point `p2` of task `t2`.

Figure 6: The four elementary operators for building scheduling problem instances.

The technical details of how the caregivers' prescriptions are cast into a scheduling problem is outside the scope of this paper. It is sufficient to mention that the expressiveness of the temporal problem which is cast is completely captured by the four basic operators shown in figure 6.

What we would like to emphasize here is that such a versatile specification formalism allows us to model with very high precision the behavioural constraints for the assisted person. This ability to describe real-

[3] As well known in the scheduling community, the introduction of maximum time lag constraints increases problem complexity from P to NP.

[4] Similar attempts at using core solving technology in domestic and health-care environments have been made (e.g. [McCarthy and Pollack, 2002; Pollack et al., 2002]).

ity with the required degree of granularity makes it possible to always maintain the desired level of flexibility in the specification of the necessary constraints. Indeed, this implies a low level of invasiveness because the synthesized behavioral pattern is never constrained beyond the real requirements prescribed by the caregivers.

Clearly, this versatility comes at the cost of a high complexity of the specification formalism. Indeed, the four operators shown above are rather straightforward, but building a complex scheduling problem using these operators can be a demanding task even for a scheduling expert. Moreover, modeling behavioral constraints in the context of the RDE in this fashion would turn out to be not only tedious but also definitely out of reach for someone not proficient in scheduling.

A key issue is thus represented by the fact that the monitoring framework should be designed to meet the requirements of different types of end-users, each having different needs: for instance, a doctor might be interested in monitoring activities which pertain to health control, while the assisted person's relatives might instead be concerned with the recreational aspects of the person's daily life. In order to enable these different users to easily interact with the supervision framework we have deployed in the RDE, we employ a knowledge representation layer for problem modeling, built around the core scheduling technology which implements the CM module. This layer allows the end-user to easily specify behavioural constraints for the assisted person while ignoring the technicalities of how these constraints are cast into the underlying core scheduling formalism[5]. In the following section we describe by means of a simplified example how the introduction the knowledge representation layer makes our monitoring technology accessible to the caregivers.

4.1 Modeling Framework

In order to provide the caregivers with a modeling tool which hides the technology-specific details while maintaining the necessary expressiveness, we proceed in two steps:

Domain definition. The first step is to define the types of tasks which are to be monitored and the types of constraints which can bind them. This equates to formalizing the types of medical requirements and behavioral patterns which can be prescribed by the human supervisors. The result of this requirement analysis is what we call a *domain description*. A domain encapsulates the scheduling-specific knowledge for the definition of the behavioral constraints, and provides usable "building blocks" for the particular category of caregiver to use. These building blocks, called *constructs*, constitute a terminology which is tailored to the expertise of the particular caregiver.

Instantiation. The caregivers can at this point employ the particular domain which has been built for them to define the constraints for the assisted person. A physician, for instance, may use the "RDE-medical-requirement" specification terminology specified in the domain which was created for such purposes. A domain definition process which is correctly carried out yields a collection of constructs which match the supervisors' usual terminology, and mask completely the scheduling-specific knowledge otherwise needed for schedule specification. The particular requirements for the assisted person are thus defined in the form of construct instantiations, which are consequently passed on to the monitoring system.

Once the nominal schedule is established by the caregivers, all execution-time variations to the schedule are taken into account by the execution monitor: by polling the sensors, the execution monitor gathers information on the real state of execution of the tasks, and employs the CM to propagate any variations. The key idea is that if any of these variations violate a constraint then the proper actions are triggered by the EM (such as alarms, reminders, and so on).

4.2 RDE Domain Formalization

We now show a simplified example domain specification which defines some typical behavioral and medical requirements of the assisted person. As mentioned above, this domain defines a set of constructs any instantiation of which is an "encoding" of a set of requirements to which the assisted person's routine should adhere. In the following paragraph, we omit the details of the construct definitions, limiting the presentation to a simplified description of how the constructs define the underlying scheduling problem.

Domain definition. Let us start with the basic construct for defining the assisted person which is being supervised:

```
(:construct assisted_person
  :parameters (name) ...  )
```

[5] The scheduling specific details as how this compilation occurs are outside the scope of this paper, and are described in [Cesta et al., 2005].

This construct defines a binary resource corresponding to the assisted person. This reflects the assumption that the assisted person carries out at most one task (of the tasks which are monitored) at any instant in time. This is guaranteed by the fact that every construct in this domain uses exactly one unit of this binary resource. It should be clear that behaviors in which there is some degree of concurrency can be modeled by increasing the capacity of this resource.

Another requirement of the monitoring system is to oversee the dietary habits of the assisted elderly person. To this end, we define the following three constructs:

```
(:construct breakfast
  :parameters (person start end) ... )
(:construct lunch
  :parameters (person start end min_bfast
    max_bfast) ... )
(:construct dinner ... )
  :parameters (person start end min_lunch
    max_lunch) ... )
```

The reason for modeling breakfast, lunch and dinner (rather than a single `meal` construct) is because the caregivers need to ascertain the regularity of the assisted person's diet. For instance, through the specification of the `min_lunch` and `max_lunch` parameters, it is possible to model the upper and lower bounds between one meal and another. Thus, the instantiation `(dinner 1200 1260 180 360)` in the problem definition (time units are in seconds) equates to stating that (1) the assisted person's nominal time for dinner is from 8 pm to 9 pm, (2) the assisted person should have dinner at least three hours after lunch, and (3) he or she should not have dinner more than six hours after lunch.

In addition to the dietary constraints, medical requirements are also specified by means of the `medication` construct:

```
(:construct medication
  :parameters (person product dur min_time
    max_time) ... )
```

The construct prescribes that a medication cannot be taken before `min_time`, nor after `max_time`, which in turn are user definable parameters of the construct. This is achieved by constraining the start time-points of the task with the beginning of the time horizon. Similarly, a construct which imposes lower and/or upper bounds on medication with respect to meals is provided:

```
(:construct meal_bound_medication
  :parameters (person product dur meal
    min max) ... )
```

For example, by specifying `(meal_bound_medication roger aspirin 5 lunch 0 25)`, we model that Roger can take an Aspirin potentially immediately after lunch, but without exceeding twenty-five minutes.

Instantiation. A problem specification based on the domain described above is shown below:

```
(define (problem test_prob)
  (:domain RDE)
  (:specification
    (assisted_person jane)
    (breakfast jane 480 510)
    (lunch jane 780 840 240 360)
    (dinner jane 1170 1290 300 360)
    (meal_bound_medication jane aspirin 5
      dinner 0 20)
    (medication jane herbs 10 720 1200)
    (medication jane laxative 5 1020
      1260)))
```

It is interesting to point out some of the design decisions which were made in the domain definition. Notice that all tasks have a fixed duration, a fact which may seem counter-intuitive in this domain. For instance, we have no reason to believe that Jane's breakfast lasts half an hour, nor can we commit to any other projected duration since it will always be wrong. On the other hand, establishing a lower or upper bound on the duration of her meals would just as well be unfounded. Thus, this uncertainty is dealt with by the CM, which dynamically adapts the duration of the tasks to the sensors' observations. The durations of the tasks are thus kept fixed in the problem specification since the execution monitor does not trigger an alarm when they are not respected. An alarm may however be triggered in the event that the sensed deviation from the nominal duration causes other serious violations of behavioural constraints in the nominal schedule. In general, the constraints modeled in the domain can be treated variety of ways: some constraints, such as task durations in the specific example shown above, are "soft", meaning that their purpose is solely that of modeling the assisted person's nominal behaviour; other constraints, such as the relationship between meals and medication in the above example, are "hard", meaning that if they are violated, this represents a contingency which calls for a specific event (such as an alarm, a notification and so on). In the light of these considerations, the constructs defined in the domain must be seen as elements of a language with which a caregiver can express (1) which events in the daily routine he or she would like to supervise (e.g., Jane should take an Aspirin every day), (2) how these events are related to each other in terms of "causality" (e.g., since Aspirin needs to be taken with a full stomach, having dinner is a precondition for taking an Aspirin), and (3) the degree to which the assisted person should comply to the nominal schedule (e.g., Jane cannot wait more than twenty minutes after she has finished dining to take her Aspirin).

5 Conclusion and Future Work

In this paper we have described some aspects related to the design of an intelligent domestic environment for the care of elderly people. We have mainly focused on the design choices which minimize the level of invasiveness of the embedded technology. We have shown how this goal is pursued both in the development of the hardware components and in the implementation of the supervision framework. As we have seen, endowing domestic robots with more "human-centered" features, such as intelligent obstacle avoidance schemes and intuitive human-robot interfaces, is critically important if robotic components are to be accepted in domestic environments. Similarly, we strive to provide caregivers with intelligent monitoring tools which are also extremely configurable around the very particular requirements of a particular assisted person. We argue that adaptability is a determining factor for the successful deployment of ambient intelligence in domestic environments.

The work we have presented in this article represents a first step towards a fully-customizable supervisory system, and is part of a larger effort started in 2003 with the ROBOCARE project, in which the issues related to human-robot interaction are extremely relevant. While the question of broadening the scope of application of robots for the care of the elderly is still a very open issue, we believe that one important reason which justifies a wider utilization in contexts such as the RDE lies in concealing their qualities as technological aides behind a friendly appearance.

Acknowledgements

This research is partially supported by MIUR (Italian Ministry of Education, University and Research) under project ROBOCARE (A Multi-Agent System with Intelligent Fixed and Mobile Robotic Components).

References

S. Bahadori, A. Cesta, L. Iocchi, G.R. Leone, D. Nardi, F. Pecora, R. Rasconi, and L. Scozzafava. Towards Ambient Intelligence for the Domestic Care of the Elderly. In P. Remagnino, G.L. Foresti, and T. Ellis, editors, *Ambient Intelligence: A Novel Paradigm*. Springer, 2004. To appear.

A. Cesta, C. Cortellessa, F. Pecora, and R. Rasconi. Mediating the Knowledge of End-Users and Technologists: a Prolem in the Deployment of Scheduling Technology. In *Proceedings of the International Conference on Artificial Intelligence and Applications (AIA'05), Innsbruck, Austria*, 2005.

A. Cesta, G. Cortellessa, A. Oddi, N. Policella, and A. Susi. A Constraint-Based Architecture for Flexible Support to Activity Scheduling. In *LNAI 2175*, 2001.

A. Cesta and R. Rasconi. Execution Monitoring and Schedule Revision for O-OSCAR: a Preliminary Report. In *Proceedings of the Workshop on Online Constraint Solving at CP-03, Kinsale Co. Cork*, 2003.

A. Farinelli and L. Iocchi. Planning trajectories in dynamic environments using a gradient method. In *Proc. of the International RoboCup Symposium 2003*, Padua, Italy, 2003.

M.V. Giuliani, M. Scopelliti, and F. Fornara. Coping Strategies and Technology in Later Life. In *Proceedings of Workshop on Robot Companions, AISB'05 Convention, Hatfield, UK*, 2005.

C.E. McCarthy and M.E. Pollack. A Plan-Based Personalized Cognitive Orthotic. In *Proceedings of the 6th International Conference on AI Planning and Scheduling*, 2002.

M.E. Pollack, C.E. McCarthy, S. Ramakrishnan, I. Tsamardinos, L. Brown, S. Carrion, D. Colbry, C. Orosz, and B. Peintner. Autominder: A Planning, Monitoring, and Reminding Assistive Agent. In *Proceedings of the 7th International Conference on Intelligent Autonomous Systems*, 2002.

M. Scopelliti, M.V. Giuliani, A.M. D'Amico, and F. Fornara. If I had a robot: Peoples' representation of domestic robots. In Keates, S. and Clarkson, P.J. and Langdon, P.M. and Robinson, P., editor, *Design for a more inclusive world*, pages 257–266. Springer-Verlag, 2004.

Embodied social interaction for robots

Henrik I Christensen* & Elena Pacchierotti*

* Centre for Autonomous Systems
Kungl Tekniska Hgskolan
10044 Stockholm, Sweden
{hic,elena-p}@nada.kth.se

Abstract

A key aspect of service robotics for everyday use is the motion of systems in close proximity to humans. It is here essential that the robot exhibits a behaviour that signals safe motion and awareness of the other actors in its environment. To facilitate this there is a need to endow the system with facilities for detection and tracking of objects in the vicinity of the platform, and to design a control law that enables motion generation which is considered socially acceptable. We present a system for in-door navigation in which the rules of proxemics are used to define interaction strategies for the platform.

1 Introduction

Service robots are gradually entering our everyday life. Already now more than 1 000 000 vacuum cleaners are in use in houses around the world (Karlsson, 2004). We are also starting to see robots being deployed as hospital logistic aids such as those provided by FMC Technologies for transportation of meals and linen. Not to mention the AIBO dog type robots that are provided by Sony. Gradually we are starting to see service types robots for assistance to people in terms of everyday tasks such as mobility aid, pick-up of objects, etc. As part of operation in public spaces it is essential to endow the robots with facilities for safe navigation in the vicinity of people. Navigation entails here both the safe handling of obstacles, going to specific places, and maneuvering around people present in the work area. For the interaction with human we see at least two modes of interaction: i) instruction of the robot to perform specific tasks incl. generation of feedback during the command dialogue, and ii) the embodied interaction in terms of motion of the robot. The embodied (non-verbal) interaction is essential to the perception of safety when the robot moves through the environment. The speed of travel that is considered safe very much depends upon the navigation strategy of the overall system.

Several studies of interaction with people have been reported in the literature. Nakauchi and Simmons (2000) report on a system for entering a line for a conference in which there is a close proximity to other users. Here the robot has to determine the end of the line and align with other users. Althaus et al. (2004) report on a system in which group dynamics is studied so as to form natural distances to other people in a group during informal discussions. The control involves entering and exiting the group and alignment with other actors in the group. Passage of people in a hallway has been reported in Yoda and Shiota (1996, 1997). However few of these studies have included a directly analysis of the social aspects. They have primarily considered the overall control design.

In the present paper we study the problem of physical interaction between a robot and people during casual encounters in office settings. The encounters are with people that are assumed to have little or no direct model of the robots actions, and the interaction is consequently assumed to be with naive users. The encounters are in terms of meeting and passing robots that operate in room or corridor settings. Similar studies have been performed with users in professional environments such as hospitals, but we are unfortunately unable to report on the results of these studies.

The paper is organised with an initial discussion of social interaction during passage and person-person interaction in an informal setting in Section 2. Based on these considerations a strategy for robot control is defined in Section 3. To enable the study of behaviours in real settings a system has been implemented which allows us to study the problem. The implementation is presented in Section 4. Early results on the evaluation of the system are presented in Section 5. These early results allow us to identify a number of issues that require further study. A discussion of these challenges is presented in Section 6.

Finally a number of conclusions and option issues are provided in Section 7.

2 Physical Passage 101

The spatial interaction between people has been widely studied in particular in psychology. The studies go back several centuries, but in terms of formal modelling one of the most widely studied model is the one presented in Hall (1966) which is frequently termed *proxemics*. The literature on proxemics is abundant and good overviews can be found in Aiello (1987) and Burgoon et al. (1989). The basic idea in proxemics is to divide space around the person into four categories:

Intimate: This ranges up to 30 cm from the body and interaction within this space might include physical contact. The interaction is either directly physical such as embracing or private interaction such as whispering.

Personal: The space is typically 30-100 cm and is used for friendly interaction with family and for highly organised interaction such as waiting in line.

Social: the range of interaction is here about 100-300 cm and is used for general communication with business associated, and as a separation distance in public spaces such as beaches, bus stops, shopping, etc.

Public: The public space is beyond 300 cm and is used for no interaction or in places with general interaction such as the distance between an audience and a speaker.

It is important to realize that personal spaces vary significantly with cultural and ethnic background. As an example in in Saudi Arabia and Japan the spatial distances are much smaller, while countries such as USA and the Netherlands have significant distances that are expected to be respected in person-person interaction.

Naturally one would expect that a robot should obey similar spatial relations. In addition there is a need to consider the dynamics of interaction. The passage and movement around a person also depends on the speed of motion and the signaling of intentions, as is needed in the passage of a person in a hallway. As an example when moving frontally towards a robot one would expect the robot to move to the side earlier, where as a side-side relation is safer, due to the kinematic constraints. Consequently one would expect that the proxemics relations can be modelled as elliptic areas around a person as shown in Figure 1.

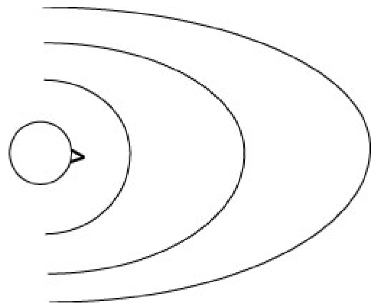


Figure 1: The interaction zones for people moving through a corridor setting

Video studies of humans in hallways seem to indicate that such a model for our spatial proxemics might be correct (Chen et al., 2004). One should of course here point out that the direction of passage also is an important factor. The “patterns of motion” is tied to social patterns of traffic in general. I.e. in Japan, UK, Australia, . . . the passage is to the left, while in most other countries it is to the right of the objects in a hallway. The general motion of people is closely tied to these patterns.

3 Design of a control strategy

Given that proxemics plays an important role in person-person interaction, it is of interest to study if similar rules apply for interaction with robots operating in public spaces. To enable a study of this a number of basic rules have been defined. The operation of a robot in a hallway scenario is presented here. Informally one would expect a robot to give way to a person when an encounter is detected. Normal human walking speed is 1-2 m/s which implies that the avoidance must be initiated early enough to signal that it will give way to the person. In the event of significant clutter one would expect the robot to move to the side of hallway and stop until the person(s) have passed, so as to give way. To accommodate this one would expect a behaviour that follows the rules of:

1. upon entering the social space of the person initiative a move to the right (wrt to the robot reference frame).
2. move far enough to the right so as not to enter into the personal space of the person while passing the person.

3. await a return to normal navigation until the person has passed by. A too early return to normal navigation might introduce uncertainty in the interaction.

Using the rules of proxemics outlined in Section 2, one would expect the robot to initiate avoidance when the distance is about 3 meters to the person. The avoidance behaviour is subject to the spatial layout of environment. If the layout is too narrow to enable passage outside of the personal space of the user (i.e. with a separation of at least 1 meter) the robot should park at the side of the hallway. The strategy is relatively simple but at the same time it obeys the basic rules of proxemics.

4 Implementation of a passage system

To test the basic rule based design presented in Section 3 a prototype system has been implemented on a Performance PeopleBot with an on-board SICK laser scanner, as shown in Figure 2. The system was

Figure 2: The PeopleBot system used in our studies

designed to operate in the hallways of our institute which are about 2 meters wide, so the hallways are relatively narrow. To evaluate the system there is a need to equip it with methods for:

- Detection and tracking of people.
- Navigation in narrow spaces with significant clutter.
- Path planning with dynamically changing targets to circumvent people and other major obstacles.

Each of the methods are briefly outlined below.

4.1 Detection of people

The detection of people is based on use of the laser scanner. The laser scanner is mounted about 20 cm above the floor. This implies that the robot will either detect the two legs of the person or a single skirt. To allow detection of people scan alignment is performed Gutmann and Schlegel (1996) which enable differencing of scans and detection of motion. The scan differencing is adequate for detection of small moving objects such as legs. Using a first order motion model it is possible to estimate the joint motion of the legs or the overall motion of a single region (the skirt). Tracking is complicated by partial occlusions and significant motion of legs, but the accuracy of the tracking is only required to have an accuracy of ± 10 cm to enable operation. The tracker is operating at a speed of 6Hz, which implies that the motion might be up to 30 cm between scans. The ambiguity is resolved using the first order motion model in combination with fixed size validation gates (Bar-Shalom and Fortmann, 1987). The detection function generates output in terms of the position of the centroid of the closest person. In the event of more complex situations such as the presence of multiple persons a particle filter can be used as for example presented by Schulz et al. (2001).

4.2 Rules of interaction

The basic navigation of the system is controlled by a trajectory following algorithm that drives the system towards an externally defined goal point. A collision avoidance algorithm based on the Nearness Diagram method (Minguez and Montano, 2004) drives the robot safely to the final location. The environmental information is generated in form of a local map that integrates laser and sonar data. During interaction with a person the following strategy is used:

- As soon as the robot enters in the social space of the person, determine if there is space available for a right passage (given knowledge of the corridor provided by the localisation system).
- If passage is possible define a temporary goal point about 1 meter ahead and 1 meter to the right with respect to the current position of the robot.
- Upon entering an area ± 10 cm of the temporary goal point, define a new intermediate goal point that allows the robot to pass the person.
- Upon passage of the person resume the navigation task.

- If no passage is possible park the robot close to the right side and resume the navigation task once the person(s) have passed down the hallway.

4.3 The implemented system

The methods outlined above have been implemented on the PeopleBot (minnie) in our laboratory. The system uses an on board Linux computer and the control interface is achieved using the Player/Stage system (Vaughan et al., 2003) for interfacing. The SICK laser scanner and the sonar data are fed into a local mapping system for obstacle avoidance. In addition the laser scanner is fed into a person detection / tracking system. The output from the mapping system is used by the nearness diagramme based trajectory follower that ensures safe passage through cluttered environments. All the software runs in real-time at a rate of 6Hz. The main bottleneck in the system is the serial line interface to the SICK scanner.

5 Early evaluation

The system has been evaluated in real and simulated settings. The tests in real settings have involved both hallway environments and open spaces such as a department kitchen or large living room. To illustrate the operation of the system a single run is presented here.

In figure 3 is shown a setup in which the robot (blue trajectory) is driving down a hallway. The robot is about 3 m away from the person and is thus entering the social space of the approaching person.

At this point in time the robot selects a point to the right of the person and initiates an avoid maneuver. The turn is abrupt to clearly signal that it is give way to the person. The trajectory is shown in Figure 4. The red cross clearly marks the temporary goal of the robot.

As the robot proceeds on the passage trajectory, it passes the user (actually the user disappears from the sensor's field of view), as is shown in Figure 5.

Upon completion of the passage behaviour the robot resumes its original trajectory which is the reason for the sharp turn towards its final destination, as shown in Figure 6.

The results presented here are preliminary and to fully appreciate the behaviour of the robot for operation in public spaces there is a need to perform a user study. It is here also of interest to study how velocity of motion and variations in the distance will be perceived by people that encounter the robot. Such stud-

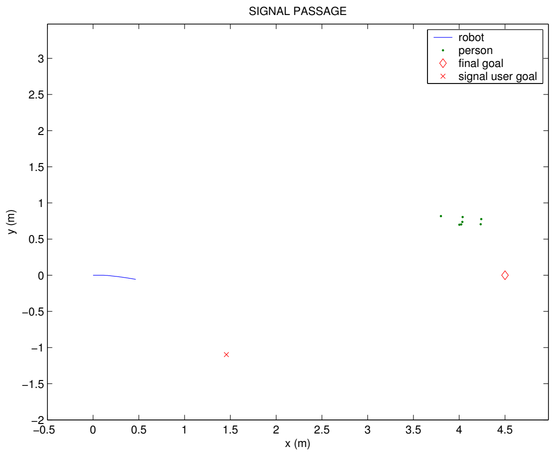


Figure 3: The initial detection of an encounter. The robot is driving towards the diamond marker when it detects a person moving in the opposite direction.An intermediate goal is defined that allows the robot to steer to the right

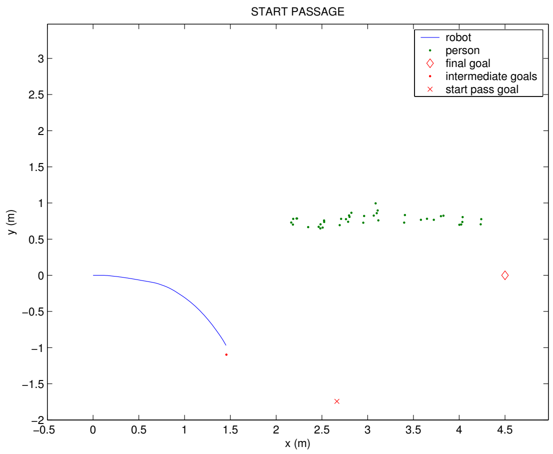


Figure 4: The initial avoid trajectory of the robot as it signals that it is giving way to the approaching person

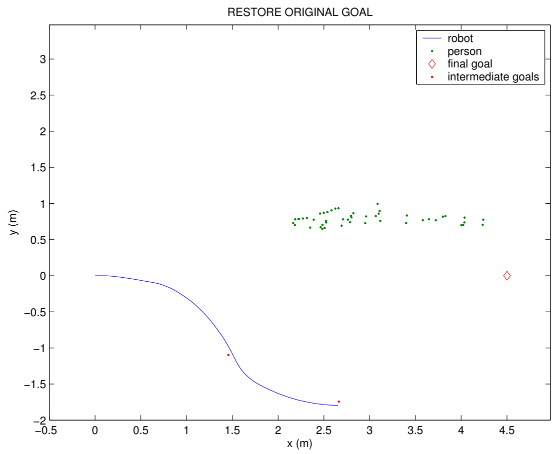


Figure 5: The passage of the person is continued until the person disappears from the field of view of the sensor

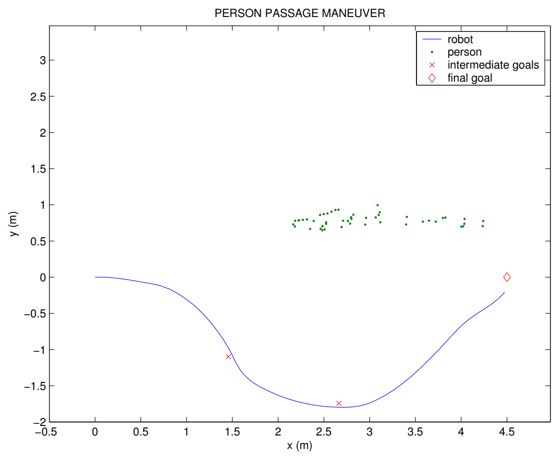


Figure 6: The completion of the passage behaviour

ies must be performed before any final conclusions on the proposed method can be given.

6 Challenges in embodied interaction

A passage behaviour is merely one of several behaviours that are required in the design of a system that operates in public spaces and interacts with naive users. The motion of the robot is crucial to the perception of the system. Simple jerky motion results in a perception of instability, and smoothness is thus an important factor. For operation in daily life situations there is further a need to consider the direct interaction with people so as to receive instructions from a user. As part of such actions there is a need to consider other social skills such as

- How to approach a person in a manner that signals initiation of a dialogue?
- If following a person, how fast can you approach a person from behind before it is considered tail gating?
- When entering into a group how is the group structure broken to enable entry?
- In a tour scenario where a person directs the person around, the robot is required to follow at a certain distance after the user, but when receiving instructions there might be a need to face the user to interpret gestures and to use speech to receive instructions. How can both be achieved in a manner that is respectful and at the same time not too slow?
- In office buildings there might be a need to utilize elevators to enable access to multiple floors. How is the robot to behave for entering and exiting the elevator? Often elevators are crammed spaces and there is limited room to allow correct behaviour. If the robot is too polite it might never be admitted to an elevator in which there are people present. Many robots have a front and as such are required to enter the elevator and turn around, which in itself poses a challenge in terms of navigation. How can the robot signal intent to enter an elevator without being considered rude?

The embodied interaction with people is only now starting to be addressed and it is an important factor to consider in the design of a system, as both the physical design and motion behaviours are crucial to the acceptance of a final system by non-expert users.

7 Summary

As part of human robot interaction there is a need to consider the traditional modalities of interaction such as speech, gestures and haptics, but at the same time the embodied interaction, the body language of the robot, should be taken in account. For operation in environments where users might not be familiar with robots this is particularly important as it will be in general assumed that the robot behaves in a manner similar to humans. The motion pattern of a robot must thus be augmented to include the rules of social interaction. Unfortunately many of such rules are not formulated in a mathematically well-defined form, and thus there is a need for transfer these rules into control laws that can be implemented by a robot. In this paper the simple problem of passage of a person in a hallway has been studied and a strategy has been designed based on definitions borrowed from proxemics. The basic operation of a robot that utilizes these rules has been illustrated. The hallway passage is merely one of several different behaviours that robots must be endowed with for operation in spaces populated by people. To fully appreciate the value of such behaviours there is still a need for careful user studies to determine the utility of such methods and to fine-tune the behaviours to be socially acceptable.

Acknowledgements

This research has been sponsored by the Swedish Foundation for Strategic Research through its Centre for Autonomous Systems, the CEC as part of Cognitive Systems for Cognitive Assistants – CoSy. The financial support is gratefully acknowledged. The work has benefited from discussions with Prof. K. Severinson-Eklundh, Ms. E. A. Topp, Mr. H. Hüttenrauch, and Mr. A. Green.

References

J. R. Aiello. Human spatial behaviour. In D. Stokels and I. Altman, editors, *Handbook of Environmental Psychology*. John Wiley & Sons, New York, NY, 1987.

P. Althaus, H. Ishiguro, T. Kanda, T. Miyashita, and H. I. Christensen. Navigation for human-robot interaction tasks. In *Proceedings of the IEEE International Conference on Robotics and Automation*, volume 2, pages 1894–1900, April 2004.

Y. Bar-Shalom and T. Fortmann. *Tracking and Data Association*. Academic Press, New York, NY., 1987.

J. Burgoon, D. Buller, and W. Woodall. *Nonverbal Communication: The unspoken dialogue*. Harper & Row, New York, NY, 1989.

D. Chen, J Yang, and H. D. Wactlar. Towards automatic analysis of social interaction patterns in a nursing home environment from video. In *6th ACM SIGMM Int'l Workshop on Multimedia Information Retrieval*, volume Proc of ACM MultiMedia 2004, pages 283–290, New York, NY, October 2004.

J-S. Gutmann and C. Schlegel. Amos: comparison of scan matching approaches for self-localization in indoor environments. In *Proc. of the First Euromicro on Advanced Mobile Robot*, pages 61–67, 1996.

E.T. Hall. *The Hidden Dimension*. Doubleday, New York, 1966.

J. Karlsson. *World Robotics 2004*. United Nations Press/International Federation of Robotics, Geneva, CH, October 2004.

J. Minguez and L. Montano. Nearness Diagram Navigation (ND): Collision avoidance in troublesome scenarios. *IEEE Trans on Robotics and Automation*, 20(1):45–57, Feb. 2004.

Y. Nakauchi and R. Simmons. A social robot that stands in line. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, volume 1, pages 357–364, October 2000.

D. Schulz, W. Burgard, D. Fox, and A. B. Cremers. Tracking multiple moving objects with a mobile robot. In *CVAP-01*, Kauai, HW, December 2001. IEEE.

R.T. Vaughan, B. Gerkey, and A. Howard. On device abstraction for portable, reusable robot code. In *IROS-03*, pages 2121–2127, Las Vegas, NV, Oct. 2003.

M. Yoda and Y. Shiota. Analysis of human avoidance motion for application to robot. In *Proceedings of the IEEE International Conference on Robot and Human Communication*, pages 65–70, November 1996.

M. Yoda and Y. Shiota. The mobile robot which passes a man. In *Proceedings of the IEEE International Conference on Robot and Human Communication*, pages 112–117, September 1997.

Coping strategies and technology in later life

M. V. Giuliani[*] M. Scopelliti[*†] F. Fornara[†]

[*] Institute of Cognitive Sciences and Technologies,
National Research Council,
Via Nomentana, 56, Rome, Italy

[†] Department of Social & Developmental Psychology,
University of Roma "La Sapienza"
Via dei Marsi, 78, Rome, Italy

v.giuliani@istc.cnr.it m.scopelliti@istc.cnr.it ferdinando.fornara@uniroma1.it

Abstract

The study presented in this paper aims at understanding to what extent elderly people are likely to accept a technological aid providing personal assistance in common everyday activities. Acceptability requirements seem to be the main concern for designers and producers. This does not refer only to physical and functional characteristics, but also to the overall integration between the technological devices and the psychological environment of the home, which is embedded in a variety of familiar behaviours and routines. In this perspective, the present research focused on strategies envisioned by elderly people as appropriate to cope with their diminished ability to perform everyday activities at home. The aim was to understand to what extent a technological device can be successfully applied to domestic tasks, and what everyday activities may fit such a strategy. The sample consisted of 123 elderly people living in Rome. We administered a questionnaire focussing on preferred strategies for performing common domestic tasks, and on attitudes towards new technologies and home modification. Results show that the adoption of a strategy, including the introduction of technological devices, is highly related to the specific problem being coped with, while personal factors are relevant only in specific situations. With increasing age, people are more inclined to give up, and higher educational levels correspond to more frequent technological solutions.

1 Introduction

In which areas of the everyday life of elderly people would a robot companion be more welcome?

A previous study, carried out as a part of the Robocare project (see Cesta *et al.'s* presentation in this conference) with the aim of assessing people's attitudes and preferences towards a domestic robot, revealed that the elderly have a conflicting view of such a device (Scopelliti, Giuliani, D'Amico & Fornara, 2004; Scopelliti, Giuliani & Fornara, forthcoming). Older people seem to recognise its potential usefulness in the house, but they are somewhat afraid of potential damages caused by the robot and of intrusion in their privacy. As regards physical shape and behaviour of the robot, they clearly express a preference towards a serious looking small robot, with a single cover colour and slow movements. In addition, most of them would like it not to be free to move inside the house and would expect it to be programmed in a fixed way to execute specific tasks. With respect to younger age groups, our old respondents do not appreciate the stimulating and intriguing side of an autonomous agent and tend to emphasise the practical benefits. However, when asked about the specific tasks the robot could perform in their home, people's answers are somewhat vague or unrealistic. In fact, robots are still too far away from everyday life to be easily distinguished from other technological aids, and the attitude towards them mirrors the attitude towards new technology in general.

The key point to be underlined is that the elderly do not show an a priori opposition to technological innovations, but that they are more likely to accept the change in everyday routines implied by the introduction of technological devices only when the practical benefits are evident. On the other hand, the assessment of benefits is not only related to the actual capability of a machine to perform a task, but also to the value people attribute to that task, and to the alternatives which are available.

Hence, an important aim is to understand what the deeper needs of elderly users are and the solutions envisioned to satisfy these needs. Ignoring these aspects would pose serious difficulties for the adoption of potentially useful devices.

In the present research, a wider approach was adopted, in which technological innovations are considered along a continuum, where a domestic robot is situated at the extreme pole, in that it can perform tasks with some degree of autonomy. Hence, the focus is on the characteristics of the situation in which a technological device is likely to be privileged with respect to other solutions in everyday domestic life, instead of on the characteristics of the technological device.

A central feature to emphasise is the relationship between adopted strategies, successful aging and life satisfaction. With reference to the theoretical model proposed by Brandtstadter & Renner (1990), two general coping strategies for maintaining life satisfaction are distinguished: the first is assimilation, involving active modification of the environment in order to reach personal goals; the second is accommodation, involving a more passive acceptance of life circumstances and obstacles, and a personal adaptation to the environment. Following this distinction, adaptive strategies can be put along a continuum ranging from the most assimilative to the most accommodative ones. Some studies (Wister, 1989; Brandtstadter & Renner, 1990) showed that old people tend to shift from assimilative to accommodative strategies as age increases. However, the use of both these strategies was positively related to life satisfaction.

A more comprehensive framework, grounded in environmental psychology, was provided by Slangen-de Kort, Midden & van Wagenberg (1998), who focused on the categorization of the activity that is adapted. Referring to daily domestic activities, a distinction between adaptation of the physical environment (i.e., modification of the home, use of assistive devices), the social environment (formal help, e.g., paid housekeeping, and informal help, e.g., help from friends), and the person him- or her-self (e.g., changes of behaviour, "give-up" reaction) was made.

Strategies of adaptation of the physical environment are considered the most assimilative; strategies of personal adaptation (particularly the "give-up" reaction) are categorized as the most accommodative ones.

Following this conceptual framework, the present study addresses a further issue, which is related to new technologies. More specifically, the use of technological devices is added as a specific assimilative choice for adapting the physical environment to personal needs. Furthermore, the investigation of the effect of increasing age on attitudes and behavioural intentions towards technology in general and in specific everyday situations is one of the objectives of this study.

2 The study

2.1 Objectives

This study aimed at finding answers to the following questions.

1) What are the main dimensions of elderly people's attitude towards new technolog ies?

2) Which personal (i.e., age, gender, educational level, income), psychological (i.e., perceived health, competence, openness to home changes), environmental (i.e., home safety and comfort) and situational (i.e., typology of problems) factors are more related to the choice of adaptation strategies in different situations?

3) Which personal, psychological and environmental factors are associated with attitudes and behavioural intentions towards changes in the domestic setting, referring to both spatial modifications in the environment and, more specifically, the introduction of technological devices?

2.2 Tools

We developed two different versions of a questionnaire, for male and female respondents. The questionnaire addressed several topics, and was organized in four sections.

a) The first section included a set of 8 scenarios, each of them describing an old person (a man in the male-version, a woman in the female-version) who finds difficulties in coping with a specific everyday situation. The eight situations are the following: 1 – Playing cards. Feeling unsafe to go to a friend's house to play cards; 2 – Telephone call. Having hearing difficulties in using the telephone; 3 – Medicine. Forgetting when to take daily medicines; 4 – Newspaper. Eyesight difficulties in reading; 5 – Cleanings. Housekeeping; 6 – Bathtub. Getting in and out the bathtub; 7 – Intruders. Fear of intrudersgetting into home; 8 – Home accidents. Feeling unsafe about accidents in the domestic setting.

Respondents were asked to suggest one possible solution to the problem to the scenario's actor, by choosing among different options representing adaptation strategies pertaining to the following macro-categories: 1) accommodation, i.e. give-up behaviour; 2) use of social resources, i.e. searching for either 2a) "formal help", from volunteers, health-care associations, paid assistant, etc., or 2b) ‘informal help”, from relatives, friends, neighbours, etc.; 3) adaptation of physical environment, either 3a) changing the spatio-physical setting, or 3b) using technological assistive devices.

The alternative solutions vary on a continuum from purely accommodative to purely assimilative, and follow a random order in each scenario response-set.

b) The second section included a set of 8 instrumental everyday activities. Only activities usually performed by both male and female elderly people were selected for assessment. Four of these activities require a cognitive effort (remembering to take a medicine, remembering to switch off the gas, managing money, keeping oneself well-informed about what’s happening in the world); the remaining four require a motion effort (house keeping or home maintenance, cutting toe nails, climbing or going down the stairs, kneeling or bending). The activities cover different problem/ability types, such as mnemonic functioning, performing complex cognitive tasks, homecare, self care, flexibility of body motion. For each target activity respondents are asked to assess: 1) their degree of autonomy on a dichotomous response scale (by oneself/help by others); 2) their ease of performing on a 5-step Likert-type response scale (from “not at all” to “very much”); 3) their overall satisfaction about the way the activity is performed on a 5-step Likert-type response scale (from “not at all” to “very much”). In addition, overall satisfaction towards health was measured on a 5-step Likert-type response scale.

c) The third section focused on the home environment. It included: 1) two short scales, respectively measuring perceived safeness and perceived comfort of home spaces (i.e., hall, kitchen, bathroom(s), bedroom(s), living room) through a 5-step Likert-type response scale; 2) a series of items measuring both attitudes and intentions towards possible home modifications (response scales are both dichotomous and Likert-type); 3) some items about attitudes and intentions towards technological modifications in the home.

d) The final section included a 5-step agree/disagree Likert-type Attitude Scale towards new technologies, borrowed from a previous study (Scopelliti *et al.*, 2004) and questions about socio-demographics (gender, age, education, income, housemates, etc.).

2.3 Sample and procedure

We contacted a sample of 123 elderly subjects, aged from 62 to 94 years (Mean= 74.7). The selection was made in order to cover a wide range of age in later life. Participants were people living in an urban environment, well balanced with respect to gender (M=61, F=62), age group (younger than 75 = 63, older than 75 = 60) and educational level.

The questionnaire was administered in a face-to-face interview. This procedure was adopted in order to overcome the difficulties of the majority of respondents with a pen-and-pencil survey.

2.4 Results

A Factor Analysis performed on the 12 items of the Attitude towards new technologies Scale yielded two independent dimensions, which explain 47.8% of the total variance. A Positive Attitude, summarizing the advantages provided by technologies (you don’t get tired, you don’t waste time, you can perform a lot of activities, you are not dependent on others, etc,), is opposed to a Negative Attitude, referring to a general uneasiness and a slight mistrust with technology (devices break down too often, instructions are difficult to understand, I do not trust, etc.). The two dimensions show a good internal consistency (Positive Attitude: Cronbach’s α = .80; Negative Attitude: Cronbach’s α = .69). The two dimensions proved to be coexistent aspects in elderly people’s representation, showing a somewhat ambivalent image of new technologies. Age, gender, income, and educational level did not show any significant difference with reference to both Positive and Negative Attitude towards new technologies.

How does this general attitude apply to everyday situations? A two-fold analysis of proposed scenarios was performed, in order to outline consistent cross-scenario strategies and to understand what the independent variables associated with specific behaviours in each situation are.

On the whole, the strategy of relying on technological interventions is one of the most preferred, second only to spatio-physical changes of the setting (Figure 1).

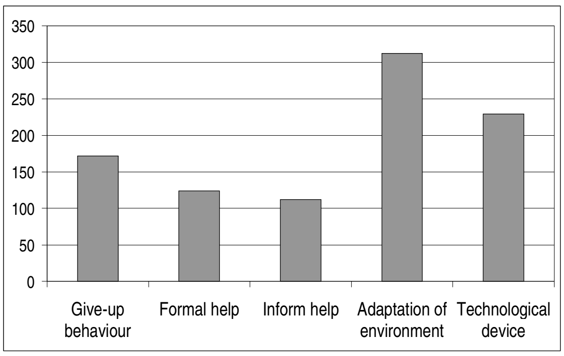


Figure 1. Distribution of strategies

The analysis of the relationships between individual variables and cross-scenarios strategies show that the only significant association is related to the opposition between the "give-up" strategy and the technological choice. The old elderly (over 74 years) tend to adopt a give up behaviour significantly more often than the young elderly (under 74 years) ($F_{(1, 121)}$ = 7.03, p<.01) and conversely the young elderly are more likely to rely on technological aids ($F_{(1, 121)}$ = 13.19, p<.001). A similar result emerges as regards educational level with the higher educated respondents relying on technology significantly more than less educated respondents, who are more likely to adopt a "give-up" behaviour ($F_{(2, 114)}$ = 5.22, p<.01). Socio economic level shows a significant effect on the chosen strategy, low-income respondents being more likely to adopt a "give-up" behaviour ($F_{(1, 113)}$ = 6.03, p<.05), while high-income people were more likely to ask for a formal help ($F_{(1, 113)}$ = 5.65, p<.05). Gender did not show any significant association with coping strategies.

By contrast, the relationship between strategies and scenarios is highly significant (Chi-square = 495.42, df = 28, p<.01), showing that the choice of a strategy is highly dependent on the situation. The "give-up" reaction is frequently adopted in the Playing Cards and Newspaper scenarios, rarely in the Bathtub, Cleanings and Intruders scenarios. The formal help is frequently required only in the Cleanings and Intruders scenarios. The informal help is a viable solution in the Playing Cards scenario alone. The technological strategy is very often indicated, emerging as ineffective only in the Playing Cards, Newspaper and Cleanings scenarios. The environmental change, as already mentioned, represents a relevant strategy across all different scenarios.

Some consistencies emerged as regards the scenarios categories. In fact, difficulties experienced in discretionary leisure activities, such as Playing cards and Newspaper reading, are most likely to generate accommodation, while in situations related to safety (such as Intruders and Home accidents scenarios) or health and personal care (Medicine and Bathtub scenarios) people usually strive to find an alternative solution, mainly based on changing the environment.

The analysis of the influence of independent variables on coping strategies in each scenarios shows a highly variable pattern.

Age was found to influence coping strategies in a couple of activities, namely in the Telephone call (Chi-square = 16.70, df = 5, p<.01) and in the Medicine (Chi-square = 9.88, df = 5, p<.05) scenarios. More specifically, it was found that the old elderly are more likely to give up in the Telephone call scenario than the young elderly; conversely, the young elderly are more likely to adopt the technological strategy (Figure 2). In the Telephone call scenario, the technological strategy consisted in a special device displaying the verbal communication on a monitor.

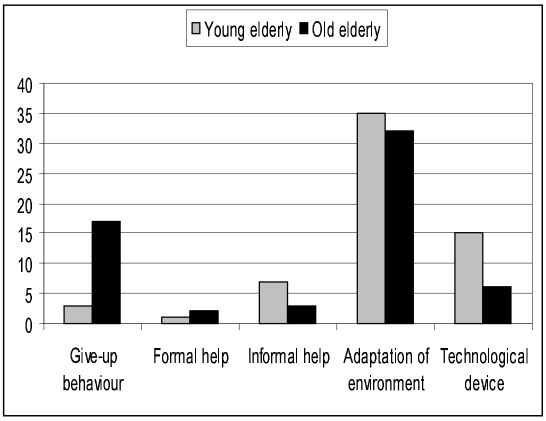


Figure 2: Telephone Call. Effect of Age

With respect to the Medicine scenario, again, the young elderly are more likely to adopt the technological strategy than the old elderly (Figure 3).

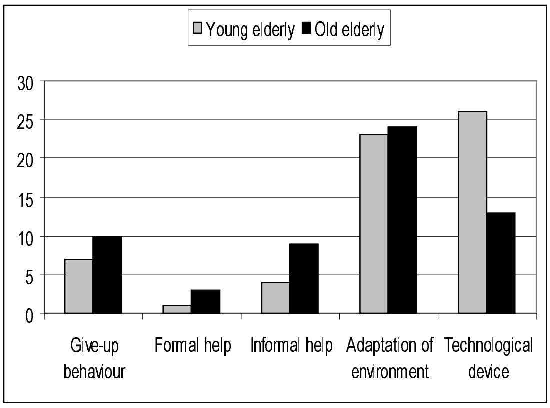


Figure 3: Medicine. Effect of Age

No *gender* differences were found in choosing strategies in the eight scenarios.

The *educational level*, as already mentioned, emerged as having a strong impact on adaptation strategies. The opposition between a "give-up" reaction and the choice of a technological strategy respectively by higher educated and lower educated respondents emerged in the Playing Cards (Chi-square = 22.32, df = 10, p<.05), Telephone call (Chi-square = 19.46, df = 10, p<.05), Medicine (Chi-square = 19.62, df = 10, p<.05), Intruders (Chi-square = 18.41, df = 10, p<.05) and Home accidents (Chi-square = 29.08, df = 10, p<.01) scenarios. Not surprisingly, in the Cleanings scenario, a higher educational level is positively associated (while a lower educational level is negatively associated) with the choice of a paid assistant for housekeeping activities.

Income was found to influence the adopted strategy in the Home Accidents scenario (Chi-square = 21.99, df = 4, p<.001). Low-income people are significantly more likely to show an accomodative solution (to relocate to their children's house) or to re-organise the domestic environment; high-income respondents are more willing to adopt the technological strategy (to install a tele-care system) (Figure 4).

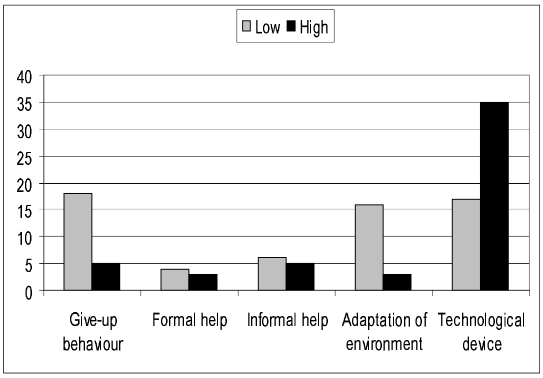


Figure 4: Medicine. Effect of Age

With reference to psychological variables we used a median-split on the health condition response in order to divide the sample in two groups, respectively with bad and good perception of overall personal health conditions.

On the whole, *perceived health* was not found to be an influential variable, with the exception of the Home accidents scenario (Chi-square = 11.22, df = 5, p<.05). In this situation, a worst perception of one's health is associated with a "give-up" reaction and a choice for relocation to the relatives' house, while a better perception would rather direct respondents towards the technological strategy (Figure 5).

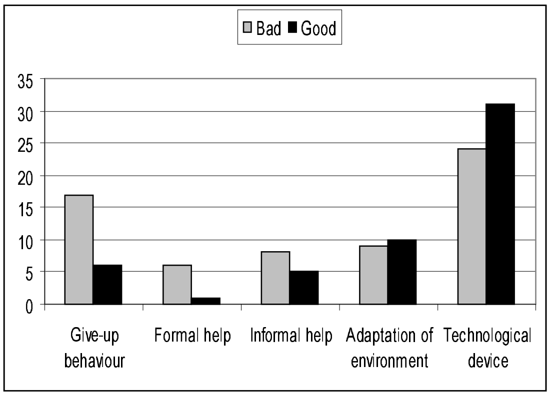


Figure 5: Home accidents.
Effect of Perceived Health

In respect to *perceived competence*, responses are heavily polarized towards the side of autonomy (above 90% of respondents answered being able to perform everyday domestic activities by themselves), leaving little room for a comparison by statistical analyses among people with different response to this variable.

The *attitude towards home modifications* was not found to be related to the chosen strategy. People with a positive and a negative attitude towards a change in the domestic setting did not show any significant difference in the strategies adopted in the eight proposed scenarios.

On the other hand, the *attitude towards a technological change* in the home setting emerged as a rather important variable in influencing the preferred strategy.

Both in the Telephone scenario, (Chi-square = 17.91, df = 5, p<.01) and in the Medicine scenario (Chi-square = 14.41, df = 5, p<.05) people with a negative attitude towards technological changes at home emerged as being significantly more likely to show a "give-up" reaction than people with a positive attitude. The latter, on the other hand, were shown to use the technological strategy more often than the former in the second scenario (Figure 6 and Figure 7).

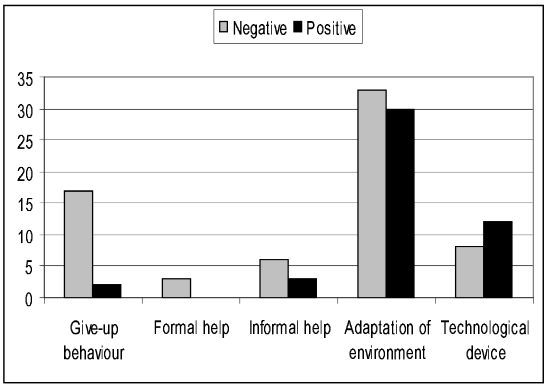


Figure 6: Telephone call.
Effect of Attitude towards technological change

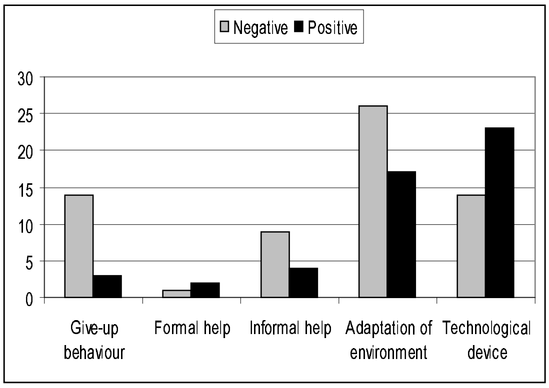


Figure 7: Medicine.
Effect of Attitude towards technological change

Finally, a significant effect was found in the Bathtub scenario (Chi-square = 11.41, df = 5, p<.05) (Figure 8). Again, a "give-up" was mainly shown by people with a negative attitude towards technological change.

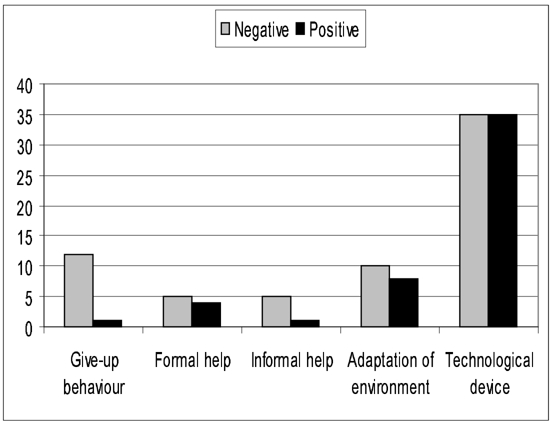


Figure 8: Bathtub.
Effect of Attitude towards technological change

Environmental variables (e.g. *perceived comfort and safety*) were also found to significantly influence the choice of a strategy in a few scenarios. In the Playing Card scenario (Figure 9), when respondents perceive their house as being uncomfortable, they are more likely to adopt a give-up strategy (Chi-square = 15.58, df = 5, p<.01).).

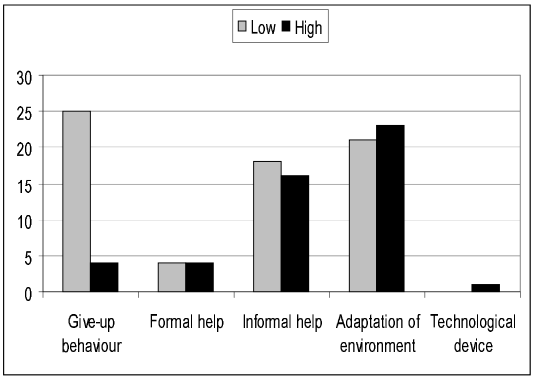


Figure 9: Playing Cards.
Effect of Perceived Comfort

In the Telephone call (Chi-square = 15.58, df = 6, p<.05) and Home accident (Chi-square = 11.46, df = 5, p<.05) scenarios, perceived comfort encourages the technological strategy, while low comfort is associated to giving-up or relying in informal help.

In a parallel way, a higher perception of safety is significantly associated (Chi-square = 13.31, df = 5, p<.05) with a preference for the technological strategy in the Home accident scenario (Figure 10). Interestingly, people considering their house as being somewhat unsafe show a moderate (even though not significant) tendency to decide for a relocation, and a significant preference for informal help (asking for a neighbour's assistance), which is probably perceived as a much more practical solution.

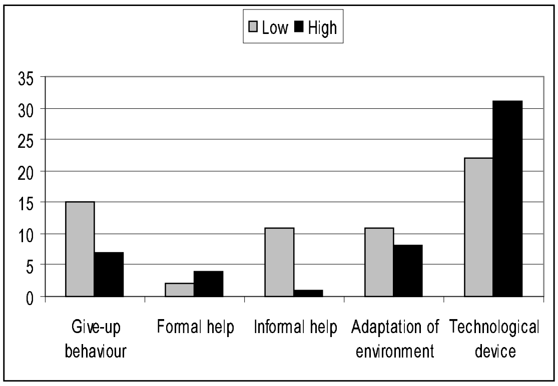


Figure 10: Home accidents.
Effect of Perceived Safety

In the Playing Card scenario the perception of potential risks favours the "give-up" behaviour, while the perception of safety encourages people to invite friends at home, instead of going out for playing (Chi-square = 21.60, df = 6, p<.01).

With respect to the third question, that is the *attitude towards home modifications*, also including technological devices, it is on the whole rather positive, and it is strongly anchored in practical considerations, which frequently emerged as fundamental aspects in elderly people's perspective (Scopelliti *et al.*, 2004, forthcoming). Neither gender ($F_{(1, 95)}$ = .46, n.s.) nor educational level ($F_{(2, 90)}$ = .03, n.s.), nor income ($F_{(1, 110)}$ = .01, n.s.), perceived health ($F_{(1, 95)}$ = .01, n.s.), home comfort ($F_{(1, 93)}$ = .30, n.s.), or safety ($F_{(1, 92)}$ = .11, n.s.) showed any significant association with the attitude. Conversely, a significant effect was found with respect to age, with the old elderly being much more reluctant to accept any kind of environmental change than the young elderly ($F_{(1, 95)}$ = 15.85, p<.001).

2.5 Discussion

In accordance with previous findings (Scopelliti *et al.*, 2004), the general attitude towards new technologies was found to be rather positive, with a homogeneous distribution of positive and negative evaluations across age, gender, and educational level groups. Conversely, significant differences between different age groups had emerged in previous research comparing young people, adults and elderly people (Scopelliti *et al.*, 2004). It is possible to argue that, for people over 65 years, age is no longer an important factor in shaping people's attitudes towards innovative technologies.

The more relevant finding in identifying acceptability requirements of domestic technology is that elderly people do not act in an idiosyncratic way for dealing with everyday problems at home. Instead, they choose the best solution depending on the situation. On the whole, assimilative strategies were shown to be a frequently chosen solution to cope with increasing difficulties in performing everyday tasks at home. The widespread stereotype that elderly people would be hostile to changes, both in general and even more with the introduction of technological devices turned no longer to be true.

The apparently contradictory result that people somewhat concerned about their home safety tend to adopt more accommodative strategies, could be explained with the hypothesis that people who consider their house as being safer have actually adopted some technological device inside their own domestic space, and they are aware of the benefits it can provide. So also in the proposed scenario they think that this can be the best option.

A tendency to exhibit a give-up reaction was found among the old elderly, showing that difficulties are often perceived as a normal condition they have to live with and passively accept. On the other hand, the young elderly more often try to find an adaptive solution to everyday problems, frequently relying on technology.

Interestingly, people prefer an accommodative strategy in the Medicine scenario, involving personal health, than in the Intruders scenario, involving safety. In the Home accidents scenario respondents do not give up the behaviour, but often ask for somebody else's (usually their relatives) control over it, by choosing to relocate. These results show a central concern for safety in elderly people's life. Social relationships can be a useful resource only for specific activities, like cleaning and playing cards. In the Cleanings scenario they ask for help from a formal and paid assistant, and this choice is in agreement with a common practice; in the Playing cards scenario, they ask for an informal help, showing the intrinsically social value of this activity. Home adaptations to personal needs emerged as a frequent strategy for overcoming problems, dispite some differences among situations. Technology is frequently accepted, particularly in the Telephone call scenario, because it involves hearing devices which seem to be extremely familiar to the elderly. Conversely, it is hardly chosen in the Cleaning scenario, for which human help seems to be more practical, and in the Playing cards scenario, because the social dimension of the activity would be hampered. In the Bathtub and Intruders scenarios environmental modifications (a special tub and an armoured door) are preferred to technology, probably because they are easier to install in the home.

The effect of educational level was found across different scenarios, so showing the key role of this variable in accounting for the choice of strategy, far more relevant than income. This result, which enriches previous findings about a relationship between negative emotional reaction to domestic robots and lower educational level (Scopelliti *et al.*, forthcoming), suggests that the possibility of controlling technological devices is an essential requirement for their acceptability. Low educated people are less confident in their ability to master a novel device. As a consequence, designers and producers have to consider that ease of use and adequate training are as much important as practical advantages provided by new technology.

Our results are compatible with a cohort effect, rather than a mere effect of aging, so we will need to pose the question of what attitudes towards technologies and environmental adaptation strategies elderly people will have in the near future. On the one hand, it is possible to use these findings in order to improve technological devices in the light of actual needs and behaviours of aging people; on the other hand, it seems to be necessary to merge together data from different age-groups in order to achieve an effective development of future technology. A central issue in this regard is to consider the psychological implications of whatever technological modification might be made in the home setting, implying both fixed and mobile devices such as robots. Moreover, a deeper knowledge on how people usually handle everyday activities and how they interact with the domestic setting, makes it easier to understand how new technology can fit the home environment.

Acknowledgements

This research is partially supported by MIUR (Italian Ministry of Education, University and Research) under project RoboCare (A Multi-Agent System with Intelligent Fixed and Mobile Robotic Components).

References

Brandtstadter, J. & Renner, G. (1990). Tenacious goal pursuit and flexible goal adjustment: Explications and age-related analysis of assimilation and accommodation strategies of coping. *Psychology and Aging*, 5, 58-67

Scopelliti, M., Giuliani, M. V., D'Amico, A. M. & Fornara, F. (2004). If I had a robot: Peoples' representation of domestic robots. In Keates, S., Clarkson, P. J., Langdon, P. M. & Robinson, P. (Eds.), *Design for a more inclusive world* (pp.257-266). London: Springer-Verlag.

Scopelliti, M., Giuliani, M. V. & Fornara, F. (forthcoming). Robots in a domestic setting: A psychological approach. *Universal Access in the Information Society*.

Slangen-de Kort, Y. A. W., Midden, C. J. H. & van Wagenberg, A. F. (1998). Predictors of the adaptive problem-solving of older persons in their homes. *Journal of Environmental Psychology*, 18, 187-197.

Wister, A. V. (1989). Environmental adaptation by persons on their later life. Research on Aging, 11, 267-291. Yiannis Aloimonos and Zoran Duri´c. Estimating the heading direction using normal flow. *International Journal of Computer Vision*, 13(1):33–56, 1994.

Communication Robots for Elementary Schools

Takayuki Kanda[*] Hiroshi Ishiguro[*†]
[*]ATR Intelligent Robotics and Communications Labs. [†]Osaka University
Kyoto, Japan Osaka, Japan
kanda@atr.jp ishiguro@ams.eng.osaka-u.ac.jp

Abstract

This paper reports our approaches and efforts for developing communication robots for elementary schools. In particular, we describe the fundamental mechanism of the interactive humanoid robot, Robovie, for interacting with multiple persons, maintaining relationships, and estimating social relationships among children. The developed robot Robovie was applied for two field experiments at elementary schools. The first experiment purpose using it as a peer tutor of foreign language education, and the second was purposed for establishing longitudinal relationships with children. We believe that these results demonstrate a positive perspective for the future possibility of realizing a communication robot that works in elementary schools.

1 Introduction

Recently, many researchers have been struggling to realize a *communication robot*, which is considered as a robot that participates in human daily life as a peer-type partner, communicates with humans as naturally as humans do by making bodily gestures and utterances, and supports humans with its communication tasks. Research activities toward communication robots have led to the development of several practical robots, such as therapy tools (Dautenhahn & Werry, 2002; Wada, Shibata, et al., 2004) and those for entertainment (Fujita 2001), and such robots are enlarging their working scope in our daily lives.

We believe that elementary schools are a promising field of work for a communication robot. The robot could be a playmate with children, although its interaction ability is limited in comparison to humans' and it would have very few social skills. As these fundamental abilities of robots improve, we can enhance their role: they will probably be useful for education support and understanding and building human relationships among children as friends. In future, it perhaps will help to maintain safety in the classroom such as by moderating bullying problems, stopping fights among children, and protecting them from intruders. That is, *communication robots for elementary schools* can be a good entry point for studying how robots participate in human daily life as peer-type partners.

We have developed a communication robot called Robovie that autonomously interacts with humans by making gestures and utterances as a child's free-play (Kanda et al., 2004 a); however, Robovie is confronted with three major problems in elementary schools: 1) difficulties in sensing in the real world, 2) difficulties in maintaining relationships with humans for long periods, and 3) difficulties in social interaction with many people.

We are addressing these problems via the following approaches. For the first problem, we believe that ubiquitous sensors are very helpful in reducing the burden of recognition in the real world. For example, with RFID tags (a kind of ubiquitous sensor) a robot can recognize individuals and call their names during interaction, which greatly promotes the interaction (Kanda et al., 2003). For the second problem, we have employed a design policy of interactive behaviors, such as a pseudo-learning mechanism and talking about personal matters (Kanda et al. 2004 b). For the third problem, we are trying to enhance its social skills. Currently, the robot identifies individuals to adapt its interactive behaviors to each of them (Kanda et al., 2003), and estimates human relationships by observing the humans' interaction around it (Kanda et al., 2004c).

The developed robot was used for two field experiments in elementary schools. We believe our experiments are novel as the first trials of applying interactive humanoid robots for human daily lives for a long period. The first experiment's purpose was to apply a robot to motivate Japanese children to study English (Kanda et al., 2004 d). The robot demonstrated positive effects for the motivating purpose; however, it is one interesting finding that

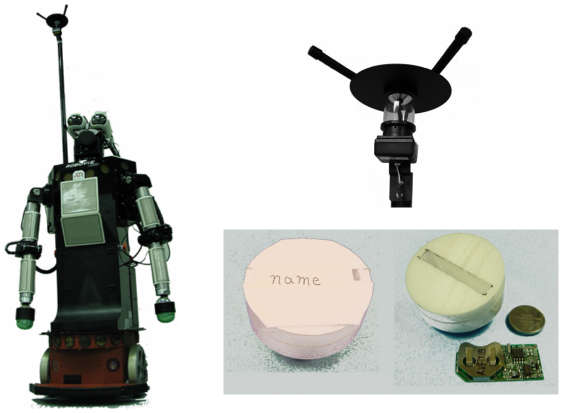


Figure 1: Robovie (left) and wireless tags

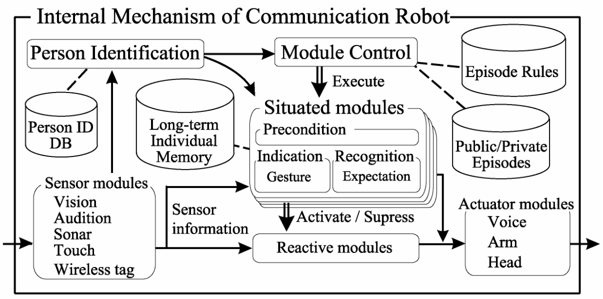


Figure 2: Software architecture of *Robovie*

the robot started to become boring after a week during the two weeks of the experiment's duration. The second experiment's purpose was to sustain long-term relationships between children and the robot, and a mechanism was added to the robot to assist long-term interaction (Kanda et al., 2004 b). As a result, it could maintain active interaction in a classroom for a few weeks, and sustained long-term relationships with some of the children for the two months of the experiment.

Meanwhile, we analyzed its performance regarding friendship estimation among children (Kanda et al., 2004 c) for these two experiments, finding better estimation for the children who interacted with it for a long time. That is, the robot's ability for long-term relationships seems to positively affect its estimation performance, and the estimated result may also promote the establishment of relationships with children.

Although the most parts of the three difficulties still remain as an open challenge, we are optimistic for the future of communication robots, because we believe that the difficulties will be gradually solved through the approach of field experiments. For example, the ability of communication robots for long-term interaction was improved between these two experiments in elementary schools, which seemed to demonstrate a positive perspective for this future direction. Namely, by placing robots in daily-life fields even with a limited task as a part of an experiment, the abilities lacked and problems faced will become clearer, enabling us to improve the fundamental abilities of robots.

2 Robot system

2.1 Robovie

Figure 1 shows the humanoid robot "Robovie." This robot is capable of human-like expression and recognizes individuals by using various actuators and sensors. Its body possesses highly articulated arms, eyes, and a head, which were designed to produce sufficient gestures to communicate effectively with humans. The sensory equipment includes auditory, tactile, ultrasonic, and vision sensors, which allow the robot to behave autonomously and to interact with humans. All processing and control systems, such as the computer and motor control hardware, are located inside the robot's body.

2.2 Person identification with wireless ID tags

To identify individuals, we used a wireless tag system capable of multi-person identification by partner robots. Recent RFID (radio frequency identification) technologies have enabled us to use contactless identification cards in practical situations. In this study, children were given easy-to-wear nameplates (5 cm in diameter), in which a wireless tag was embedded. A tag (Fig. 1, lower-right) periodically transmitted its ID to the reader installed on the robot. In turn, the reader relayed received IDs to the robot's software system. It was possible to adjust the reception range of the receiver's tag in real-time by software. The wireless tag system provided the robots with a robust means of identifying many children simultaneously. Consequently, the robots could show some human-like adaptation by recalling the interaction history of a given person.

2.3 Software Architecture

Figure 2 shows an outline of the software that enables the robot to simultaneously identify multiple persons and interact with them based on an individual memory for each person. Our approach includes non-verbal information on both robots and humans, which is completely different from linguistic dialog approaches. To supplement the current insufficient sensor-processing ability, we employed an active interaction policy, in which the robots initiate interaction to maintain communicative relationships with humans. The basic components of the system are *situated modules* and *episode rules*. *Module control* sequentially executes *situated modules* according to the current situation and execution orders defined by

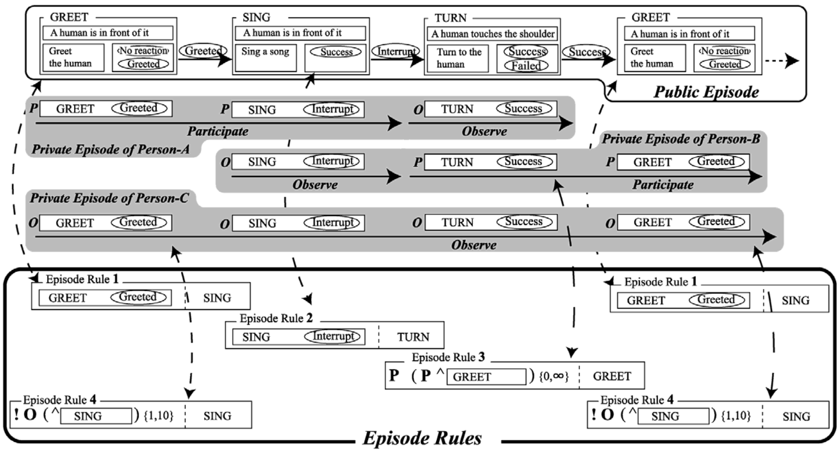


Figure 3: Illustrated example of episodes and episode rules for multiple persons

the *episode rules*. It is a completely bottom-up design that is quite different from others. Developers create *situated modules*, which execute a particular task in a particular situation, and *episode rules* that represent their partial execution order. The mechanism of interaction among humans is not yet known, so a top-down design approach is not yet possible.

The architecture includes four databases: *Person ID DB* to associate people with tag IDs, *episode rules* to control the execution order of *situated modules*, *public and private episodes* to sustain communications with each person, and *long-term individual memory* to memorize information about individual people. By employing these databases, the robot can track students' learning progress such as their previous answers to game-like questions.

The reactive modules handle emergencies in both movement and communication. For example, the robot gazes at the part of its body being touched by a human to indicate that it has noticed the touch, but continues talking. This hierarchical mechanism is similar to subsumption (Brooks 1986). In the situated and reactive modules, inputs from sensors are pre-processed by sensor modules such as English speech recognition. Actuator modules perform low-level control of actuators. In the following, we explain the *situated modules, person identification*, and *module control* in more detail.

Situated Modules.

As with an adjacency pair (a well-known term in linguistics for a unit of conversation such as "greeting and response" and "question and answer"), we assume that embodied communication forms by the principle of the action-reaction pair. This involves certain pairs of actions and reactions that also include non-verbal expressions. The continuation of the pairs forms the communication between humans and a robot.

Each situated module is designed for a certain action-reaction pair in a particular situation and consists of precondition, indication, and recognition parts. By executing the precondition part, the robot determines whether the situated module is in an executable situation. For example, the situated module that performs a handshake is executable when a human is in front of the robot. By executing the indication part, the robot interacts with humans. In the handshake module, the robot says "Let's shake hands" and offers its hand. The recognition part recognizes a human's reaction from a list of expected reactions. The handshake module can detect a handshake if a human touches its offered hand.

Person Identification.

Clark classified interacting people into two categories: participants, who speak and listen, and listeners, who listen only (Clark, 1996). Similar to Clark's work, we classify humans near the robot into two categories: participants and observers. The *person identification* module provides persons' identities, as well as their approximate distance from the robot. Since the robot is only capable of near-distance communication, we can classify a person's role in interaction based on his/her distance. As Hall discussed, there are several distance-based regions formed between talking humans (Hall, 1990). A distance of less than 1.2 m is "conversational," while a distance from 1.2 m to 3.5 m is "social." Our robot recognizes the nearest person within 1.2 m as the participant, while others located within a detectable distance of the wireless tag are observers.

Module Control (Episodes and Episode Rules)

We define an *episode* as a sequence of interactive behaviors taken on by the robot and humans. Internally, it is a sequence of *situated modules*. *Module control* selects the next *situated module* for execution by looking up *episodes* and *episode rules*. There are "public" and "private" *episodes* as shown in **Fig. 3**. The *public episode* is the sequence of all executed *situated modules*, and the *private episode* is an individual history for each person. By memorizing each person's history, the robot adaptively tailors its behaviors to the participating or observing persons.

The *episode rules* are very simple so that developers can easily implement many rules quickly. They guide the robot into a new episode of interaction and also give consistency to the robot's behaviors. When the robot ends an execution of *the* current *situated module*, all *episode rules* are checked to determine the most appropriate next *situated module*. Each *situated module* has a unique identifier called a ModuleID. The *episode rule* "*<ModuleID A=result_value>ModuleID B*" stands for "if module *A* results in *result_value*, the next execution will be module *B*." Then "<…><…>" stands for the sequence of previously executed *situated modules*. Similar to regular expressions, we can use selection, repetition, and negation as elements of *episode rules*.

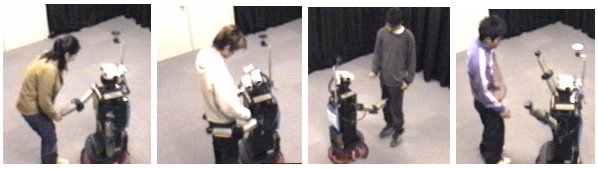
Furthermore, if "P" or "O" is put at the beginning of an *episode rule*, that *episode rule* refers to *private episodes* of the current participant or observers. Otherwise, the *episode rules* refer to *public episodes*. If the first character in the angle bracket is "P" or "O," this indicates that the person experienced it as a participant or an observer. Thus, "<P ModuleID=result value>" is a rule to represent that "if the person participated in the execution of ModuleID and it resulted in the result value." Omission of the first character means "if the person participated in or observed it."

Figure 3 is an example of *episodes* and *episode rules*. The robot memorizes the *public episode* and the *private episodes* corresponding to each person. *Episode rules* 1 and 2 refer to the *public episode*. More specifically, *episode rule* 1 realizes sequential transition: "if it is executing GREET and it results in Greeted, the robot will execute the situated module SING next." *Episode rule* 2 realizes reactive transition: "if a person touches the shoulder, the precondition of TURN is satisfied, after which the robot stops execution of SING to start TURN." Also, there are two *episode rules* that refer to *private episodes*. *Episode rule* 3 means that "if all modules in the current participant's *private episode* are not GREET, it will execute GREET next." Thus, the robot will greet this new participant. *Episode rule* 4 means "if the person hears a particular song from the robot once, the robot does not sing that song again for a while."

2.4 Implemented interactive behaviors

General design

The objective behind the design of Robovie is that it should communicate at a young child's level. One hundred interactive behaviors have been developed. Seventy of them are interactive behaviors such as shaking hands, hugging, playing paper-scissors-rock, exercising, greeting, kissing, singing, briefly conversing, and pointing to an object in the surroundings. Twenty are idle behaviors such as scratching the head or folding the arms, and the remaining 10 are moving-around behaviors. In total, the robot can utter more than 300 sentences and recognize about 50 words.



(a) shake hands (b) hug (c) paper-scissors-rock (d) exercise

Figure 4: Interactive behaviors

The interactive behaviors appeared in the following manner based on some simple rules. The robot sometimes triggered interaction with a child by saying "Let's play, touch me," and it exhibited idling or moving-around behaviors until the child responded; once the child reacted, it continued performing friendly behaviors as long as the child responded. When the child stopped reacting, the robot stopped the friendly behaviors, said "good bye," and restarted its idling or moving-around behaviors.

Design for long-term interaction

Moreover, we utilized the person identification functions to design interactive behavior for long-term interaction. The first idea was calling the children's names. In some interactive behaviors, the robot called a child's name if that child was at a certain distance. For instance, in an interactive behavior, the robot speaks "Hello, Yamada-kun, let's play together" when the child (named Yamada) came across to the robot. These behaviors were useful for encouraging the child to come and interact with the robot.

The second idea is pseudo-learning. The more a child interacts with the robot, the more types of interactive behavior it will show to the child. For example, it shows at most ten behaviors to a child who has never interacted with it, though it exhibits 100 behaviors to a child who has interacted with it for more than 180 minutes. Since the robot gradually changes interaction patterns along with each child's experience, the robot seems as if it learns something from the interaction. Such a pseudo-learning mechanism is often employed by the interactive pet robots like Aibo.

The third idea is having the robot confide personal-themed matters to children who have often interacted with it. We prepared a threshold of interacting time for each matter so that a child who played often with the robot would be motivated to further interact with the robot. The personal matters are comments such as "I like chattering" (the robot tells this to a child who has played with it for more than 120 minutes), "I don't like the cold" (180 minutes), "I like our class teacher" (420 minutes), "I like the Hanshin-Tigers (a baseball team)" (540 minutes).

3 Field experiments

This section reports on our previous field experiments with Robovie. The first experiment was purposed for foreign language education (Kanda et al., 2004 d), but there was no mechanism for long-term interaction implemented in Robovie. The second experiment was purposed for promoting longitudinal interaction with the mechanism for long-term interaction (Kanda et al. 2004 b). This section also reports on performances of friendship estimation (Kanda et al., 2004 c) among children through these two experiments.

3.1 First experiment: Peer tutor for foreign language education

3.1.1 Method

We performed two sessions of the experiment at an elementary school in Japan for two weeks. Subjects were the students of three sixth grade classes. There were 109 sixth grade students (11-12 years old, 53 male and 56 female). The session consisted of nine school days.

Two identical Robovie robots were placed in a corridor connecting the three classrooms, although there were no mechanisms for long-term interaction as reported in Section 2.4 implemented at that time. Children could freely interact with both robots during recesses. Each child had a nameplate with an embedded wireless tag so that each robot could identify the child during interaction.

3.1.2 Results

Since the results were reported in (Kanda et al. 2004 d) in detail, here we only briefly describe the results, with a particular focus on longitudinal interaction.

Results for Long-term Relationship

Figure 5 shows the changes in relationships among the children and the robots during the two weeks for the first-grade class. We can divide the two weeks into the following three phases: (a) first day, (b) first week (except first day), and (c) second week.

(a) First day: great excitement. On the first day, many children gathered around each robot. They pushed one another to gain a position in front of the robot, tried to touch the robot, and spoke to it in loud voices.

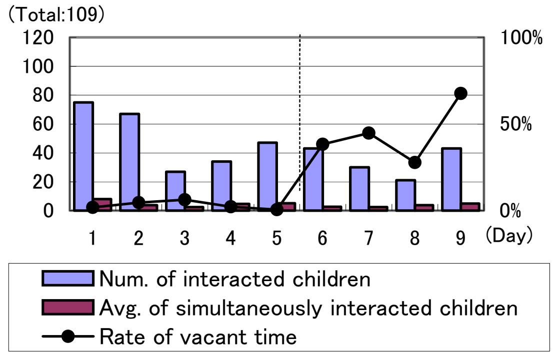


Figure 5: Transition of number of children playing with the robot (First experiment)

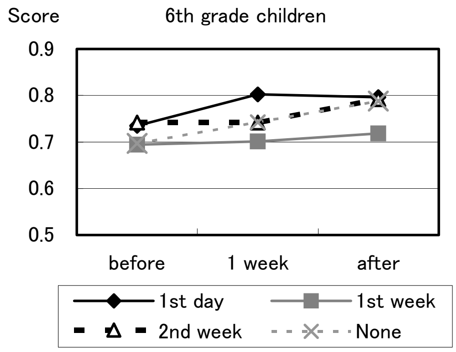


Figure 6: Improvement of children's English listening test scores (First experiment)

(b) First week: stable interaction. The excitement on the first day soon waned, and the average number of simultaneously interacting children gradually decreased. In the first week, someone was always interacting with the robots, so the rate of vacant time was still quite low. The interaction between the children and the robots became more like inter-human conversation. Several children came in front of the robot, touched it, and watched its response.

(c) Second week: satiation. It seemed that satiation occurred. At the beginning, the time of vacancy around the robots suddenly increased, and the number of children who played with the robots decreased. Near the end, there were no children around the robot during half of the daily experiment time. The way they played with the robots seemed similar to the play style in the first week. Thus, only the frequency of children playing with the robot decreased.

Results for Foreign Language Education

We conducted an English listening test three times (before, one week after, and two weeks after the beginning of the session). Each test quizzed the students with the same six easy daily sentences used by the robots: "Hello," "Bye," "Shake hands please," "I love you," "Let's play together," and

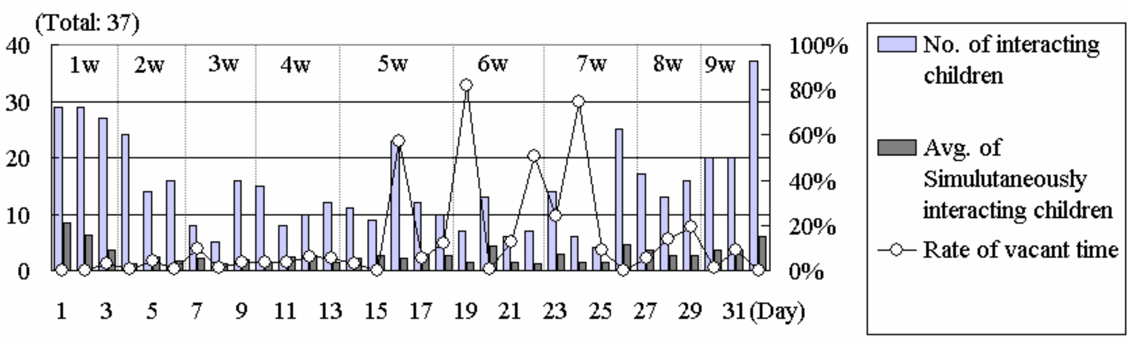


Figure 7: Transitions of the interaction between children and the robot (Second experiment)

(a) Beginning of the first day: Children formed a line

(b) showing nameplate

(c) "I can't see" behavior preferred

Figure 8: Scene of the second experiment

"This is the way I wash my face" (phrase from a song), but in different orders.

As a result, there were statistically significant improvements in their listening tests and the improvements were related to the interaction patterns of children (**Figure 6**: score represents the rate of correct answers in the listening test). Although the improvements were still quite low (less than 10% in the rate of correct answers), we believe that these results suggests a possibility of realizing a future communication robot that works in an elementary school and is equipped with a powerful language education ability.

3.2 Second experiment: Longitudinal interaction

3.2.1 Method

We performed an experiment at an elementary school in Japan for two months. Subjects were 37 students (10-11 years old, 18 male and 19 female) who belonged to a certain fifth-grade class. The experiment lasted for two months, including 32 experiment days. (There were 40 school days, but eight days were omitted because of school events.) We put the robot into a classroom, and the children were able to freely interact with it during the 30-minute recess after lunch time.

We asked the children to wear nameplates in which a wireless tag was embedded so that the robot could identify each child. The robot recorded the recognized tags during interaction to calculate each child's interacting time with it, which is used for later analysis of their interaction and friendship estimation. We also administered a questionnaire that asked the children' friendship with other children.

3.2.2 Results

Observation of Long-term Interaction

Figure 7 indicates the transition of interaction with children. The dotted lines separate the nine weeks during the two-month period. We classify the nine weeks into three principal phases and explain the interaction's transitions during those two months by describing these phases.

First phase (1st-2nd week): Robovie caused great excitement

Children were crowded around the robot on the first and second days (Fig. 8-a). During the first two weeks, it still seemed so novel to the children that someone always stayed around the robot, and the rate of vacant time was nearly 0, while the number of gathered children gradually decreased.

Second phase (3rd-7th week): Stable interaction to satiation

About ten children came around the robot every day, and some of them played with the robot. When it was raining, the children who usually played outside played with the robot, and, as a result, the number of children interacting with it increased. During these five weeks, the number of interacting children gradually decreased and vacant time increased. The "confiding of personal matters" behavior first appeared in the fourth week, with this behavior com-

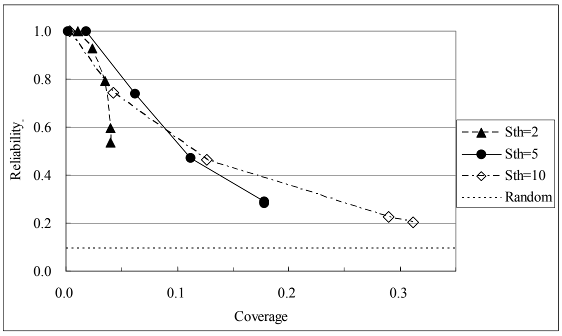


Figure 9: Friendship estimation results for the first experiment

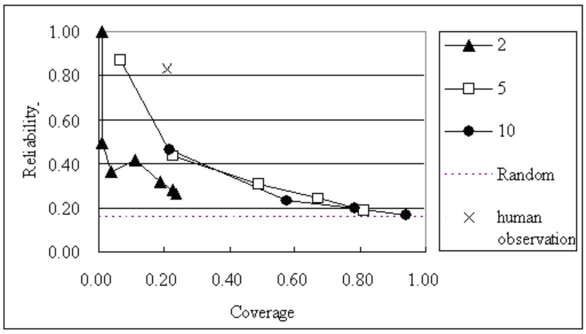


Figure 10: Friendship estimation results for the second experiment

ing into fashion among them. In this phase, we observed the following interesting scene.

- Child A observed the "confiding of personal matters" and told her friend, "the robot said that if you play with it for a long time, it will tell you a secret."
- Child B told the robot, "Please tell me your secret (personal matters)!"
- Although Child C asked the robot about personal matters, the robot did not reveal any. Child D was watching the scene and told child C the robot's personal matters that the robot had told child D before.

The robot gradually performed new behaviors according to the pseudo-learning mechanism, and these behaviors caught their attention.

- When the robot's eye was hidden (Fig. 8-c), it brushed off the obstacle and said "I can't see." This new behavior was so popular that many children tried to hide the robot's eyes.
- The robot started singing a song, and the observing children sang along with it.

Third phase (8th-9th week): Sorrow for parting

The number of children who came around the robot increased during these two weeks, though the number of children who played with the robot did not increase. Many of them simply came around and watched the interaction for a while. We believe that the teacher's suggestion affected them: on the first day of the eighth week, the class teacher told the students that the robot would leave the school at the end of the ninth week.

The "confiding of personal matters" behavior became well-known among the children, and many children around the robot were absorbed in asking the robot to tell these matters. They made a list of the personal matters they heard from the robot on the blackboard.

3.3 Friendship estimation

3.3.1 Method

We have proposed a method of friendship estimation by observing interaction among children via a robot (Kanda et al., 2004 c). This subsection briefly reports the method of estimation and estimation performances for these two field experiments.

Algorithm for reading friendly relationships

From a sensor (in this case, wireless ID tags and receiver), the robot constantly obtains the IDs (identifiers) of individuals who are in front of the robot. It continuously accumulates the interacting time of person *A* with the robot (T_A) and the time that person *A* and *B* simultaneously interact with the robot (T_{AB}, which is equivalent to T_{BA}). We define the estimated friendship from person *A* to *B* (*Friend*(*A*→*B*)) as:

$$Friend(A \rightarrow B) = if\,(T_{AB} / T_A > T_{TH}), \quad (1)$$

$$T_A = \Sigma\, if\,(observe(A) \text{ and } (S_t \le S_{TH}))\cdot \Delta t, \quad (2)$$

$$T_{AB} = \Sigma\, if\,(observe(A) \text{ and } observe(B) \text{ and } (S_t \le S_{TH}))\cdot \Delta t, \quad (3)$$

where *observe*(*A*) becomes true only when the robot observes the ID of person *A*, *if*() becomes 1 when the logical equation inside the bracket is true (otherwise 0), and T_{TH} is a threshold of simultaneous interaction time. We also prepared a threshold S_{TH}, and the robot only accumulates T_A and T_{AB} so that the number of persons simultaneously interacting at time t (S_t) is less than S_{TH} (Eqs. 2 and 3). In our trial, we set Δt to one second.

3.3.2 Results

Based on the mechanism proposed, we estimated friendly relationships among children from their interaction with the robot and analyzed how the estimation corresponds to real friendly relationships. Since the number of friendships among children was fairly small, we focused on the appropriateness (coverage and reliability) of the estimated relationships. We evaluated our estimation of friendship based on reliability and coverage, which are defined as follows.

Reliability = number of correct friendships in estimated friendships / number of estimated friendships

Coverage = number of correct friendships in estimated friendship / number of friendships from the questionnaire

Figures 9 and **10** indicate the results of estimation with various parameters (S_{TH} and T_{TH}) for these experiments. In the figures, *random* represents the reliability of random estimation, where we assume that all relationships are friendships (for example, since there are 212 correct friendships among 1,332 relationships, the estimation obtains 15.9% reliability with any coverage for later experiment). In other words, *random* indicates the lower boundary of estimation. Each of the other lines in the figure represents an estimation result with a different S_{TH}, which has several points corresponding to different T_{TH}. There is obviously a tradeoff between reliability and coverage, which is controlled by T_{TH} ; S_{TH} has a small effect on the tradeoff.

As a result, our method successfully estimated 5% of the friendship relationships with greater than 80% accuracy (at "S_{TH}=5") and 15% of them with nearly 50% accuracy (at "S_{TH}=10") for the first experiment (**Fig. 9**). It also successfully estimated 10% of the friendship relationships with nearly 80% accuracy and 30% of them with nearly 40% accuracy for the second experiment (**Fig. 10**).

4 Discussions and Future directions

4.1 Role of communication robots for elementary school

One promising role for communication robots in elementary schools is as a peer tutor. Through its interaction ability, Robovie had a positive effect as the peer tutor for foreign language education. However, current robots' abilities for interacting with humans are still very limited, strongly restricting the performances of the robot for language education task or other education tasks.

A more realistic role is currently behaving as a kind of friend with children and potentially bringing mental-support benefits, which is similar to a therapy robot (Wada, Shibata, et al., 2004). It is perhaps a substitution for pet animals but, what is more, we can design and control the robot's behavior so that it can more effectively produce benefits.

We believe that the mental-support role will be integrated into the robot's role as a peer tutor. For example, a communication robot might be able to *maintain safety* in a classroom. That is, the robot will be friend with children and, at the same time, report the problems such as bullying and fighting among children to the teacher so that teacher can change the robot's behavior to moderate the problems.

Figure 11: An interaction scene between children and a robot with a soft skin sensor

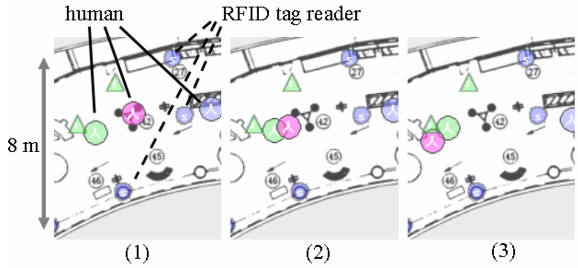


Figure 12: Position estimation with several RFID tag readers

4.2 Sensing in real world

It is difficult to prepare a robust sensing ability for robots in real world. Regarding Robovie, at elementary school, image processing and speech-recognition functions worked not as well as they did in the laboratory. Contrary, tactile sensing worked robustly. We believe that it is one of the most promising future directions, at least next several years, to use more tactile-based interaction for communication robots in elementary schools. **Figure 11** shows our robot with a soft skin sensor, which features a precise recognition capability in both spatial and temporal resolution on tactile sensing.

However, very limited information can be obtained only by using sensors attached to a robot's body. For example, a robot has difficulty in correctly identifying the person it is interacting with from among hundreds of candidates, which stands in contrast to robots being able to consistently recognize individuals with RFID tags (a kind of ubiquitous sensor). With RFID, the robot can call a child's name in interaction, which greatly promotes interaction.

Moreover, ubiquitous computing technology offers greater potential, in particular with sensors attached to an environment. **Figure 12** shows an example where humans' positions were recognized by using several RFID tag readers embedded in an environment. As those examples illustrate, it is important to make environments more intelligent so that a communication robots can behave as if it is more intelligent.

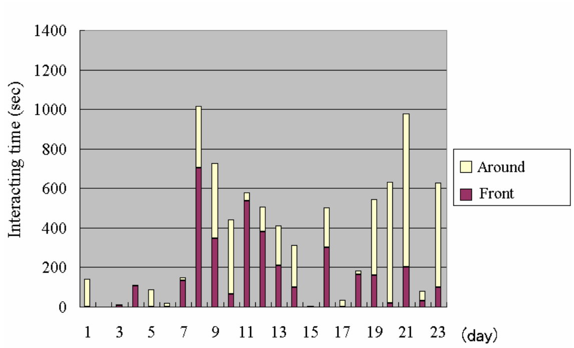


Figure 13: An example of analyzing longitudinal interaction (transition of a child's place)

4.3 Longitudinal interaction

Robots have a strong novelty effect. In other words, since robots are very novel for typical people, people are eager to interact with robots in the beginning, but rapidly become bored with them.

The second experiment indicated that the behavior design for longitudinal interaction (described in Sec. 2.4) contributed to keep children's interest longer. We believe that such an approach of adding contents (interactive behaviors) will be effective. However, this direction is gradually falling into the region of art rather than engineering.

There is other approach we should try: establishment of user models on longitudinal interaction. In these two experiments, we have observed three phases of interaction "great excitement," "stable interaction," and "saturation." If the robot can identify each person's phase of longitudinal interaction from sensory input, it can easily adjust its behavior to keep interaction more stable. For example, if a person is becoming saturated, it would exhibit some new behaviors. **Figure 13** is an example of analysis on long-term interaction. Here, we can see a change in the user's interaction patterns.

4.4 Social skills

Toward advancing the social skills of communication robots, we implemented two functions. One is that by identifying individuals around the robot, it alters its interactive behaviors to adapt to each person. The other is friendship estimation based on the observation of interaction among humans with RFID tags. Although current estimation performance is still quite low, we believe that we can improve it by using other sensory information, such as distance between people. This is one of our important future works.

Meanwhile, even with current performance, friendship estimation probably enables us to promote interaction between children and Robovie. For example, it would say "please take *child A* to play together," to *child B*, where *child A* and *B* are estimated as friends, thus it can make the interaction more enjoyable. Such positive relationships are rather easy to use.

On the contrary, it is difficult to identify negative relationships. For example, rejected and isolated children are identified by analyzing sociograms (a graph about social networks), which requires accurate estimation of relationships among all children. If more accurate estimation could be realized so that we can analyze sociograms based on the estimation, usage of such estimations could form the basis of interesting research themes on the social skills of communication robots. We believe that it will require a more interdisciplinary research, because in psychology, there is much knowledge about humans' strategy on communication, such as Heider's balance theory. At the same time, ethical problems should be more considered when robots start to estimate negative relationships or intervene in humans' relationships.

Research into the social skills of communication robots will be very important when the robots eventually participate in human society, and the functions we described here will be probably contribute a small part to developing social skills. We hope that there will be much research performed on this topic.

5 Conclusion

This paper reported our approaches and efforts to develop communication robots for elementary schools. The developed robot, Robovie, was applied to two field experiments at elementary schools. The result from the first experiment indicated that a communication robot will be able to support human activity with its communication abilities. The result from the second experiment indicated that we can promote longitudinal relationships with children by preparing some software mechanisms in the communication robot. In addition, the result from the friendship estimation indicated that a communication robot will be able to possess some social skills by observing human activities around it. We believe that these results demonstrate a positive perspective for the future possibility of realizing a communication robot that works in elementary schools.

Acknowledgements

We wish to thank the teachers and students at these two elementary schools where we conducted the field experiments for their agreeable participation and helpful suggestions. We also thank Prof. Naoki Saiwaki, Rumi Sato, Takayuki Hirano, and Daniel Eaton, who helped with these field experiments in the elementary schools. This research was supported in part by the National Institute of Information and Communications Technology of Japan, and also supported in part by the Ministry of Internal Affairs and Communications of Japan.

References

Dautenhahn, K., and Werry, I., A quantitative technique for analyzing robot-human interactions, *IEEE/RSJ International Conference on Intelligent Robots and Systems,* pp. 1132-1138, 2002.

Wada, K., Shibata, T., Saito, T., and Tanie, K., Effects of Robot-Assisted Activity for Elderly People and Nurses at a Day Service Center, *Proceedings of the IEEE*, Vol.92, No.11, pp. 1780-1788, 2004.

Fujita, M., AIBO; towards the era of digital creatures, *International Journal of Robotics Research, 20,* pp. 781-794, 2001.

Kanda, T., Hirano, T., Eaton, D., and Ishiguro, H., Person Identification and Interaction of Social Robots by Using Wireless Tags, *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS2003)*, pp. 1657-1664, 2003.

Kanda, T., Ishiguro, H., Imai, M., Ono, T., Development and Evaluation of Interactive Humanoid Robots, *Proceedings of the IEEE*, Vol.92, No.11, pp. 1839-1850, 2004 a.

Kanda, T., Sato, R., Saiwaki, N., and Ishiguro, H., Friendly social robot that understands human's friendly relationships, *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS2004)*, pp. 2215-2222, 2004 b.

Kanda, T., and Ishiguro, H., Friendship estimation model for social robots to understand human relationships, *IEEE International Workshop on Robot and Human Communication (RO-MAN2004)*, 2004 c.

Kanda, T., Hirano, T., Eaton, D., Ishiguro, H., Interactive Robots as Social Partners and Peer Tutors for Children: A Field Trial, *Journal of Human Computer Interaction*, Vol. 19, No. 1-2, pp. 61-84, 2004 d.

Brooks, R. A., A Robust Layered Control System for a Mobile Robot, *IEEE J. of Robotics and Automation*, 1986.

Clark, H. H., *Using Language*, Cambridge University Press, 1996.

Hall, E., *The Hidden Dimension*, Anchor Books/Doubleday, 1990.

My Gym Robot

Marti P.[*] Palma V.[*] Pollini A.[*] Rullo A.[*]
[*]Communication Science Dpt., University of Siena
Via dei Termini 6, 53100 Siena, Italy
marti@unisi.it
{palma,pollini,rullo}@media.unisi.it

Shibata T.[†]
[†] Intelligent Systems Institute, AIST, PRESTO, JST
1-1-1 Umezono, Tsukuba, Ibaraki, 305-8568 Japan
shibata-takanori@aist.go.jp

Abstract

The paper describes an experimental study to investigate the potential of robotic devices in enhancing physical rehabilitation. Nowadays "robotic active assist" therapy for movement recovery mainly works at the level of repetitive, voluntary movement attempts by the patient, and mechanical assistance by the robot. Our study describes the experience of physical rehabilitation of a 2 year child with severe cognitive and physical-functional delays. The therapeutic protocol consisted in movement recovery sessions performed with the help of a baby seal robot in which repetitive exercises were combined with cognitive tasks based on sensorial and emotive stimulation. The paper describes the results of the study and offers a reflection of possible future directions of robot assisted therapy.

1 Introduction

Several recent studies have suggested that robotic devices can enhance physical movement recovery in particular in stroke patients (Aisen et al., 1997; Volpe et al., 2000; Reinkensmeyer et al., 2000b). In these studies, patients have performed repetitive movement exercises with robotic devices attached to their limbs. The robotic devices have physically assisted in limbs movement using a variety of control approaches. This "robotic active assist" therapy has been shown to improve arm movement recovery in acute stroke patients (Burgar, et al., 2000) and chronic stroke patients (Reinkensmeyer et al., 2000a) according to coarse clinical scales and quantitative measures of strength and active range of motion. Despite these promising results, key research questions remain unanswered like: can robotic assistance in physical rehabilitation be used so that patients can assume an active and spontaneous role in the exercise? Can physical rehabilitation turn to be a stimulating and more engaging activity?

The robotic therapy methods used so far can be viewed as consisting of two key components: repetitive, voluntary movement attempts by the patient, and mechanical assistance by the robot. In our experiment we used Paro to introduce a more spontaneous and engaging activity and to stimulate the patient in the execution of coordinated movements. Furthermore, we complemented the objective of pure physiotherapy exercises with a human-robot interaction investigation, studying the potentiality of Paro to engage patient and to stimulate explorative behaviour (Dautenhahn et al., 2002; Dautenhahn and Werry 2004; Baron – Cohen et al., 1999; Baron-Cohen 2001).

The paper describes recent outcomes of a joint project conducted by the University of Siena, Italy, in collaboration with AIST (Advanced Industrial Science and Technology research), Tsukuba, Japan.

The project experimented a baby seal robot named Paro (Saito et al., 2003), for the treatment of young patients with various cognitive disabilities (Down syndrome, Autism, Angelman syndrome, Sensory-motor coordination etc.) (Marti et al., submitted). Paro has been developed to have physical interaction with human beings. Robot's appearance is from a baby of harp seal, with a white and soft fur. The seal robot has tactile, vision, audition, and posture sensors beneath the artificial fur and is able to exhibit three kinds of behaviors: proactive, reactive, and physiological. Pro-active behaviors are generated considering internal states, stimuli, desires, and a rhythm of the day. The basic behavioral patterns include some poses and some motions. The seal robot reacts to sudden stimulation like turning the head towards a source of sound and behaves following the rhythm of a day with some spontaneous desires such as sleep and tiredness.

2 Paro: a baby seal robot

Paro was designed by Shibata (Shibata et al., 2001) using a baby harp seal as a model (see Figure 1). Its surface is covered with pure white fur and its weight is around 2.8Kg. The robot is equipped with several sensors and actuators to determine its behaviour. As mentioned above, Paro has the appearance of a baby harp seal. Previous attempts to develop cat-robot and dog-robot (Shibata et al., 1999) demonstrated the inadequacy of these models in supporting interaction dynamics. The physical appearance of these robots turned out to be unsuccessful in meeting human being expectations during the interaction. The unlikeness from real cats and dogs was so evident to compromise any possibility of engagement with the robots. The baby seal model was therefore attempted.

Figure 1: The seal robot Paro

The choice was inspired by the idea to reproduce an unfamiliar animal that could barely create expectations in the human agent during the interaction. The design of Paro tried to balance the need to guarantee the likeliness with a real baby seal with the capability to stimulate exploration and sustain interaction. In this perspective a considerable effort was devoted to the design of eyes and gaze and all the facial expressions in general. The body is equally harmonious and balanced in all its parts.

In designing Paro, a particular attention was devoted to create an impressive tactile experience, a fundamental perceptual source of stimuli and information during the interaction (Woodward et al., 2001; Smith and Heise 1992). Its surface is covered with pure white and soft fur. Also, a newly-developed ubiquitous tactile sensor is inserted between the hard inner skeleton and the fur to create a soft, natural feel and to permit the measurement of human contact with Paro. The robot is equipped with the four primary senses: sight (light sensor), audition (determination of sound source direction and speech recognition), balance and the above-stated tactile sense. Its moving parts are as follows: vertical and horizontal neck movements, front and rear paddle movements and independent movement of each eyelid, which is important for creating facial expressions.

The combination of these technical features provides the robot with the possibility to react to sudden stimulation. For example, after a sudden loud sound, Paro turns the head in the direction of the sound.

Along with the reactive behaviour described above, Paro has a proactive-behaviour generation system consisting of two different layers: a Behaviour-Planning layer and a Behaviour-Generation layer. By addressing it's internal states of stimuli, desires, and rhythms, Paro generates proactive behavior.

Behavior-planning layer

This consists of a state transition network based on the internal states of Paro and its desires, produced by its internal rhythm. Paro has internal states that can be named with words indicating emotions. Each state has numerical level which is changed by stimulation. The state also decays in time. Interaction changes the internal states and creates the character of Paro. The behavior-planning layer sends basic behavioral patterns to behavior-generation layer. The basic behavioral patterns include several poses and movements. Here, although the term "proactive" is used, the proactive behavior is very primitive compared with that of human beings. Paro's behavior has been implemented similar to that of a real seal.

Behavior generation layer

This layer generates control references for each actuator to perform the determined behavior. The control reference depends on magnitude of the internal states and their variation. For example, parameters can change the speed of movement and the number of instances of the same behavior. Therefore, although the number of basic patterns is finite, the number of emerging behaviors is infinite because of the varying number of parameters. This creates life-like behavior. This function contributes to the behavioral situation of Paro, and makes it difficult for a subject to predict Paro's action.

The behavior-generation layer implemented in Paro adjusts parameters of priority of reactive behaviors and proactive behaviors based on the magnitude of the internal states. This makes the robot's behaviour appropriate to the context, being able to alternate reactions to external stimuli and generation of behaviours for gaining attention. Moreover, Paro has a physiological behaviour based on diurnal rhythm. It has several spontaneous needs, such as sleep, that affects its internal states and, consequently, the perceived behaviour.

In order to keep traces of the previous interactions and to exhibit a coherent behavior, Paro has a function of reinforcement learning. It has positive value on preferred stimulation such as stroking. It also has negative value on undesired stimulation such as beating. Paro assigns values to the relationship between stimulation and behavior. Gradually, Paro can be tuned to preferred behaviors. Eventually, the technical features allow Paro to engage distant interactions, in this being aware of contextual information.

3 The clinic case

Paro has been recently tried out as a therapeutic tool in non pharmacological protocols. The paper reports about one of a series of experimental studies conducted in the Rehabilitation Unit at ‘Le Scotte’ Hospital in Siena, to assess the efficacy of such a tool in complementing existing therapeutic protocols. In particular we describe the case of JC, a 2 year old child with severe cognitive and physical-functional delays. Till now, the child has not received a clear diagnosis: from his birth he had a clear delay in cognitive and physical development in particular language was not present at any level, even in pre-verbal production. The origin of his delay has to be attributed to a genetic disorder not better identified yet. Thus, all the descriptions about his pathology are basically a sum of the symptoms he showed during the treatment.

At the beginning of our study, JC was not able of coordinate movements, he had difficulties in trunk muscles control and when he sat down he was only able to slowly turn his bust and body even if never in complete rotation. So whenever present, these movements were dim and unsteady. JC had a particular head conformation with small ears close to the head, and squint and divergent eyes. Other relevant aspects were physical weakness and easiness to fall ill. He was often ill with temperature and headaches and this allowed only spotted therapeutic intervention. JC was able to produce few sounds similar to crying or mumbling but not to vocalize. He could articulate several facial expressions from vexation to pleasure.

The most severe aspects of his developmental delay were the physical impairment and motor coordination. JC was not able to perform particular gestures such as clapping because his left hand was slower than the right one. As for complex physical motor abilities he was not able to walk on all fours since he could not get down on hands and knees and walk or move. He seemed clearly to feel comfortable when lying even if grovelling was a considerable effort for him since arms and legs had to be coordinated.

In previous therapy he was mainly trained to catch and release objects, and every exercise was done involving the right side of his body first and only when the movement was completely acquired and understood the therapist referred to the left side, the most impaired one, but with scarce results.

The objective of our experimental study was to investigate the role of Paro in complementing the therapeutic treatment in the acquisition of basic motor-functional schemas like autonomously sitting and lying by controlling the body. In particular we were interested in investigating if the activity with the robot could support physical coordination and balancing attainment, and how postures control might evolve or vary along the sessions. However, since previous training was very similar to physiotherapy exercises concentrated only on the repetition of motor routines, we aimed to explore other aspects of child-robot interaction in particular the role of Paro in stimulating interest and engage the child during the activity.

3.1 Previous treatment: the Bobath method

JC was previously treated following the Bobath method, a consolidate concept developed by the early 1950’s to address neurological disability came into being, most of which had a theoretical basis in neurophysiology. Indeed the post war era brought an increasing awareness of the need for rehabilitation, leading a burgeoning of interest in all aspects of rehabilitation. The Bobath is a system of therapeutic exercises designed to inhibit spasticity and to aid in the development of new reflex responses and equilibrium reactions. The exercises are performed by modifying postures that progress from simple movements to more complex ones in a sequence based on the neurological development of the infant.

The main aim of treatment is to encourage and increase the child's ability to move and function in as normal way as possible. More normal movements cannot be obtained if the child stays in few positions and moves in a limited or disordered way. Therefore it is fundamental to help the child to change his abnormal postures and movements so that he or she is able to comfortably adapt to the environment and develop a better quality of functional skills. The method consists of training the child to acquire key behavioural patterns of movement and positioning like:

- head control
- grasping
- “parachute reactions”, that is the capability to protect himself in case of fall
- trunk turning

- equilibrium reaction in case of fall
- equilibrium control during the movement.

JC was treated following the Bobath concept for one year, two times a week lasting one hour each. The treatment was performed by physically and cognitively stimulating the child using toys and coloured pillows. The treatment aimed at supporting the development of motor and postural basic schemas through specific body movements facilitated by the therapists.

Figure 2 shows some exercises and positions simulated with a doll. These are similar to physiotherapy exercises in which motor routines are privileged to cognitive or symbolic therapy.

Figure 2: Bobath method in practice

Usually when adopting the Bobath method, the therapists use several objects to engage the patient in the activity, like balls to favour grasping and throwing, or sound objects to attract the child attention. Also JC was treated following the same protocol, but after one year the child was still acquiring the first autonomous basic motor-functional schemas like sitting and lying controlling the body. More in detail, after the treatment the child exhibited the following behaviour:

- Asymmetric postural behaviour characterized by a more reactive and developed right side of the body in respect to the left one.
- As for complex motor patterns JC was still not able to walk on all fours. Each time the therapist tried to support this activity JC remained still without taking part to any proposed tasks. When on all fours as in crawling, he was only able to control the gaze direction through head movements.
- When lying JC was not able to move his body but only to turn the head unintentionally.
- When supine he was only able to extend and flex the legs together without controlling them independently.
- He could turn the trunk without being able to stay lateral.
- The exploration of the surrounding environment was still based on the oral experience of putting objects close to the mouth without being able of direct manipulation (e.g. grasping). Controlled catching and releasing skills like releasing objects of different size were not present.
- He did not show any interest in the surrounding environment including people.
- He remained seated only when supported by pillows or by the therapist.

4 The experimental study

Our experimental study stems from a recent tradition of research in robot assisted therapy. A significant part of this field of research focuses on the use of robot as therapeutic tool for autistic children. A pioneering study was carried out by Weir and Emanuel (1976) who used a remote-controlled robot as a remedial device for a seven year old boy with autism. More recently, other researches made use of robots for rehabilitation (Plaisant et al., 2000) and more specifically with autistic children (Michaud et al., 2000; Michaud and Théberge-Turmel, 2002; Dautenhahn et al., 2002; Dautenhahn and Werry, 2004). Francois Michaud and his team at Université de Sherbrooke developed different typologies of mobile autonomous robotic toys with the aim of supporting autistic children in the acquisition of social and communication basic skills. Bumby and Dautenhahn (1999) investigated the modality by which people (especially children) may perceive robots and what kind of behaviour they may exhibit when interacting with them. They analysed human-robot interaction and the potential of robotic devices as therapeutic support (Dautenhahn and Billard, 2002).

The idea of our study came basically from three considerations:

- The encouraging results of previous studies on robot assisted therapy (Reinkensmeyer et al., 2000b; Volpe et al., 2000; Dautenhahn and Werry, 2004) and pet therapy (Boldt and Dellmann-Jenkins, 1992; Lago et al., 1989; Friedmann et al., 1980).
- The idea of designing engaging rehabilitation activities that combine physical and cognitive rehabilitation (a specific requirement raised from interviews and focus groups with therapists).
- The specific characteristics of Paro to impart a positive mental effect on humans, such as joy and relaxation through physical interaction, touch and spontaneous actions (Saito et al., 2003).

In fact Paro is characterized by having "agentivity cues" that are physical features and behaviours that favour the attribution of intentions to the robot. These are basically physical, perceptual and behavioural features.

Physical and perceptual cues of agentivity in Paro include morphology and texture. Designers put a considerable effort in designing eyes, gaze and the facial expressions in general. As for the texture, Paro gives an impressive tactile experience when stroked. Its surface is covered with pure white and soft fur and tactile sensors are inserted between the hard inner skeleton and the fur to create a soft, natural feel and to permit the measurement of human contact with Paro. The behavioural cues of Paro include eye direction, head turn, and self-propelled motion, cues that infants select and detect when reproduced or simulated by an agent during the interaction. Paro has sight, audition, balance and tactile sense, it is also able of vertical and horizontal neck movements, front and rear paddle movements and independent movement of each eyelid, which is important for creating facial expressions.

Furthermore the proactive-behaviour generation system creates a life-like behaviour of the robot and makes it difficult for a subject to predict Paro's action. Another feature that distinguishes agents from non-agents is the ability to engage in contingent and reciprocal interactions with other agents (Johnson, 2000). The behavior-generation layer implemented in Paro makes the robot's behaviour appropriate to the context, being able to alternate reactions to external stimuli and generation of behaviours for gaining attention. Moreover, the physiological behaviour based on diurnal rhythm generates several spontaneous needs, such as sleep, whilst the function of reinforcement learning on preferred or undesired stimulation allows to gradually tune Paro to preferred behaviors and eventually engage distant interactions, in this being aware of contextual information.

All these characteristics made of Paro an extremely interesting candidate for our study. During brainstorming sessions with the therapists, the need for engaging the patient at different levels, physical and emotional, was a basic requirement. The control of emotional reactions and the exhibition of consistent behaviours were also indicated as a first step for improving the child capabilities.

From these considerations we designed an experimental study based on two main hypotheses. We postulated a positive effect of Paro in sustaining JC in the acquisition of basic skills in the following areas:

- *Physical-functional area* of motor basic movements control, such as:
  - Prone to lateral turning
  - Supine to prone turning (or vice versa)
  - Autonomously sitting down
  - Sitting down with external aid
  - Kneeling
  - Sitting down to lying
  - Clapping hands

  and *postural patterns* mainly equilibrium control and basic postural patterns acquisition.
- *Engagement area* related to the control of emotional expressions of the child and attention on the robot during the activity. Smiling, crying and attention are meaningful manifestations of engagement.

It is important to highlight that the objectives related to the two mentioned areas were all tried out during the treatment with the Bobath method (even if in a non systematic way indeed quantitative data were never collected). Unfortunately these objectives turned to be unsuccessful as shown in the description of the child conditions at the beginning of the study reported in session 3.1.

4.1 The methodology

The experimental study was conducted over a period of three months with a weekly occurrence as in the current therapeutic protocol. It was articulated into six sessions, each one lasting about one hour (depending on the health status of the child) and conducted under two different conditions:

- *Paro-passive*, in which the robot was used as a stuffed puppet. This condition reproduced the previous sessions in which the Bobath method was applied with the support of a toy (in the following charts we will refer to this condition as "Session OFF").
- *Paro-interactive*, in which the robot was switched on and so fully operational (we will refer to these with "Session ON"). This condition was used to compare the behavioural characteristics of Paro with those of a stuffed puppet, like in the case of the previous treatment.

The activities were designed to be as much similar as possible to the exercises performed following the Bobath method. Whereas in the previous treatment pillows and toys were used to support the exercises, in our experiments these were substituted by Paro in the two conditions.

The experimental sessions were organised in two groups:

- 2 sessions under the Paro-passive condition alternating free exploration and rehabilitation exercises.
- 4 sessions (a 2+2 cycle of iterative sessions) under the Paro-interactive condition

characterized by free exploration and rehabilitation exercises.

The exercises took place mainly on a foam–rubber mattress inside the rehabilitation room and the therapist used a wooden bench to support JC in standing up (Figure 3). All sessions were video-recorded.

Figure 3: some rehabilitation exercises performed with Paro

Some therapists were involved in the definition of indicators for quantitative measures and in their interpretation during the data analysis. We defined observation grids based on a set of micro-behaviours (Dautenhahn et al., 2001; Camaioni et al., 1997, Zappella, 1990) to collect quantitative data about the occurrence of each indicator during the activity. Tables 1 and 2 below contain extracts of indicators (micro-behaviours) related to the physical and engagement areas.

Table 1: example of micro-behaviours related to the physical-functional area

Physical-Functional Area	
Postures control	Leaning toward a given object
	Leaning to hold his parents
	Sitting down (head and arms balancing)
	Staying on all fours (head and trunk balancing)
	Lying (movements coordination)
	Prone to lateral turning
	Supine to prone turning (or vice versa)
	[…]

Table 2: example of micro-behaviours related to the engagement area

Engagement Area	
Emotional reactions	Surprise
	Fear
	Impatience
	Avoidance
Attention (following PARO's gaze)	with eyes
	turning the head
	turning the body
Crying to	The presence of the robot
	Robot behaviours
Smiling / laughing at	the presence of the robot
	robot behaviours
	Clapping his hands to show joy
	[…]

4.2 Results

The data analysis combined a quantitative and qualitative approach (Dautenhahn et al., 2002). Video analysis and semi-structured interviews with therapists were used to collect and analyse quantitative and qualitative information. The final results emerged from the combination of these two methodologies. For example, the quantitative data were commented by the therapists who helped us in interpreting ambiguous or unclear video sequences.

This kind of analysis was extremely useful to elaborate meaningful correlations between results of different micro-behaviours. The micro-behaviours were analysed in sequences of 10 seconds measuring occurrences of meaningful events contained in the grid.

In the following the results are presented in relation to the two areas physical-functional and engagement.

4.2.1 Physical-functional area

Motor basic movements

The video-analysis mainly focused on dodging, turning and sitting movements to address the emergent behavioural responses JC exhibited all along the experimental sessions.

Figure 4 shows the trend of body functional movements of JC. This label includes all micro-behaviours of the physical-functional area.

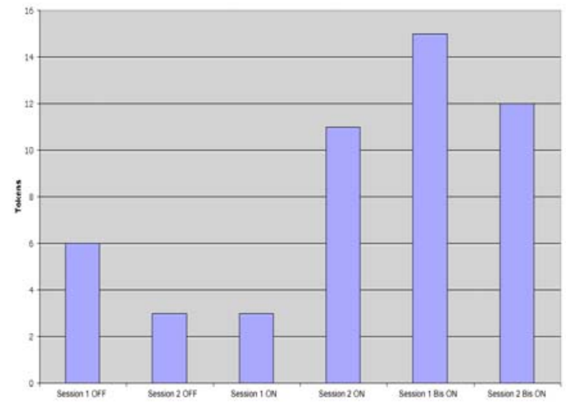


Figure 4: trend of body functional movements

As shown in the graph, the number of body movements performed by JC increased in the Paro active condition. The therapist interpreted the third session as a reflection period JC took to “study” the robot before engaging any interaction. She considered this behaviour as a symptom of cognitive development in the child.

Postural patterns

In children with neurological delay, postural patterns are strictly related to the acquisition of equilibrium and to the capability of manipulating objects maintaining a balanced posture.

For JC the manipulation of an object in a balanced position was a difficult task. Indeed he neither had a correct posture of the back muscles nor the ability to get balanced moving the weight of his body onto the basin.

In the previous treatment with the Bobath Method supported by toys/pillows, JC was able to maintain the equilibrium but the manipulation of objects in a balanced position was never reached. With Paro the child showed an improvement of this capability.

In data analysis we combined the indicators related to the manipulation of the robot with those of the physical-functional area and calculated the co-occurrence of manipulation events with the time in which JC was able to maintain a balanced position.

We introduced the correlation coefficient to determine the relationship between the two sets of indicators. In order to clarify how the correlation coefficient was calculated, we can take as an example the average temperature of a location and the use of air conditioners. It is possible to examine the relationship between these items determining the correlation coefficient of the array1 and array2, where

Array1 (x) is a cell range of values

Array2 (y) is a second cell range of values

and CORREL ρ (array1,array2)

In calculating the correlation coefficient between the two aggregated data, manipulation events and balance position time, we applied the following correlation coefficient equation:

$$\rho_{x,y} = \frac{\text{Cov}(X, Y)}{\sigma_x * \sigma_y}$$

where : $-1 \leq \sigma_{xy} \leq 1$

and : $\text{Cov}(X, Y) = -\frac{1}{n} \sum_{j=1} (x_j - \sigma_x)(y_j - \sigma_y)$

A positive value of the correlation coefficient (within 0 and +1) was found between the two aggregated data.

Manipulation and balancing had a strong correlation when the robot was switched on: in fact the correlation coefficient was nearly to + 1, while in the Paro passive condition the correlation coefficient was nearly zero.

More in detail the correlation coefficient in the Paro passive condition was: -0,0087487

Whilst in the *Paro* interactive condition was: 0,76513824

Figure 5 shows the evolution of the correlation along the six experimental sessions.

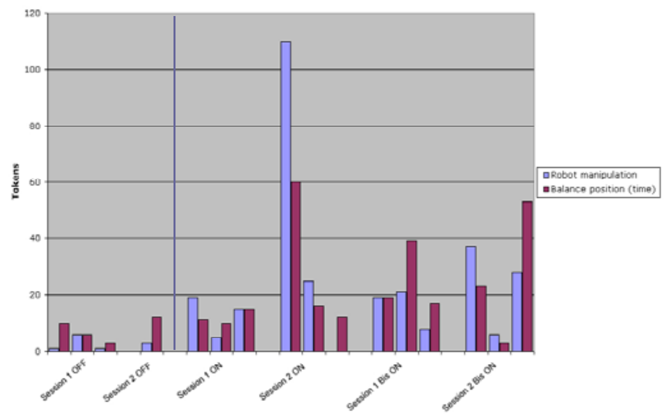


Figure 5: trend of correlation between manipulation and balance

Even for manipulation/ balance, the third session seemed to reproduce the same “observation period” we detected for the basic movements acquisition.

The fourth session presents a peak in the activity, but it is important to notice that in the following sessions the robot manipulation happened when the child was able to stay in the balanced position for a longer period. This means that once JC was able to sit autonomously long enough, he was also able to manipulate the robot. The therapist interpreted this as an acquired new skill of JC that gave him the opportunity to observe and explore the robot. In this case the improvement of posture control corresponded to an improvement of manipulation skills.

4.2.2 Engagement area

As explained above the design of engaging rehabilitation experiences was one of the main motivations of our study. The control of attention, his emotional reactions and the exhibition of consistent behaviours were considered a first step for improving the child capabilities.

Attention
Beside the emotional reactions, the Bobath method suggests to consider the attention as a meaningful indicator of engagement. Indeed keeping attention presupposes a strong motivation in establishing a relationship with an object or an agent.
Attention on the robot was measured in relation to the occurrence of three micro-behaviours:

- following Paro's gaze with the eyes,
- following Paro's gaze turning the head, that presupposes coordination between eyes and neck muscles and a certain motivation in following Paro's actions.
- following Paro's gaze turning the body. This task presupposes a strong motivation to observe, discover and find a target.

Figure 6 shows the occurrence of these micro-behaviours in the six experimental sessions.

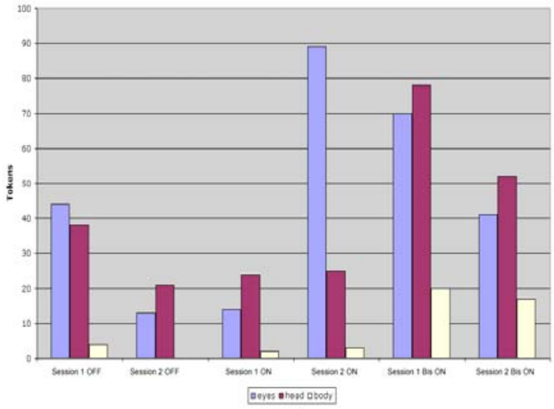


Figure 6: trend of attention behaviours

In the first two sessions, JC seemed to be quite curious to look at Paro's gaze turning the eyes and the head but not enough to move the body.
The third session, even if quite similar to the second, was interpreted by the therapist as an observation period JC took to "study" the robot and get familiar with him. The following sessions JC was much more reactive and in particular in the last two he was able to move his body toward the robot up to twenty times in the same session.
It's important to notice that following Paro's gaze moving the body implies the activation and coordination of motor and cognitive skills since this behaviour is triggered by the intention of reaching and manipulating an object.

Emotional reactions
Smiling, crying, expressions of joy and clapping hands were considered meaningful manifestations of engagement, in particular clapping hands that was a very difficult task for JC.
Figure 7 shows the related outcomes.

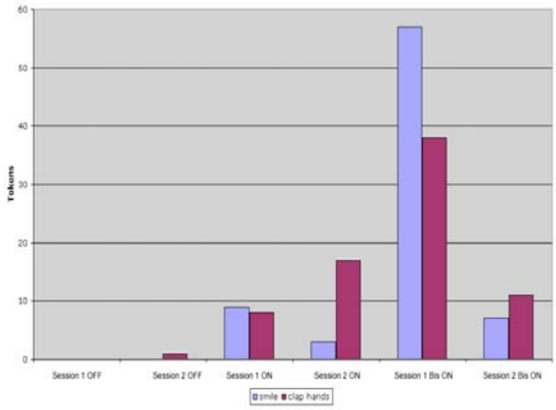


Figure 7: smiling and clapping hands

These results are particularly encouraging since JC was never able to clap the hands before. Furthermore, since JC did not have any speech, the expressions of joy he produced during the activity with Paro could be interpreted as an attempts to communicate quite strong feelings of engagement with the robot. An effect that was totally absent in the Paro passive condition.

5 Discussion and conclusions

The outcomes of our experiment seem to corroborate the initial hypotheses that inspired our study. The introduction of Paro in the Bobath protocol seemed to strengthen the efficacy of the method in rehabilitating physical-functional skills through actively involving the child in engaging exercises.

At present, JC is no longer inserted in a rehabilitative path. He regularly attends the kindergarten and seems to have steadily mastered and maintained the skills acquired during the treatment, in particular the physical-functional ones.

In our research we compared physical with behavioural characteristics of Paro investigating how they differently affect the therapy. In the Paro passive condition only the physical cues were exhibited by the robot whereas in the interactive condition both the physical and behavioural ones were present. Comparing the two conditions we observed that in the passive condition, the sensorial experience was not sufficiently supported to meet the objectives of the treatment. In the interactive condition, the physical and behavioural cues together had a positive influence on the child performance. Furthermore, in the interactive condition the child showed novel behaviours not previously emerged. In addition, the high number of occurrences of such behaviours could be interpreted as a stable acquisition of these skills.

Of course since the results of our experimental study were limited in time and restricted to one subject, they cannot be readily generalised. However they are certainly noteworthy if considered in light of a series of experiments conducted in Italy and Japan. In Italy, Paro was tried out at 'Le Scotte' Hospital in Siena with patients various typologies of neurological diseases. In particular, we made an experiment with two twin sisters affected by the Angelman syndrome, a very rare genetic disease. The results of the experiment show the increase of dyadic (child-robot) and triadic communications (child-robot-therapist) in the Paro interactive condition (Marti et al., submitted).

Other interesting results come from experiments performed in Japan (Saito et al., 2003) in which Paro was introduced to a health service facility for the elderly. The study showed that the introduction of Paro as a mental commit robot produced a good influence for calming the patients and reducing nursing staff's stress.

These experiments confirm the versatility of Paro to be used with efficacy in different contexts and for different purposes, opening new perspectives to the application of robot-assisted therapy.

Acknowledgements

We would like to thank the therapists of the Functional Rehabilitation Unit of the "Le Scotte" hospital in Siena, in particular Adriana Salvini, who supported the study with enthusiasm and professionalism. A special thanks goes to JC and his family to have shared with us their experience of life, their concerns and hopes; and to Filippo Fanò for his valuable support to all phases of the research.

References

Aisen, M.L., Krebs, H.I., Hogan, N., McDowell, F. and Volpe, B.T. The effect of robot-assisted therapy and rehabilitative training on motor recovery following stroke, *Archives of Neurology*, 54(4), pp. 443-6. 1997.

Baron – Cohen S., Howlin P., Hadwin J. *Teoria della Mente e autismo. Insagnare a comprendere gli stati psichici dell'altro* Centro studi Erikson, Trento, 1999.

Baron-Cohen, S. Theory of mind in normal development and autism. *Primse*, 34, pp.174-183. 2001.

Boldt M.A., Dellmann-Jenkins M.: The impact of companion animals in later life and considerations for practice. *Journal of Applied Gerontology*, 11: 228-239, 1992.

Bumby, K., Dautenhahn, K. Investigating Children's Attitudes Towards Robots: A Case Study. In *Proceedings CT99, The Third International Cognitive Technology Conference*, August, San Francisco. 1999.

Burgar, C.G., Lum, P.S., Shor, P.C. and Van der Loos, H.F.M., Development of robots for rehabilitation therapy: The Palo Alto VA/Stanford experience, *Journal of Rehabilitation Research and Development,* 37(6), pp. 663-673. 2000.

Camaioni L., Bernabei P., Levi G., Di Falco M., Paolesse C. Lo sviluppo socio – comunicativo nei primi due anni di vita di bambini con autismo: possibilità di una diagnosi precoce. in *Psicologia clinica dello sviluppo* Anno I, Agosto N. 2, Nucleo monotematico, Il Mulino Ed., Bologna, 1997.

Dautenhahn, K., Werry, I., Harwin W., Investigating a Robot as a Therapy Partner for Children with Autism, *Proc. AAATE, 6TH European Conference for the Advancement of Assistive Technology*, 3-6 September in Ljubljana / Slovenia. 2001

Dautenhahn K., Billard A. Games children can play robota, a Humanoid robotic doll. In *Proceedings 1st Cambridge Workshop on Universal Access and Assistive Technology [CWUAAT] (incorporating 4th Cambridge Workshop on Rehabilitation Robotics),* Trinity Hall, University of Cambridge, Cambridge, United Kingdom. In: S Keates, S., Clarkson, P.J., Langdon, P.M., Robinson, P. (Eds.), *Universal Access and Assistive Technology,* London, Springer-Verlag. 2002.

Dautenhahn, K., Werry, I., Rae, J., Dickerson, P., Stribling, P., Ogden, B. Robotic Playmates: Analysing Interactive Competencies of Children with Autism Playing with a Mobile

Robot. In Dautenhahn, K., Bond, A., Canamero L., Edmonds B. (Eds.), *Socially Intelligent Agents – Creating Relationships with Computers and Robots* , pp. 117-124, Dordrecht, Kluwer Academic Publishers. 2002.

Dautenhahn, K., Werry I. Towards Interactive Robots in Autism Therapy: Background, Motivation and Challenges. In *Pragmatics and Cognition, 12:1 2004* , pp. 1–35. 2004.

Emanuel, R., Weir, S. Catalysing Communication in an Autistic Child in a Logo-like Learning Environment. In *AISB (ECAI)* 1976:, pp.118-129. 1976.

Friedmann E., Katcher A.H., Lynch J.J., et al.: Animal Companions and one-year survival of patient after discharge from a coronary care unit. *Public Health Report*, 95:307-312, 1980.

Johnson, S.C. The recognition of mentalistic agents in infancy. In *Trends in Cognitive Science 4*, pp. 22-28. 2000.

Lago D., Delaney M., Miller M., et al.: Companion animals. attitudes toward pets, and health outcomes among the elderly, A long-term follow-up. *Anthrozoos*, 3:25-34, 1989.

Marti P., Shibata T., Fanò F., Palma V., Pollini A., Rullo A, Agentivity in social interaction with robots. Submitted to *Interaction Studies: Epigenetic Robotics*

Michaud, F., Lachiver, G., Lucas, M., Clavet, A. Designing robot toys to help autistic children - An open design project for Electrical and Computer Engineering education, In *Proceedings American Society for Engineering Education (ASEE) Conference*, St-Louis Missouri. 2000.

Michaud, F., Théberge-Turmel, C. Mobile robotic toys and autism. In *Socially Intelligent Agents - Creating Relationships with Computers and Robots*, Dautenhahn, K., Bond, A., Canamero, L., Edmonds, B. (Eds), Kluwer Academic Publishers. 2002.

Plaisant, C., Druin, A., Lathan, C., Dakhane, K., Edwards, K., Vice, J.M., Montemayor, J. A Storytelling Robot for Pediatric Rehabilitation. Revised version. In *Proceedings ASSETS '00*, Washington, ACM, New York, pp. 50-55. HCIL,CS-TR-4183, UMIACS-TR-2000-65. 2000

Reinkensmeyer, D.J., Kahn, L.E., Averbuch, M., McKenna-Cole, A., Schmit, B.D., and Rymer, W.Z., Understanding and treating arm movement impairment after chronic brain injury: Progress with the ARM Guide, *Journal of Rehabilitation Research and Development*, 37(6), pp. 653-662. 2000a.

Reinkensmeyer, D., Hogan, N., Krebs, H., Lehman, S. and Lum, P., Rehabilitators, robots, and guides: New tools for neurological rehabilitation, in *Biomechanics and Neural Control of Posture and Movement*, J. Winters and P. Crago, Editors. Springer-Verlag. p. 516-533. 2000b.

Volpe, B.T., Krebs, H.I., Hogan, N., Edelstein, L., Diels, C., and Aisen, M., A novel approach to stroke rehabilitation: robot-aided sensorimotor stimulation, *Neurology*, 54(10), pp. 1938-44. 2000.

Shibata, T., Mitsui, T. Wada, Touda, A., Kumasaka, T., Tagami, K., Tanie, K., Mental Commit Robot and its Application to Therapy of Children, Proc. of the IEEE/ASME International Conf. on AIM'01, 182, CD-ROM Proc. 2001

Saito T., Shibata T., Wada K., Tanie K., Relationship between Interaction with the Mental Commit Robot and Change of Stress Reaction of the Elderly" in *Proceedings IEEE International Symposium on Computational Intelligence in Robotics and Automation* July 16-20, Kobe, Japan.2003.

Shibata, T., Tashima, T., Tanie, K. Emergence of Emotional Behavior Through Physical Interaction Between Human and Robot. In *Proceddings of ICRA 1999*, pp. 2868-2873. 1999.

Smith, L.B., Heise., D. Perceptual similarity and conceptual structure. In Burn, B. *Percepts, concepts and categories,* Ed. Elseiver. 1992.

Zappella., M. (Eds) *Il metodo Portage. Guida pratica per l'educazione precoce. Schede.* Edizioni Omega, Torino. 1990

Woodward, A.L., Sommerville, J.A., Guajardo, J.J. How infants make sense of intentional action. In Malle, B., Moses, L., Baldwin, D. (Eds.), *Intentions and intentionality: Foundations of Social Cognition,* pp.149-169, Cambridge, MA, MIT Press. 2001.

Classifying Types of Gesture and Inferring Intent

Chrystopher L. Nehaniv
*Adaptive Systems Research Group
School of Computer Science
Faculty of Engineering and Information Sciences
University of Hertfordshire
College Lane, Hatfield Herts AL10 9AB
United Kingdom
C.L.Nehaniv@herts.ac.uk

Abstract

In order to infer intent from gesture, a rudimentary classification of types of gestures into five main classes is introduced. The classification is intended as a basis for incorporating the understanding of gesture into human-robot interaction (HRI). Some requirements for the operational classification of gesture by a robot interacting with humans are also suggested.

1 Introduction: The Need for Classifying Gesture

The word *gesture* is used for many different phenomena involving human movement, especially of the hands and arms. Only some of these are interactive or communicative. The pragmatics of gesture and meaningful interaction are quite complex (cf. Kendon (1970); Mey (2001); Millikan (2004)), and an international journal *Gesture* now exists entirely devoted to the study of gesture. Applications of service or 'companion' robots that interact with humans, including naive ones, will increasingly require human-robot interaction (HRI) in which the robot can recognize *what* humans are doing and to a limited extent *why* they are doing it, so that the robot may act appropriately, e.g. either by assisting, or staying out of the way. Due to the situated embodied nature of such interactions and the non-human nature of robots, it is not possible to directly carry over methods from human-computer interaction (HCI) or rely entirely on insights from the psychology of human-human interaction. Insights from proxemics and kinesics, which study spatial and temporal aspects of human-human interaction (Hall, 1983; Condon and Ogston, 1967; Kendon, 1970) and some insights of HCI, e.g. recognizing the diversity of users and providing feedback acknowledgment with suitable response timing (e.g. (Shneiderman, 1998)), may also prove to be extremely valuable to HRI. Notwithstanding, the nascent field of HRI must develop its own methods particular to the challenges of embodied interaction between humans and robots. New design, validation, evaluation methods and principles particular to HRI must be developed to meet challenges such as *legibility*, making the robot's actions and behaviour understandable and predictable to a human, and '*robotiquette*', respecting human activities and situations (e.g. not interrupting a conversation between humans or disturbing a human who is concentrating or working intensely — without sufficient cause), as well respecting as social spaces, and maintaining appropriate proximity and levels of attention in interaction. Part of meeting these challenges necessarily involves some understanding human activity at an appropriate level. This requires the capabilities of recognizing human gesture and movement, and inferring intent. The term "*intent*" is used in this paper in a *limited* way that refers to *particular motivation(s) of a human being that result in a gestural motion as relevant for human-robot interaction*.

In inferring the intent from a human's gesture it is helpful to have a classification of which type of gesture is being observed. Without a sufficiently broad classification, understanding of gesture will be too narrow to characterize what is happening and appropriate responses will not be possible in many cases.

While this paper does not attempt a comprehensive survey of the role and recognition of gesture in human-robot interaction, it does suggest inherent limitations of approaches working with a too narrow notion of gesture, excluding entire classes of human gesture that should eventually be accessible to interactive robots able to function well in a human social environment.

The questions of how gestures are acquired and come to be recognized as meaningful by particular individuals in the course of their development (ontogeny of gesture and its recognition), and conventionalized, elaborated, or lost within particular cultures (evolution of gesture) are large and deep issues, but will not be addressed within the scope in this paper.

Knowing how to recognize and classify gesture may also serve to inform the design of robot behaviour, including gestures made by the robot to achieve legibility and convey aspects of the robot's state and plans to humans. This in turn will contribute to robot interaction with humans that is legible, natural, safe, and comfortable for the humans interacting with the robot.

2 Classification of Gestures

The following is a rough, tentative classification. Gestures are classed into five major types with some subtypes.

To begin to approach the complexity of gesture in the context of situated human-robot interaction, the rough classes of gesture described below are developed in order to provide a broad level of description and the first steps toward a pragmatic, operational definition that could be used by an autonomous system such as a robot to help it (1) to infer the intent of human interaction partners, and, as an eventual goal, (2) to help the robot use gestures itself (if possible) to increase the legibility of its behaviour.

Ambiguity of Gesture. It should be stressed that a single specific instance of a particular the kind of physical gestural motion could, depending on context and interaction history, reflect very different kinds of human intents. It will not always be possible to infer intent based solely on based the mechanical aspects of human movements (such as changes in joint angles) without taking context into account.

To approach this problem, a classification of gesture for inferring intent and assisting in the understanding of human activity should closely relate gesture with limited categories of intent in situated human activity. The classes of the tentative classification presented here thus correspond to and allow the (limited) attribution of intent on the part of humans. The classification is developed as an aid for helping robots to achieve limited recognition of situated human gestural motion so has to be able to respond appropriately if required, while these robots are working in an environment of ambient human activity (such as a home or office), in which, at times, the robots are also assisting or cooperating with the humans. *Applications of this classification will require the mapping of physical aspects of gestural motion in interactional contexts to the five gestural classes (and their subtypes) suggested here.*

2.1 Five Classes (with Subtypes)

1. **'Irrelevant'/Manipulative Gestures.** These include *irrelevant gestures*, *body / manipulator motion*, *side-effects of motor behaviour*, and *actions on objects*. Broadly characterized, manipulation by a human is here understood as *doing something to influence the non-animate environment or the human's relationship to it (such as position).* Gestural motions in this class are manipulative actions (in this sense) and their side effects on body movement. These 'gestures' are neither communicative nor socially interactive, but instances and effects of human motion. They may be salient, but are not movements that are primarily employed to communicate or engage a partner in interaction. Cases include, e.g. motion of the arms and hands when walking; tapping of the fingers; playing with a paper clip; brushing hair away from one's face with one's hand; scratching; grasping a cup in order to drink its contents. (Note it may be very important to distinguish among the subtypes listed above for robot understanding of human behaviour.)

2. **Side Effect of Expressive Behaviour.** In communicating with others, motion of hands, arms and face (changes in their states) occur as part of the overall communicative behaviour, but without any specific interactive, communicative, symbolic, or referential roles (cf. classes 3-5)
Example: persons talk excitedly raising and moving their hands in correlation with changes in voice prosody, rhythm, or emphasis of speech.

3. **Symbolic Gestures.** Gestural motion in symbol gesture is a *conventionalized signal in a communicative interaction.* It is generally a member of a limited, circumscribed set of gestural motions that have specific, prescribed interpretations. A symbolic gesture is used to trigger certain actions by a targeted perceiver, or to refer to something or substitute as for another signal according to a code or convention. Single symbolic gestures are analogous to discrete actions on an interface, such as clicking a button.

Examples: waving down a taxi for it to stop; use of a conventional hand signals (a command to halt indicated open flat hand; a military salute); nodding 'yes'; waving a greeting 'hello' or 'goodbye'.

Note that the *degree of arbitrariness* in such gestures may vary: The form of the gesture may be an arbitrary conventional sign (such as a holding up two fingers to mean 'peace', or the use of semaphores for alphabetic letters). On the other hand, a symbolic gesture may resemble to a lesser or greater extent iconically or, in ritualized form, a referent or activity.
Further examples: holding up two fingers to indicate 'two'; opening both (empty) hands by turning palms down to indicate a lack of something. Nearly all symbolic gestures are used to convey *content* in communicative interactions.

4. **Interactional Gestures.** These are gesture used to *regulate interaction with a partner*, i.e. used to initiate, maintain, invite, synchronize, organize or terminate a particular interactive, cooperative behaviour: raising a empty hand toward the partner to invite the partner to give an object; raising the hand containing an object toward the partner inviting them to take it; nodding the head indicating that one is listening. The emphasis of this category is neither reference nor communication but on gestures as mediators for cooperative action.[1] Interactional gestures thus concern regulating the form of interactions, including the possible regulation of communicative interactions but do not generally convey any of the content in communication. Interactional gestures are similar to class 1 manipulative gestures in the sense that they influence the environment, but in contrast to class 1, they influence the "animated environment" – *doing something to influence human agents (or other agents) in the environment*, but not by conveying symbolic or referential content.[2]

5. **Referential/Pointing Gestures.** These are *used to refer to or to indicate objects (or loci) of interest* – either physically present objects, persons, directions or locations the environment – by pointing (*deixis*[3] – showing), or indication of locations in space being used as proxies to represent absent referents in discourse.

Table 1 summarizes the five classes.

2.2 Target and Recipient of a Gesture

If a gesture is used interactively or communicatively (classes 2-5), it is important to recognize whether the gesture is directed toward the current interaction partner (if any) — which may the robot, another person (or animal) present in the context, or possibly neither (*target*). If pointing, what is the person pointing to? Who is the pointing designed to be seen by? (*recipient*) If speaking, to whom is the person speaking? If the gesture is targetted at or involves a contact with an object, this suggests it may belong to class 1 (or possibly 5, even without contact). A gesture of bringing an object conspicuously and not overly quickly toward an interaction partner is manipulative (in the sense explained in the discussion of class 1, since an object is being manipulated), but it may well at the same time also be a solicitation for the partner to take the object (class 4). Similarly if the partner has an object, an open hand conspicuously directed toward the partner or object may be a solicitation for the partner to give the object (class 4).

2.3 Multipurpose Gestures

It is possible for a single instance of a particular gesture to have aspects of more than one class or to lie intermediate between classes. As mentioned above, handing over an object is both class 1 and 4. And, for example, holding up a yellow card in football has aspects of classes 1 and 3, object manipulation and

[1] Note that we are using the word "cooperative" in a sense that treats regulating communication or interaction as an instance of cooperation.

[2] Some more subtle examples include putting one's hand on another person's arm to comfort them. Such actions, and others involving physical contact, may be quite complex to interpret as understanding them may require understanding and modeling the intent of one person to influence that state of mind of another. At this point, we class simply them with interactional gestures recognizing that future analysis may reveal deep issues of human-human interaction and levels of complexity beyond the rudimentary types of human intent considered here. A special case worthy of note is human contact with the robot, unless this is directly a manipulation of the robot's state via an interface - e.g. via button presses — which would fall into class 3 (symbolic gesture), non-accidental human contact with the robot is likely to be indicative of an intent to initiate or regulate interaction with the robot (class 4). Physical contact between humans might also involve expression of affection (kissing), or aggression (slapping, hitting) – which generally indicate types human-human interaction it would be better for a robot to steer clear of!

[3] Deixis can involve a hand, finger, other directed motion, and/or eye gaze. Checking the eye gaze target of an interaction partner is commonly used to regulate reference and interaction; it develops and supports joint attention already in preverbal infants. Language, including deictic vocabulary (e.g. demonstratives such as the words "these" and "that"), and other interactional skills, typically develop on this scaffolding (see Kita (2003)).

conventional symbolic signal. Many ritualized symbolic gestures (class 3) also can be used to initiate or regulate interaction (class 4), e.g. the 'come here' gesture: with palm away from the recipient, moving the fingers together part way toward the palm; waving forearm and open hand with palm facing recipient to get attention. More complex combinations are possible, e.g. a gesture of grasping designed by the human to be seen by a recipient interaction partner and directed toward a heavy or awkwardly-sharped target object as a solicitation of the partner to cooperatively carry the object with the gesturer (classes 1, 4, 5).

2.4 Ritualization: Movement into Classes 3 and 4

Gestures that originate in class 1 as manipulations of the non-animate environment and the person's relationship to it may become *ritualized* to invite interactions of certain types, e.g., cupping the hand next to the ear can indicate that person doing it cannot hear, so that the interaction partner should speak up. Originally cupping the hand near the ear served to improve a person's ability to hear sounds in the environment from a particular direction (class 1), but it may be intended to be seen by a conversational partner who then speaks up (class 4). The hand cupped at the ear can even be used as a conventionalized symbol meaning 'speak up' (clas 3). Other examples of ritualization toward regulation of interaction and also symbolic gesture include mimicking with two hands the motions of writing on a pad as a signal to a waiter to ask for the bill; miming a zipping action across the mouth to indicate that someone should be 'shut up'; or placing a raised index finger over lips which have been pre-formed as if to pronounce */sh/*.

2.5 Cultural and Individual Differences

Different cultures may differ in their use of the various types of gesture. Some symbolic gestures such as finger signs (e.g. the "OK" gesture with thumb and index finger forming a circle) can have radically different interpretations in other cultures, or no set interpretation depending on the culture of the recipient (e.g. crossing fingers as a sign of wishing for luck, or the Chinese finger signs for some numbers such as 6, 7, 8). Tilting the head back (Greece) or nodding the head (Bulgarian) are used symbolically for 'no', but would certainly not be interpreted that way in many other cultures. Cultures also differ in their types and scope of movement in (class 2) expressive gestures: Consider, for example, the differences of rhythm, prosody, hand motions, eye contact, and facial expressions accompanying speech between British, Italian, Japanese, and French speakers.

Within cultures, differences between different individuals' uses of gestures can be regional, restricted to particular social groups within the culture, and vary in particularities (such as speed, repertoire, intensity of movement, etc.) between individuals according to preference or ontogeny. Elderly and young may employ gestures in different ways.

3 Some Related Work on Recognizing Gesture and Intent

The important role of gesture for intent communication in human-robot interaction is increasingly being acknowledged, although some approaches still focus only on static hand poses rather than dynamic use of more general types of gesture in context. A survey of hand gesture understanding in robotics appears in Miners (2002).

Multimodal and voice analysis can also help to infer intent via prosodic patterns, even when ignoring the content of speech. Robotic recognition of a small number of distinct prosodic patterns used by adults that communicate praise, prohibition, attention, and comfort to preverbal infants has been employed as feedback to the robot's 'affective' state and behavioural expression, allowing for the emergence of interesting social interaction with humans (Breazeal and Aryananda, 2002). Hidden Markov Models (HMMs) have been used to classifying limited numbers of gestrual patterns (such as letter shapes) and also to generate trajectories by a humanoid robot matching those demonstrated by a human (Billard et al., 2004). Multimodal speech and gesture recognition using HMMs has been implemented for giving commands via pointing, one-, and two-handed gestural commands together with voice for intention extraction into a structured symbolic data stream for use in controlling and programming a vacuuming cleaning robot (Iba et al., 2002). Many more examples in robotics exist.

Most approaches use very limited, constrained, and specific task-related gestural repertoires of primitives, and do not attempt to identify gestural classes. They have tended to focus on a fixed symbolic set of gestures (possibly an extensible one, in which new gestures can be learned), or focus on only a few representatives from one or two of the gestural classes identified here (e.g. symbolic and pointing gestures).

Knowledge of specific conventional codes and signs can help the identification of particular signs within class 3, and also in determining that the gesture in fact belongs to class 3, i.e. is a symbolic communicative signal. Machine learning methods such as Hidden Markov Models may be used successfully to learn and classify gestures for a limited finite set of fixed gestures (e.g. (Westeyn et al., 2003)). It seems likely that HMM methods would be most successful with class 3 (symbolic gestures), but how successful they would be at differentiating between classes or within other classes remains uninvestigated at present.

4 Inferring the Intent of Gesture

Being able to classify gesture into one of the above classes gives us only a starting point for inferring the intent of the person making the gesture due to frequent ambiguity. Resolving this points to the important roles of context and interactional history. Thus, it is necessary to develop operational methods for recognizing the class of gesture in a particular context.[4]

To this it should help when

(a) the activity of the gesturer is known,

(b) previous and current interaction patterns are remembered to predict the likely current and next behaviour of the particular person,

(c) objects, humans and other animated agents in the environment are identified and tracked.

(d) the scenario and situational context are known (e.g. knowing whether a gesture occurs at a tea party or during a card game).

Knowing the above could help the robot classify the gesture and infer the intent of the human. Information on the state of human (e.g. working, thirsty, talking, ...) often can limit the possibilities.

4.1 Recognizing Intent from Gesture Given Interactional Context

If the interactional context of recent activity in which a gesture occurs is known, this can suggest possibilities for which classes (and subtypes) of gesture might be involved. Even giving data on the interactional context, including data on context, culture, individual differences, models of human activity and task aspects that relate to gesture, does not necessarily completely constrain the possible gesture nor its intent (if any). If the context suggests a particular identifying class (and subtype) for the gesture identified, this does not immediately lead to any certain knowledge of human intent behind it.

Data on the interaction history and context may help in determining the class of a gesture. If the class is known, then the set of possible gestures can remain large, or be narrowed significantly. Symbolic gestures (class 3) correspond to discrete symbols in a finite set, of which their may be only be a small number according to context or size of the given repertoire of the given symbolic gestural code. Interactional gestures (class 4) are likely to comprise a small, constrained class. Class 1 gestures are either "irrelevant", or to be understood by seeking the intent of the associated motor action or object manipulation (e.g. grasping or throwing an object, arms moving as a side effect of walking). Class 5 (referential and pointing gestures) comprise a very limited class.

4.2 Typical Interactional Context of Gestures

A programme to apply the above classification can be developed as follows.

1. Identify the many, particular gestural motions that fit within each of the five classes. Some gestural motions will appear in more than one class. For example, the same mechanical motion of putting a hand and arm forward with the forearm horizontal and the hand open could indicate preparation to manipulate an object in front of the human (class 1), to show which object is being referred to (class 5), or to greet someone who is approaching, or to ask for an object to be handed over (both class 4).
2. Gestural motions identified as belonging to several classes need to be studied to determine in which contexts they occur: determining in which class(es) particular a instance of the gesture is being used may require consideration of objects and persons in the vicinity, the situational context, and the history of interaction.
3. Systematic characterizations of a physical gestural motion together with interactional contexts in which they are occur could then be used to determine the likely class.

[4] Knowledge of the immediate context in some cases needs to be augmented by taking into account of the broader *temporal horizon* of interactional history (cf. Nehaniv et al. (2002)).

Classification of gestural classes and associated (limited) categories of human intent		
Class	Name	Defining Characteristics and Associated Intent
1	'Irrelevant' or Manipulative Gestures	Influence on non-animate environment or human's relationship to it; manipulation of objects, side effects of motor behavior, body motion
2	Side Effect of Expressive Behaviour	Expressive marking, (no specific direct interactive, symbolic, referential role) associated to communication or affective states of human
3	Symbolic Gestures	Conventionalized signal in communicative interaction; communicative of semantic content (language-like)
4	Interactional Gestures	Regulation of interaction with a partner; influence on human (or other animated) agents in environment but generally with lack of any symbolic/referential content used to initiate, maintain, regulate, synchronize, organize or or terminate various types of interaction
5	Referential/Pointing Gestures	Deixis; indicating objects, agents or (possibly proxy) loci of discourse topics, topics of interest; pointing of all kinds with all kinds of effectors (incl. eyes): referential, topicalizing, attention-directing

Table 1: **Five Classes of Gesture.** See text for explanation, details and examples. Note that some occurrences of the same physical gesture can be used different classes depending on context and interactional history; moreover, some gestures are used in a manner that in the same instance belongs to several classes (see text for examples).

4.3 Updating the Interaction History

Attribution of intent related to gesture can then feedback into understanding of the situational context, including motivational state of the human performing the gesture, and becomes part of the updated interaction history, which can then help in inferring intent from ensuing gestures and activity.

5 Conclusions

In order to infer the intent of a human interaction partner, it may be useful to employ a classification of gesture according to some major types – five in the tentative classification proposed here – whose intent may be, in the five classes, absent / directed to objects or environment, incidentally expressive, symbolic, interactional, or deictic. A summary of the classes is given by Table 1.

In order to deploy the inference of intent on robots interacting with humans it will be necessary to operationalize the distinctions between these (sometimes overlapping) classes. This may require the use of knowledge of human activity, recognition of objects and persons in the environment, and previous interactions with particular humans, as well as knowledge of conventional human gestural referencing and expression, in addition to specialized signaling codes or symbolic systems.

The classification presented here suggests some requirements for the design and implementation of systems inferring intent from gesture based on this classification. These requirements might be realized in a variety of different ways using, e.g. continuous low-key tracking or more detailed analysis, event-based and/or scenario-based recognition, and prediction of human activity based on models of human activity flows (with or without recognition of particular humans and their previous interactions), depending the particular needs of the given human-robot interaction design and the constraints and specificity of its intended operational context. Design of a robot restricted to helping always the same user in the kitchen environment would be quite different from one that should be a more general purpose servant or companion in a home environment containing several adults, children and pets, but the classification presented here is applicable in informing the design of gesture recognition for inferring intent in either type of system, and for designing other HRI systems.

Finally, effective human-robot interaction will require generation of gestures and feedback signals by the robot. The classification given here can suggest categories of robotic gestures that could be implemented to improve the *legibility* to humans of the robot's behaviour, so that they will be better able to understand and predict the robot's activity when interacting with it.

Acknowledgments

The work described in this paper was conducted within the EU Integrated Project COGNIRON ("The Cognitive Robot Companion") and was funded by the European Commission Division FP6-IST Future and Emerging Technologies under Contract FP6-002020.

The classification presented here is developed by the author in response to discussions in the COGNIRON project, especially with Rachid Alami, Kerstin Dautenhahn, Jens Kubacki, Martin Haegele, and Christopher Parlitz. Thanks to Kerstin Dautenhahn for useful comments on an earlier draft of this paper.

This paper extends and supercedes University of Hertfordshire School of Computer Science Technical Report 419 (December 2004).

References

A. Billard, Y. Epars, S. Calinon, G. Cheng, and S. Schaal. Discovering optimal imitation strategies. *Robotics & Autonomous Systems,* Special Issue: Robot Learning from Demonstration, 47(2-3): 69–77, 2004.

Cynthia Breazeal and Lijin Aryananda. Recognition of affective communicative intent in robot-directed speech. *Autonomous Robots*, 12(1):83–104, 2002.

W. S. Condon and W. D. Ogston. A segmentation of behavior. *Journal of Psychiatric Research*, 5:221–235, 1967.

Gesture. ISSN: 1568-1475. John Benjamins Publishing Co., The Netherlands `http://www.benjamins.com/cgi-bin/t_seriesview.cgi?series=GEST`.

Edward T. Hall. *The Dance of Life: The Other Dimension of Time*. Anchor Books, 1983.

Soshi Iba, Christian J. J. Paredis, and Pradeep K. Khosla. Interactive multi-modal robot programming. In *Proceedings of the 2002 IEEE International Conference on Robotics and Automation, Washington D.C., May 11-15, 2002*, 2002.

Adam Kendon. Movement coordination in social interaction: Some examples described. *Acta Psychologica*, 32:100–125, 1970.

Sotaro Kita, editor. *Pointing: Where Language, Culture and Cognition Meet*. Lawrence Erlbaum Associates, Inc, 2003.

Jacob Mey. *Pragmatics: An Introduction*. Blackwell Publishers, 2001.

Ruth Garrett Millikan. *The Varieties of Meaning: The 2002 Jean Nicod Lectures*. MIT Press/Bradford Books, 2004.

William Ben Miners. *Hand Gesture for Interactive Service Robots*. MSc Thesis, The University of Guelph, Faculty of Graduate Studies, August 2002.

Chrystopher L. Nehaniv, Daniel Polani, Kerstin Dautenhahn, René te Boekhorst, and Lola Cañamero. Meaningful information, sensor evolution, and the temporal horizon of embodied organisms. In *Artificial Life VIII*, pages 345–349. MIT Press, 2002.

Ben Shneiderman. *Designing the User Interface: Strategies of Effective Human-Computer Interaction*. Addison-Wesley, 3rd edition, 1998.

Tracy Westeyn, Helene Brashear, Amin Atrash, and Thad Starner. Georgia tech gesture toolkit: Supporting experiments in gesture recognition. In *ICMI 2003: Fifth International Conference on Multimodal Interfaces*. ACM Press, 2003.

Robots as Isolators or Mediators for Children with Autism? A Cautionary Tale

Ben Robins
Adaptive Systems Research Group,
University of Hertfordshire, UK
b.1.robins@herts.ac.uk

Kerstin Dautenhahn
Adaptive Systems Research Group, University of Hertfordshire, UK
K.Dautenhahn@herts.ac.uk

Janek Dubowski
School of Psychology & Therapeutic Studies, University of Surrey Roehampton, UK
j.dubowski@roehampton.ac.uk

Abstract

The discussion presented in this paper is part of our investigation in the Aurora project into the potential use of robots as therapeutic or educational 'toys' specifically for use by children with autism. The paper raises some cautions concerning *social isolation* and *stereotypical behaviour* frequently exhibited in children with autism. We present some examples taken from trials with the robots where the children exhibit such behaviour, and discuss possible ways of ensuring not to reinforce stereotypical behaviour and a tendency to social isolation in the children. Especially, we point out an avenue of robots becoming *social mediators* (mediating contact between children and other children or adults). The paper exemplifies interaction where social behaviour was directed at the robot which raises awareness of the goal of the research, namely to help the children to increase their social interaction skills *with other people* and not simply create relationships with a 'social' robot which would isolate the children from other humans even further.

1. Introduction

Robots and other computer-based technologies are increasingly being used in therapy and education. The discussion presented in this paper is part of our investigation in the Aurora project (AURORA, 2005) into the potential use of robots as therapeutic or educational 'toys' specifically for use by children with autism. People with autism have impaired social interaction, social communication and imagination (referred to by many authors as the triad of impairment, e.g. (Wing, 1996)). Our research focuses on ways that robotic systems can engage autistic children in simple interactive activities with the aim of encouraging basic communication and social interaction skills.

Autism is a lifelong developmental disability that affects the way a person communicates and relates to people around them. People with autism show an impairment in understanding others' intentions, feelings and mental states. They have difficulties in understanding gesture, facial expressions and metaphors and forming social relationships and relating to others in meaningful ways generally poses a big problem to them. They also have impaired imagination, i.e. the development of play and imaginative activities is limited.

Literature shows that people with autism feel comfortable in predictable environments, and enjoy interacting with computers, e.g. (Colby and Smith, 1971; Moor, 1998; Murray, 1997; Powell, 1996). Studies into the behaviour of children with autism suggest that they show a preference for interacting with objects rather than with other people. People's social behaviour can be very subtle and could seem, to those with communication problems and a deficit in mind reading skills, widely unpredictable. This can present itself as a very confusing and possibly stressful experience to children with autism, an experience that they, understandably, try to avoid. As a result, it is not just that they might demonstrate a preference for interacting with objects rather than with other people, but, as Hobson suggests, children with autism often seem to relate to a person as an object (Hobson, 2002). Different from human beings, interactions with robots can provide a simplified, safe, predictable and reliable environment where the complexity of interaction can be controlled and gradually increased.

Our previous work demonstrates that although, in experimental situations, children with autism prefer to engage with a ‘robot’ rather than a ‘human’ companion, this can be turned to their advantage (Robins, et al., 2004c; Robins, et al., 2004d). Results show that repeated exposure to a robot over a long period of time can encourage basic aspects of social interaction skills (i.e. simple imitation, turn- taking and role-switch) and can reveal communicative competence in some of the children (Robins, et al., 2004a). Imitation plays an important part in social learning both in children and adults. Nadel found significant correlations between imitation and positive social behaviour in children with autism (Nadel, et al., 1999). Her findings indicate that imitation is a good predictor of social capacities in these children, and when they are being imitated, autistic children improve their social responsiveness. Inspired by these findings, we designed our trials to progressively move from very simple exposure to the robot, to more complex opportunities for interaction, giving the children the opportunity to attempt imitation and turn-taking games with the robot. It is hoped that if a robot succeeds in engaging children with autism in a variety of interactions, including turn-taking and imitation games, then it may potentially contribute to a child’s development of interaction skills

Our previous trials also highlighted that robots (humanoid and non-humanoid) can serve as salient objects mediating joint attention between the children and other people (peers and adults) (Robins, et al., 2004b; Werry, et al., 2001). Werry *et al.* (2001) demonstrated the ability of a mobile robot to provide a focus of attention and shared attention in trials with pairs of children with autism. Here, the robot’s role as a mediator became apparent in child-teacher interactions, child-investigator interactions and child-child interactions. Furthermore, Robins et al., (2004b) showed that, in some cases, specific aspects of the robot’s behavior, such as the autonomous and predictable pattern of moving head and limbs of a humanoid robot, played a major role in eliciting skilful interaction on the part of the children with the adult present in the room at the time. The robot’s role of mediator emphasizes one of our aims, namely not to replace but to *facilitate human contact*. By being an object of shared attention, the robot may potentially become a ‘social mediator’ encouraging interaction with peers (other children with or without autism) and adults.

2. A Cautionary Tale

As described above, during all of our trials the robots were initially the main focus of the children’s attention. This was the case during the child-robot imitation and turn taking games, as well as during the trials when the robot was the object of joint attention mediating interaction between the children and other people. In this paper we focus on some cautions in this respect, which have arisen during the course of the data analysis. These cautions concern two specific but frequently related behaviours, *social isolation* and *stereotypical behaviour* which is often exhibited in children with autism.

2.1 Social Isolation

Often, children with autism are being described as socially isolated, ignoring other people near them, and often treating them as if they were objects (Hobson, 1993, 2002; Siegel, 1998; Tustin, 1990). Tustin in her review of the external descriptive diagnostic features of autism, provides a quote from Kanner that illustrates it very well: “…the people, so long as they left the child alone, figured in about the same manner as did the desk, the bookshelf, or the filing cabinet.” (Tustin, 1990). In some trials in which small groups or pairs of children with autism were exposed to the robot we have noted occasions were the children seek to have an ‘exclusive’ relationship/interaction with the robot *ignoring* their peer and the experimenter.

Examples of these behaviours from two different trials with different children can be seen below.

2.1.1 Example one

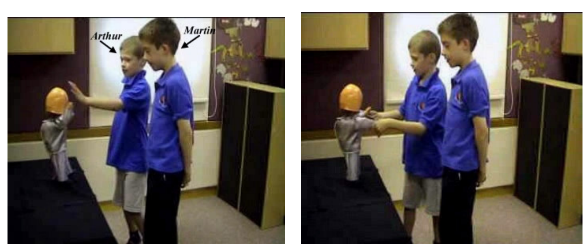


Figure 1: Arthur (left) interacting with the robot whilst Martin (right) waits for his turn.

Figure 1 above shows the beginning of the trial where Arthur (a child with autism) is interacting with the robot, in a very similar way to how he did in a previous trail (simple imitation game). Martin (a child without autism) is standing nearby awaiting his turn (all names in this paper are synomyms).

Figure 2 below shows that whilst it is Martin’s turn for interaction (the robot and the experimenter directed their attention to Martin), Arthur won't 'let go' and continued with his imitation movement, trying to get the robot's attention; and even got annoyed when this did not happen (figure 2 -right).

Figure 2: It is Martin's turn for interacting with the robot, whilst Arthur won't 'let go'.

In figure 3 below, we can see that, whilst Martin is still interacting with the robot, Arthur has stepped forward, ignoring Martin, and touches the moving hands of the robot, seeking exclusive interaction.

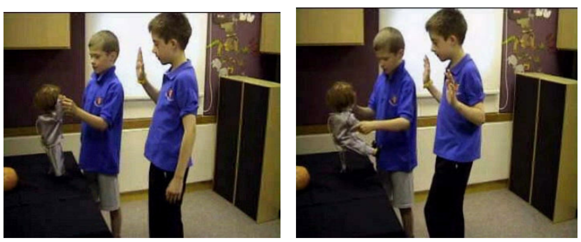


Figure 3:Arthur seeks exclusive interaction with the robot.

2.1.2 Example 2

In this example, two children with autism are playing with the robot 'together' for the first time. Each of them played with the robot individually many times in the past but here they are both exposed to the robot simultaneously.

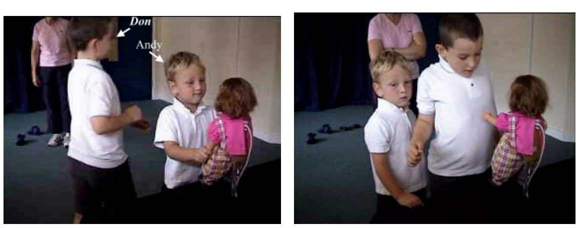


Figure 4 – Andy (left picture) and Don (right picture) Both seeking exclusive interaction with the robot.

Figure 5- Don interacting 'exclusively' with the robot, whilst Andy tries to ignore Don.

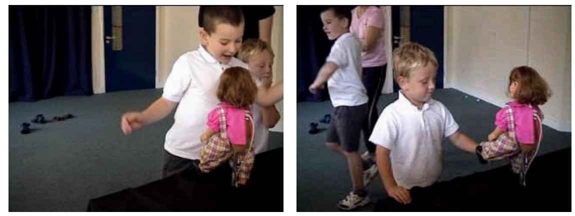


Figure 6 – Don actively seeks exclusive interaction with the robot, whilst Andy waits for exclusive opportunities to interact.

During this session, Don was asked by the teacher to show Andy how to play with the robot. Each time Don went to interact with the robot he actively ensured that he had exclusive interaction, blocking out Andy with his hands. This behaviour repeated itself on different occasions during the session, as can be seen in figures 4 (right), 5 (left), 6 (left).

Andy, on his part, was trying to ignore Don and constantly needed 'encouragement' from his teacher to look at what Don was doing (e.g. figure 5-right). He was either gazing at the robot (figure 5-left), or looking away altogether, as can be seen in figures 4 (right) and 5 (right). Andy interacted with the robot only when he had exclusive access to it, i.e. when Don had stepped away (figures 4-left, 6-right).

These situations clearly highlight that interactions in our trials need to be carefully monitored and taken into consideration when programming the robots and creating the scenarios and games to be played with the robot, to ensure that the robots encourage interaction and become *social mediators* and do not reinforce existing behaviours and become *social isolators*.

2.2 Stereotypical Behaviour

The second caution relates to the highly stereotypical behaviour also frequently noted in children with autism. These highly repetitive forms of behaviour increase social isolation and frequently become self-injurious (Van-Hasselt and Hersen, 1998; White-Kress, 2003; Hudson and Chan, 2002; Jenson, et al., 2001). Our work so far has been limited to the use of robots to develop basic interaction skills through simple imitation and turn-taking activities between the robot and child. Currently, the robots available for this kind of mediation suitable for our experiments are only capable of a relatively limited and repetitive range of movements leading to the caution that this might increase rather than decrease the incidence of these kinds of behaviours.

The following images were taken during trials where children with autism played simple turn–taking and imitation games with a small humanoid robotic doll. The Robot had a very limited range of movements, i.e. the four limbs were capable of moving up and down, and the head could move sideways. This robot's behaviour is far more stereotypical, i.e. shows little variation, as compared to a mobile robot used in other trials, as described below.

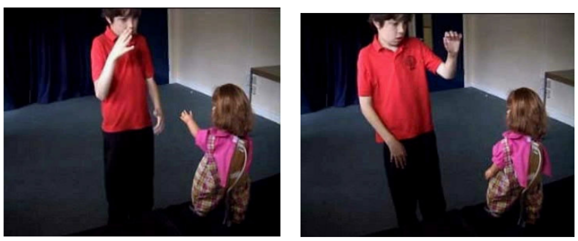


Figure 7 – Tim during a simple imitation game with the robot.

Figure 8 – Billy during a simple imitation game.

In figures 7 & 8 we can see how Tim and Billy engaged in a simple turn-taking and imitation game with the robot. The robot's movements were simple and highly repetitive, and Tim and Billy responded to them each time with almost identical movements.

In comparison, in trials with a mobile robot, where the robot was able to vary its movements during a turn-taking game, the children displayed similar, but not identical, behaviour patterns. Movements were variations of a common theme, rather than instances of a fixed behaviour repertoire. The images in figures 8 & 9 below were taken in a trial where the robot played a turn-taking game with a child. Here, the robot's behaviour varied slightly each time it approached the child or retreated from him (the angle of approach and speed differed, the robot's position relative to the child thus varied). Since the child adjusted his own movements relative to the robot's position and movements, it meant that the child repeated his response (gaze at the robot or touching the robot) each time in a slightly different manner, involving adjustments of his *whole body posture* (e.g. rolling slightly, stretching further away, using another hand etc).

Figure 8 – The robot's varied behaviour in a simple approach/avoidance game: Two instances of approach are shown.

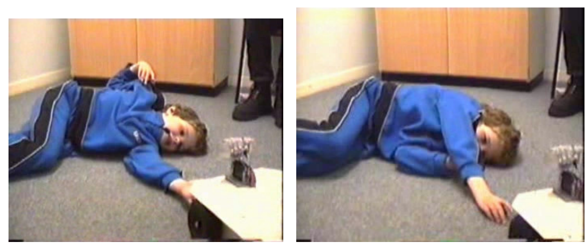


Figure 9 – The child's varied behaviour in the same game: Two instances of 'reaching out' are shown, attempts of touching the robot's front sensors which, as the child has already discovered, will make the robot approach or avoid.

In the above cases involving a mobile robot, we see two interactants that adjust their behaviour relative to, and in response to the other's behaviour, involving full-body movements and encouraging 'natural' types of movements. This situation is very different from those shown in figures 7 and 8, where the children's responses are far more stereotypical and 'mechanistic'.

Using well-defined, salient features, i.e. easy recognizable 'mechanistic' movements seems advantageous e.g. in early stages when children with autism are first being introduced to a robot. These stereotypical movements reduce the complexity of interaction (which is for the children difficult to deal with). However, in later stages, in order not to teach the children to behave like robots and to learn 'robotic movements', robots with more naturalistic, 'biological' movements would be beneficial and a suitable next step in the process of learning.

One of the advantages of using robots, as mentioned earlier, is that the complexity of interaction can be controlled. Bearing in mind the stereotypical nature of the movements of the humanoid robot which we are using, we need to ensure that, over time, we design more complex scenarios of interaction. Also, great attention needs to be paid towards the particular form and shape of movements and behaviour that we encourage in the children. After initial phases of introduction and learning, natural movements are clearly preferred over mechanistic, 'robotic' movements.

2.3 Social Behaviour: Bonding with Robots

Our approach of providing a stress free environment, with a high degree of freedom, facilitated the emergence of spontaneous, proactive, and playful interactions with the robots (Robins, et al., 2004). These interactions included, in some cases, elements of social behaviour directed at the robot.

One example of these behaviour elements occured during the last trial of a longitudinal study (Robins, et al., 2004). Here Billy ended the session running around the room and 'dancing' in front of and directed towards the robot each time he passed it (figure 10 below).

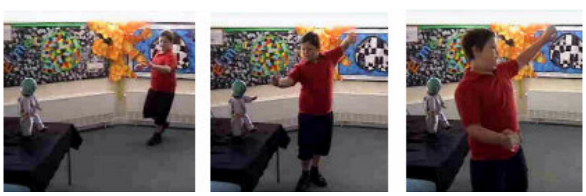


Figure 10 – Billy is 'dancing' to the robot.

Billy repeated this dance in a very similar fashion six months later during the next trial he participated in. (figure 11 below).

Figure 11 – six month later, Billy is 'dancing' again.

Another example of social behaviour displayed by Billy, is when he performed his own unique sign for *good-bye* to the robot. His teacher said at that time that it was as if he was waiting for the robot to say good-bye back to him (figure 12).

Figure 12 - Billy says 'goodbye' to the robot.

The question that must be asked throughout this research is how the children benefit from the interaction with the robots. Are they increasing their social interaction skills (*with other people*) or are we simply encouraging relationships with a 'social' robot? Billy's behaviour was clearly directed towards the robot. In non-autistic children, pretend play or play primarily targeted at other humans present in the room could serve as a possible explanation for this behaviour. However, since children with autism have impairments in these specific domains, it is unlikely that it applies to Billy. Billy very much enjoyed the interactions with the robot, he laughed and smiled during his dance. From a quality of life perspective, this enjoyment is in itself a worthwhile achievement. However, from an educational/therapeutic point of view we must ask whether this sign of 'attachment' or 'bonding' with the robot is worthwhile to pursue, reinforce, or to avoid.

For any child that is usually withdrawn and does not participate in any interaction with other people, 'bonding' with a robot could serve as leverage, and a stepping stone that could provide safety and comfort, opening the child up towards the possibilities of 'human' interactions that are far more unpredictable and complex. Thus, 'bonding with robots' could be beneficial to a child with autism, but only if it is not the ultimately goal, but an *intermediate goal* on the long path towards opening up the child towards other people[1].

3. CONCLUSION

[1]As researchers, this implies a certain responsibility and long-term commitment to this work, that is usually not supported by any existing funding initiatives.

It is not yet clear whether any of the social and communicative skills that the children exhibit during interaction with the robot would have any lasting effect and whether these skills could be generalized and applied in the children's day to day life outside the trial scenario. This aspect is part of our ongoing work. More longitudinal studies are required, together with continued monitoring of the children in their classroom and home environments. Providing experimental evidence for generalization of skills learnt in interactions with the robot is one of our current major challenges from a therapeutic/educational point of view.

From a robotics perspective the appropriate design of robots suitable in therapy and education for children with autism, including the design of suitable and naturalistic robotic movements is a major technological challenge.

References

AURORA. *URL: http://www.aurora-project.com/ last accessed 27/1/05*, 2005.

K. Colby and D. Smith. Computers in the treatment of non speaking autistic children. *Current Psychiatric Therapies* 11:1-17, 1971.

P. Hobson. Understanding persons: the role of affect. In S. Baron-Cohen, H. Tager-Flusberg, and D.J. Cohen, eds., *Understanding other minds, perspectives from autism, chap. 10*, p. 204-227. Oxford University Press, 1993.

P. Hobson. *The Cradle of Thought*. Macmillan, London, 2002.

C. Hudson and J. Chan. Individuals with Intellectual Disability and Mental llness: A Literature Review. *Australian Journal of Social Issues* 37, 2002.

Craig C. Jenson, Gene McConnachie, and Todd Pierson. Long-Term Multicomponant Intervention to Reduce Severe Problem Behaviour: A 63 Month Evaluation. *Journal of Positive Behaviour Interventions* 3, 2001.

D. Moor. Computers and people with autism. *Communication*:20-21, 1998.

D. Murray. Autism and information technology:therapy with computers. In S. Powell and R. Jordan, eds., *Autism and learning: a guide to good practice.*, 1997.

J. Nadel, C. Guerini, A. Peze, and C. Rivet. The evolving nature of imitation as a format of communication. In J. Nadel and G. Butterworth, eds., *Imitation in Infancy*, p. 209-234. Cambridge University Press, 1999.

S. Powell. The use of computers in teaching people with autism. Autism on the agenda: papers from a National Autistic Society Conference. London. In. 1996.

B. Robins, K. Dautenhahn, R. te-Boekhorst, and A. Billard. Robotic Assistants in Therapy and Education of Children with Autism: Can a Small Humanoid Robot Help Encourage Social Interaction Skills? *Universal Access in the Information Society* (in press), 2005.

B. Robins, K. Dautenhahn, R. te Boekhorst, and A. Billard. Effects of repeated exposure of a humanoid robot on children with autism. In S. Keates, J. Clarkson, P. Langdon, and P. Robinson, eds., *Designing a More Inclusive World*, p. 225-236. Springer Verlag, London, 2004a.

B. Robins, P. Dickerson, P. Stribling, and K. Dautenhahn. Robot-mediated joint attention in children with autism: A case study in a robot-human interaction. *Interaction studies: Social Behaviour and Communication in Biological and Artificial Systems* 5:2:161-198, 2004b.

B. Robins, K. Dautenhahn, and J. Dubowski. Investigating Autistic Children's Attitudes Towards Strangers with the Theatrical Robot-A New Experimental Paradigm in Human-Robot Interaction Studies? In *Proc.13th IEEE International Workshop on Robot and Human Interactive Communication - RO-MAN,Kurashiki, Japan, 20-22 September 2004.*, 2004c.

B. Robins, K. Dautenhahn, R. te Boekhorst, and A. Billard. Robots as Assistive Technology - Does Appearance Matter? In *Proc. 13th IEEE International Workshop on Robot and Human Interactive Communication - RO-MAN , Kurashiki, Japan, 20-22 September 2004.*, 2004d.

B. Siegel. *The World of the Autistic Child: Understanding and Treating Autistic Spectrum Disorders*. Oxford University Press, 1998.

F. Tustin. *The Protective Shell in Children and Adults*. Karnac, 1990.

V. B. Van-Hasselt and M. Hersen. Handbook of Psychological Treatment Protocols for Children and Adolescents. *Scandinavian Journal of Behaviour Therapy* 16:95-109, 1998.

I. Werry, K. Dautenhahn, B. Ogden, and W. Harwin. Can Social Interaction Skills Be Taught by a

Social Agent? The Role of a Robotic Mediator in Autism Therapy. In M. Beynon, C.L. Nehaniv, and K. Dautenhahn, eds., *Proc. CT2001, The Fourth International Conference on Cognitive Technology: Instruments of Mind, LNAI 2117*, p. 57-74. Springer-Verlag, Berlin Heidelberg, 2001.

Victoria E White-Kress. Self-Injurious Behaviours: Assessment and Diagnosis. *Journal of Counselling and Development* 81, 2003.

L. Wing. *The Autistic Spectrum*. Constable Press, London, 1996.

Bringing it all together: Integration to study embodied interaction with a robot companion

Jannik Fritsch, Britta Wrede and Gerhard Sagerer
*Applied Computer Science
Technical Faculty, Bielefeld University
{jannik,bwrede,sagerer}@techfak.uni-bielefeld.de

Abstract

One dream of robotics research is to build robot companions that can interact outside the lab in real world environments such as private homes. There has been good progress on many components needed for such a robot companion, but only few systems are documented in the literature that actually integrate a larger number of components leading to a more natural and human-like interaction with such a robot. However, only the integration of many components on the same robot allows us to study embodied interaction and leads to new insights on how to improve the overall appearance of such a robot companion. Towards this end, we present the Bielefeld Robot Companion BIRON as an integration platform for studying embodied interaction. Reporting different stages of the alternating development and evaluation process, we argue that an integrated and actually running system is necessary to assess human needs and demands under real life conditions and to determine what functions are still missing. This interplay between evaluation and development stimulates the development process as well as the design of appropriate evaluation metrics. Moreover, such constant evaluations of the system help identify problematic aspects that need to be solved before sophisticated robot companions can be successfully evaluated in long-term user studies.

1 Introduction

Recent research in robotics focuses on the ambitious goal of building robots that exhibit human-like interaction capabilities in order to allow for natural communication with naive users. This effort is driven by the desire to design robots that can interact outside the lab in real world scenarios such as private households or public schools. However, current systems are still far from reality.

One reason for the difficulties arising in the research process may be seen in the different research traditions in robotics. On the one hand, there is a long tradition in developing human-like functionalities such as grasping, walking, or navigating in order to enable a robot to manipulate its environment in a meaningful way. While there has been a tremendous progress in the different isolated functionalities leading to such impressive results as the walking robot ASIMO (Hirose et al., 2001), a juggling robot (Schaal and Atkeson, 1994), or artificial hands that are able to learn how to grasp things (Steil et al., 2004), there is yet no robotic platform that combines several different functionalities.

On the other hand, there are – more recently – efforts to build so-called *social robots* that are able to communicate with humans in a socially intuitive way (see Fong et al. (2003) for an overview). Here, research is focused on those aspects of social interaction that take advantage of the embodiment of a robotic system. Issues are the modelling and exploitation of joint attention and emotion for socially situated learning, spatial aspects of robot movements, robot appearance, robot personality, etc. (Breazeal et al., 2004; Salter et al., 2004; Robins et al., 2004). In all of these domains important insights were gained and impressive results have been demonstrated with respect to individual social skills. Even integrated social robots displaying different social abilities have been implemented (e.g., Breazeal et al. (2004)). However, integration with other dimensions (e.g., physical functionality, verbal communication) is yet to come.

For human-robot interaction, however, the information conveyed via verbal communication is highly relevant. Interestingly, sophisticated verbal skills that allow a deeper understanding of the user's utterances are rarely integrated in neither the more functional robots nor the social robots. Although there exist extensive literature on natural language-based human-computer communication (Allen et al., 2001, 1996; Carlson, 1996; Cahn and Brennan, 1999) the implementation of such dialogue systems on an embodied platform is still a challenge. The reason for that challenge is the high variability of the physical and communicative context in embodied communication which makes the analysis of spoken utterances difficult.

Nevertheless, the integration of verbal skills is a topic worth pursuing as it promises to endow robots with a much better capability for understanding the human's current situation and his intentions. Research on social robotics has shown impressively that integration is not only necessary but leads to a surprisingly realistic human-like robot behaviour. In this paper we argue that such integration is needed not only for different social skills but also on a broader level for all dimensions of robotic research (i.e., physical functionality, social skills, and verbal communication capabilities).

We believe that only by combining functionalities such as autonomous mobility and navigation with social and verbal communication skills it will be possible to build a robot that can actually fulfil a relevant task in the real world outside the lab with naive (but benevolent and cooperative) users. Only when a robot is capable of functioning in a real life situation realistic long-term evaluation can take place. Results from long-term evaluations are especially valuable since they point out completely new aspects for the development of robots. For example, in a long-term study of the fetch-and-carry robot CERO (Severinson-Eklundh et al., 2003; Hüttenrauch and Eklundh, 2002), issues such as the importance of focusing not only on the user himself, but also on the whole context in that the robot is 'living' in and the reactions of other people, turned out to be very important. Moreover, only long-term studies can take effects such as adaptation or continuously occurring miscommunications or malfunctionings into account and give directions for new research questions for developing robot companions that arise under real life conditions.

In this paper, we present the robot BIRON (the Bielefeld Robot Companion) as an integration platform for building a robot companion. Reporting the different steps in the development of the current system, we argue that an integrated and actually running system is necessary for the development of robot companions.

2 Capabilities of a Robot Companion and its Realization

The development of robots that are equipped with sophisticated human-robot interaction capabilities has been a field of active research in recent years. Since the beginning of research on so-called service-robots, for example as tour guides (e.g., Thrun et al. (2000)), the focus has shifted on building personal robots being suited for use in home environments (e.g., Graf et al., 2004; Bischoff, 2000; Kim et al., 2003). Such personal robots are intended to additionally fulfil communicative and social tasks. However, the maturity of the presented systems with respect to algorithmic stability and the interaction quality is often difficult to assess from publications. Although in many publications it is mentioned that the described robot is capable of interacting with a human, this interaction is often unnatural, e.g., when the user is required to use a keyboard or touch-screen for giving commands to the robot. Obviously, personal robots need to be endowed with a human-friendly interface that allows humans without technical background to interact with such a system.

One typical way of interacting with a robot is a speech interface. However, human-human interaction consists of many more modalities than speech like, e.g., gestures or eye-gaze. Combining such a variety of different modalities which are the topic of active research themselves, is a challenging integration task. Only if this task is successfully solved, the robot's capabilities and especially the interaction quality can be evaluated in user studies.

While in most of the systems reported in the literature the interaction aspects are very prominent, the concept of a robot companion goes beyond these characteristics and stresses social interaction capabilities and the ability to learn. A robot companion has not only to be able to understand natural interaction modalities such as speech and gestures, but should also be able to communicate via these modalities. Its internal representations of the environment need to be open-ended so that it can acquire new information while interacting in the physical world with communication partners.

One scenario that serves as a test-bed for carrying out research on robot companions within the EU-funded 'Cognitive Robot Companion' project (COGNIRON (2004)) is the so-called home-tour scenario that stresses the interaction and spatial learning capabilities of a robot companion. The idea of the home-tour scenario is that a user buys a robot companion at a store and unpacks it at home. In this home scenario, the user has to show the robot all the relevant objects and places in its home that are needed for later interaction (e.g.,"This is my favorite milk glas"). Note that the interaction is not only speech-based but relies heavily on gestures and context information. For example, the context information includes the current room, the viewing direction of the user, information obtained in previous interactions, and so on. In such interactions the robot companion has to dynamically extend its knowledge. As this process can never be finished, this learning is open-ended and the internal realization of the components of the robot companion has to support this open-endedness. For the realization of a robot companion the individual functionalities have to be capable of open-endedness and an appropriate storage of the acquired knowledge that supports flexible retrieval needs to be available. The interaction between the different components will become very complex if multi-modal processing of knowledge and information is required. Thus, besides the development of the individual components, their integration is a major challenge.

Such a tight integration of components that are still under development themselves is obviously nontrivial. However, waiting for the individual components to be mature is not an option, either, as the development of the individual components will be heavily influenced by testing them in an integrated system. We are convinced that only this 'embodiment' of a robot companion will result in good testing conditions to stimulate the research on the individual components. Thus, in the following we describe the lessons learned during building our robot companion BIRON.

3 BIRON – The Bielefeld Robot Companion

Before describing the development process that has led to BIRON's current capabilities in more detail, we present its hardware platform and sensors in the next section.

3.1 BIRON's Hardware Platform

The mobile robotic platform used in our lab for studying embodied interaction is a Pioneer PeopleBot from ActivMedia (see Fig. 1). The platform is equipped with several sensors to obtain information of the environment and the surrounding humans: A pan-tilt colour camera is mounted on top of the robot for acquiring images of the upper body part of humans interacting with the robot. Two far-field microphones are located at the front of the upper platform, right below the touch screen display, for localising sound sources. A SICK laser range finder is mounted at the front on the base platform.

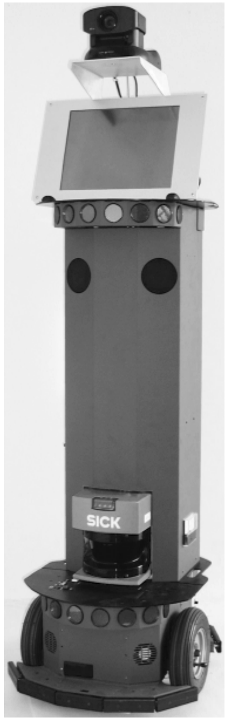


Figure 1: BIRON.

All software components are running on a network of distributed computers. The on-board PC in the robot's base (Pentium III, 850 MHz) is used for controlling the driving motors and the on-board sensors as well as for sound localisation. An additional PC inside the robot's upper extension (Pentium III, 500 MHz) is used for image processing as well as for person tracking and person attention. This second PC is connected to a 12" touch screen display on top of the robot that can be used as additional interactive device.

The two on-board PCs running Linux are linked by 100 Mbit Ethernet to a router with wireless LAN (WLAN). An additional laptop (Pentium M, 1.4 GHz) equipped with a wireless headset is linked to the on-board PCs via this WLAN. User commands given via natural speech are recorded with a wireless headset and speech processing is carried out on the laptop.

3.2 Building BIRON

When starting our activities on human-robot-interaction with a mobile robot some years ago, we first controlled only the robot's movements with speech commands. Since the user's utterances were restricted to low-level steering commands, these first interactions were very limited. The robot had no sense where the user was and whether the user was talking to the robot or to another person in the room. In order to create a more interactive environment, we first realized a multi-modal system using depth data, vision, and sound to enable our robot to track the humans in his environment (Fritsch et al., 2003). Based on the humans tracked in its surrounding and the multi-modal information associated with the individual humans, the robot was equipped with an attention mechanism to selectively pay attentions to humans looking at him and speaking at the same time (Lang et al., 2003). Such a behaviour can be seen as purely reactive in the sense that the robot paid attention to whoever was speaking and looking at the same time to the robot.

However, such a reactive behaviour is not adequate if a human wants to engage in a communicative interaction with the robot. We assume that such a one-to-one interaction is wanted by the human if he greets the robot by saying “Hello Robot”. In order to enable BIRON to understand natural language, we integrated components for speech recognition (Lang et al., 2003), speech understanding (Haasch et al., 2004) and dialog (Toptsis et al., 2004). When recognising a greeting by the human, the reactive attention behaviour needs to be deactivated to fix the attention to the communication partner. At this point the need for an integration scheme or architecture emerges for the purpose of configuring individual components, coordinating the exchange of data between components, and controling access to the hardware, e.g., the pan-tilt camera. In order to allow for an ongoing evolution of the robot, we developed a generic module, the so-called execution supervisor, that routes data between the different components and controls their configuration (Kleinehagenbrock et al., 2004). With this component the mobile robot is able to pay attention to different persons and engage in a one-to-one interaction with one user if he greets the robot by saying “Hello Robot”. From this point on, the robot focuses on this communication partner and engages in a dialog with him.

In the most recent version, the communication partner can not only get the robot’s attention but can also control the robot’s behaviour by giving commands. For example, the command “Follow me” results in the robot following the human around. This functionality seems somewhat similar to our initial activities of controlling the robot’s motion with a microphone, but in the current version a much larger number of components is involved (for more details see Haasch et al., 2004) and the robot exhibits a much more human-like behaviour. The movements of the pan-tilt camera reinforce the impression that the robot is paying attention and focusing on a communication partner, enabling humans to ‘read’ the robot’s current internal state. This somewhat human-like behaviour is very much appreciated by users (Li et al., 2004). Although at this point the robot has not yet something comparable to a ‘personality’, we consider it to be the very first version of our goal of building a robot companion.

Another crucial feature is that, through configuring the behaviour of the execution supervisor based on an XML file containing the definition of its internal states and the associated transitions, the robot companion can be extended easily when new components are available. This is crucial as a real robot companion needs many different functionalities that have to work cooperatively together to reach a good performance.

However, the possibility to add new components by making small changes in the execution supervisor does not mean that integrating new components is easy. Building large frameworks consisting of many software components is an enormous challenge. At this point software engineering aspects come into play. In order to enable integrating new components on BIRON easily and support the evolution of data structures, we use an XML-based communication framework. This framework together with a three-layer architecture including the execution supervisor forms our system infrastructure that enables the ongoing evolution of our robot companion BIRON (Fritsch et al., 2005).

4 Studying Embodied Interaction with a Robot Companion

During the different phases in the development we had the opportunity to observe different stages of the system and we examined the different kinds of interactions. In the following we report the most salient qualitative phases that our system underwent and show how the integration of different modules changes the overall quality of the system and the way it is perceived. Additionally, by reporting different stages of the evaluation and development process we show how important the interplay between evaluation and development is and that constant evaluation of the system is crucial for the further design of the system. We argue that it is necessary to build running systems that enable long term user studies in order to assess human needs and demands under real life conditions and to determine what functions are still missing.

4.1 Lessons learned from building BIRON

BIRON was developed from a remote-controlled mobile device that could be steered via speech commands such as “turn left” to a robot that can engage in (admittedly very restricted) natural interaction and understand more complex instructions such as “Follow me” or “This is a chair”. When comparing BIRON’s capabilities from the very beginning of the building phase to the current abilities, the question arises, what were the most salient phases in the development of the robot? What makes the difference between a remote-controlled toy and the appearance of an intelligent system?

One of the first steps towards the appearance of a robot with a personality was the integration of the person attention system and its control of the camera movements. The person attention module enabled the camera to actively look around for faces. Once it had found a face it was able to track this face for several seconds before moving on to another face. The effect was striking: people felt as if the robot was ‘observing’ them and looking for the most interesting person. Even though the robot was not able to understand speech at this stage, people started to talk to it and wanted to get its attention. Thus, the purposefully moving camera induced an anthropomorphising tendency in the human observers.

Another important, though technically trivial, step was to bring the speech synthesis output on-board. For technical reasons, the speech processing of BIRON takes place on a remote laptop, so that in a first phase the synthesised speech output was simply sent to loudspeakers attached to the laptop. However, this gave a surprising impression of a distributed system because the user was supposed to speak to the robot, but received the feedback from a completely dissociated device at the other end of the room. When moving the sound output to the on-board loudspeakers the appearance of the robot became suddenly more coherent and holistic.

The development of adequate verbal communication capabilities is a major challenge when building a robot companion. During the development of BIRON, we discovered however that there are ‘cheap tricks’ that suggest real intelligence. By simply repeating words uttered by the user the robot can give the impression of deep understanding. For example, when the user says “This is my computer” the robot will reply “Ok, I’m having a look at your computer”. This effect may be explained as a phenomenon of alignment (Pickering and Garrod, 2004). The theory of alignment states that the mutual understanding of two communication partners is often conveyed by the common use of prosodic, lexical, or syntactic structures from both partners. Thus, the repetition of a word would, therefore, indicate a mutual understanding in the given context.

Another intriguing effect could be achieved by moving the camera in the same direction as the pointing gesture of the user. Although this capability is currently simulated since the gesture recognition is not yet integrated (the camera always moves to a predefined position when a certain instruction is recognised), users tend to interpret this behaviour as the robot being not only able to recognise gestures, but to understand the user’s intention.

These examples show that already little and supposedly trivial communicative features can have a tremendous effect on the appearance and hence the interaction capability of a robot companion. Therefore, it is crucial to integrate and evaluate individual components in a running system.

4.2 First User Tests

For first evaluations of our integrated system we asked visitors at an open door event in our lab to interact with BIRON. For this evaluation we decided to give the users a task that they should fulfil during their interaction with BIRON in order to simulate a more task-oriented communication. Since the functional capabilities of BIRON were still limited to very basic behaviours (following and showing) we defined a meta-task in which the users should go through all interaction states of BIRON’s interaction capabilities. Note that in this stage we were not using the wireless headset yet but instead recording the user’s speech with the on-board stereo microphones.

Figure 2 shows the different states of the dialogue as a finite state machine. In order to start the interaction with BIRON the user has to greet BIRON while maintaining eye-contact with its camera. Once the user has registered to the system he can either start an object-showing sequence (‘interaction’ and ‘object’) or make BIRON follow her (‘follow’). The interaction is finished when the user says “Good bye” to BIRON or when the system fails to observe the user, e.g., by losing track of the legs or face percepts.

From this event and subsequent user interactions in our lab we collected data from 21 users. The users were mainly technically interested and had a computer science background, but were otherwise not familiar with robots in general or with BIRON in particular. In order to understand the interaction capabilities of BIRON they received an interaction chart similar to the one displayed in Figure 2. They also had the opportunity to watch demonstrations by an experienced user or interactions of other naive users with BIRON. The latter turned out to be a rich source of information for unexperienced users helping them

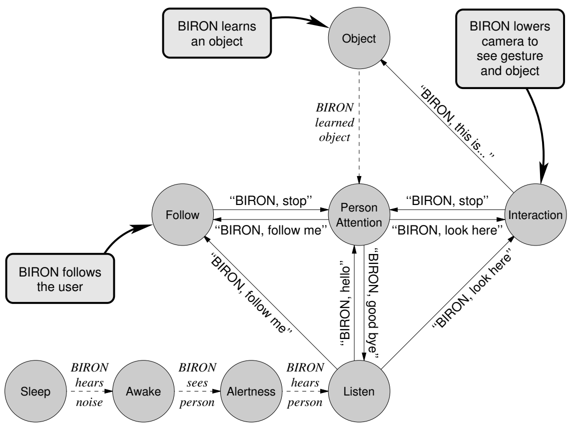


Figure 2: Speech commands and internal states of BIRON.

to learn from other users' errors or success strategies. After this introduction phase, users interacted for about 3-5 minutes with BIRON. We then asked the users to fill out a questionnaire, in which we asked them about their opinion on different aspects of the interaction with BIRON and their general attitude towards robots.

The answers given in the questionnaires indicate that the already existing social capabilities of BIRON received by far the most positive feedback. In contrast, the verbal capabilities – while being judged as very important – received the most negative feedback. Strikingly, the actual functionalities of BIRON, following the user and a simulation of an object learning behaviour, seem not to be in the focus of the users' attention. Figures 3 and 4 show the answers to the questions for the most positive and the most negative functions or characteristics of the system in detail.

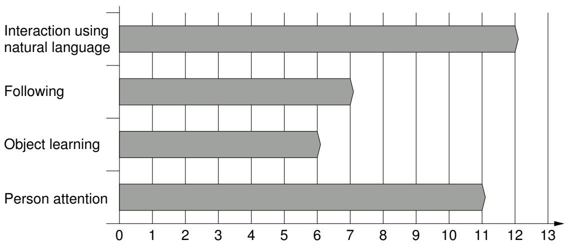


Figure 3: Histogram of answers to "What are the most interesting capabilities of the robot?" (multiple answers were possible.)

As can be seen in Figure 3, the person attention system, which is the most salient social capability of BIRON, and the verbal communication capability received most of the positive answers, whereas the following and object learning function were only named

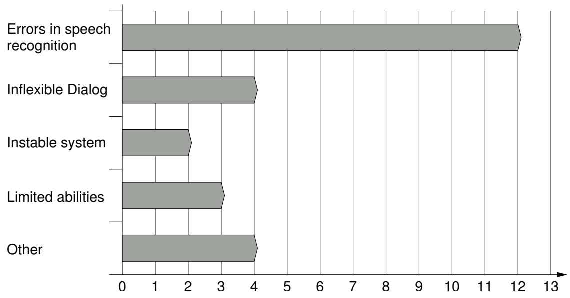


Figure 4: Histogram of answers to "What in the system didn't you like?" (multiple answers were possible.)

half as often. Also, limited abilities of BIRON were only mentioned by three users when being asked for what they did not like. On the other hand, the speech recognition errors received most of the negative feedback. It should be noted, however, that the actual mis- or non-recognition of user utterances was only partly caused by speech recognition errors. Most of the understanding errors were caused by the attention system that is responsible for switching the speech recognition on and off. Due to difficulties in tracking face or voice percepts, which are important cues for starting the speech processing, the speech recognition component was often turned on too late and turned off too early so that important parts of the utterance were missing. Switching speech processing on and off is necessary as BIRON must only listen to utterances that are directed at him. Otherwise, the speech recognition may erroneously extract some instruction from a human-human communication or background noise that is recorded via the stereo-microphones.

The judgements of the users indicate that the most critical component in our system is the speech processing module. In conveying semantic content for more complex instructions to the robot, speech appears to be the most intuitive and natural means for most of the users. Thus, if the speech processing component fails this will be immediately noticed (and commented) by the user.

From these observations we concluded that the highest priority in the further development of BIRON was the optimisation of the speech recognition system. With the purpose of increasing the signal quality by using a close-talking microphone we introduced a wireless headset. More importantly, we increased the reliability of the attention system by increasing the performance of the sound source localisation and the person tracking. These modifications turned out to improve the speech recognition and thereby the whole system performance significantly.

We were able to evaluate this more robust system at the IST Event 2004 (see IST2004) in The Hague where the robot had to perform in an uncontrolled environment with many people walking around the robot and a generally very noisy surrounding (see Fig. 5). Although most demonstrations were given by experienced users we were also able to observe interactions with naive users. We noticed that such interactions were highly affected by the performance of the speech recognition system for the individual speakers. This indicates that future systems may need to be highly adaptable to different users.

Figure 5: BIRON at the IST 2004 in The Hague.

Another interesting observation during the IST was that the motivation to interact with the robot was clearly increased simply by the fact that the robot was continuously running all the time. Also, due to the new surroundings completely new situations arose which indicated the need for more functionalities. For example, people tended to show small things by waving them in front of the camera. In our architectural design, however, we expected primarily pointing gestures for object presentation and waving things in front of the camera demands completely different recognition algorithms from our system. Other observations included people moving or speaking too fast which suggests that our dialogue system should be able to deal in a smooth and natural way with such cases of 'mal-behaviour' of users. For example, it could ask the user to speak more slowly.

In summary, our experience with different users in different situations showed that more thorough studies are needed for the development of a robust robot companion. On the one hand, evaluations are needed where people want to fulfil a real task in order to create a real task-oriented communication. On the other hand, long term studies are necessary in order to assess adaptation processes by learning effects or habituation on the user's side and to rule out artefacts created by one-time interactions such as curiosity or novelty.

5 Outlook

A very important aspect that is not yet explicitly accounted for in BIRON is the representation of its acquired knowledge. Not only need the recognition components algorithm-specific models to recognise places, objects, and persons, but these representations must also be accessible to other system components. For example, the instruction "Get John's cup from the kitchen" requires the robot to go to a position called "kitchen" and to activate the object recognition there with the model of a specific "cup" belonging to "John". Building up knowledge bases that store all these different types of information and make them accessible to the overall system is one important topic of ongoing research. It should be noted that the accessibility is closely linked to the relations between all the different kinds of information. For example, "Get the cup I used yesterday" requires identifying the object model belonging to the object the human did manipulate in the past.

Not only are these aspects related to the hidden capabilities of a robot companion important, but also its appearance. BIRON is basically a red barrel with a moving pan-tilt camera indicating roughly its focus of attention. A more natural looking alternative is some kind of humanoid robot. Here, the question arises how such a humanoid should look like and how its observable behaviour can be shaped to support a human-like interaction quality. For researching such aspects of human-robot interaction we are currently installing a humanoid torso with a face enabling simple mimics (see Fig. 6).

Independent of the actual hardware and appearance of a robot companion, integration will remain a challenging task as the coordination of a large number of software components and their ongoing evolution make the realization of robust systems difficult. Nevertheless, building integrated systems that possess only a part of the ultimately needed functionality for a human-like interaction between a robot companion and a human is already worthwhile. Evaluations of such preliminary robot companions give insights

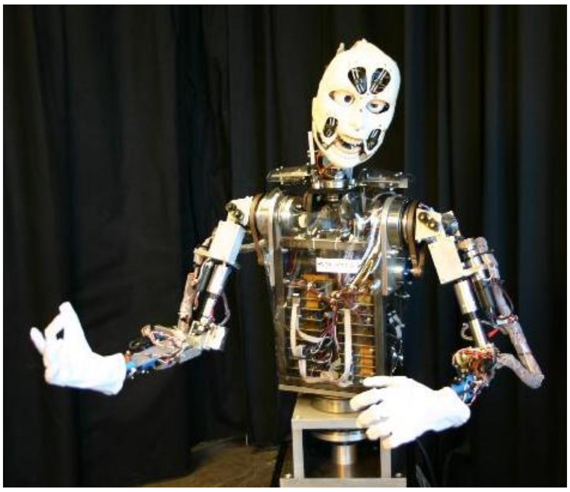


Figure 6: Humanoid torso with a head supporting simple mimics and gestures.

into the aspects central for improving the robot's performance as experienced by users.

In general, there is much ahead for the research on embodied interaction. This will require even more cooperation and interdisciplinary research as it is already being pursued. While many questions are not even touched yet, the evaluation of dependable prototypes will bring up completely new research areas.

Acknowledgements

The research described in this paper is based on work that has been supported by the European Union within the 'Cognitive Robot Companion' project (COGNIRON (2004), FP6-IST-002020) and by the German Research Foundation (DFG) within the Collaborative Research Center 'Situated Artificial Communicators' (SFB360) as well as the Graduate Programs 'Task Oriented Communication' and 'Strategies and Optimization of Behavior'.

The development of BIRON and all of its components has been carried out by a large number of people that have contributed to the system components, their integration, and the insights outlined above in various ways. We like to thank these Master and Ph.D. students listed here in alphabetical order: Henrike Baumotte, Axel Haasch, Nils Hofemann, Sascha Hohenner, Sonja Hüwel, Marcus Kleinehagenbrock, Sebastian Lang, Shuyin Li, Zhe Li, Jan Lümkemann, Jan F. Maas, Christian Plahl, Martin Saerbeck, Sami Awad, Joachim Schmidt, Thorsten Spexard, Ioannis Toptsis, Andre Zielinski.

References

J. Allen, D. Byron, M. Dzikovska, G. Ferguson, L. Galescu, and A. Stent. Towards conversational human-computer interaction. *AI Magazine*, 22(4): 27–37, 2001.

J. Allen, B. Miller, E. Ringger, and T. Sikorski. Robust understanding in a dialogue system. In *Proc. Assoc. for Computational Linguistics, June*, 1996.

R. Bischoff. Recent Advances in the Development of the Humanoid Service Robot HERMES. In *Proc. European Workshop and Masterclass on Advanced Robotics Systems Development (EUREL)*, volume 1, pages 125–134, Manchester, UK, 2000.

C. Breazeal, A. Brooks, J. Gray, G. Hoffman, C. Kidd, H. Lee, J. Lieberman, A. Lockerd, and D. Mulanda. Humanoid robots as cooperative partners for people. *International Journal of Humanoid Robots*, 2(1), 2004.

J. Cahn and S. Brennan. A psychological model of grounding and repair in dialog. In *Proc. Fall AAAI Symposium on Psychological Models of Communication*, 1999.

R. Carlson. The dialog component in the waxholm system. In *Proc. Twente Workshop on Language Technology (TWLT11) - Dialogue Management in Natural Language Systems*, Twente, 1996.

COGNIRON. The Cognitive Robot Companion, 2004. (FP6-IST-002020), `http://www.cogniron.org`.

T. Fong, I. Nourbakhsh, and K. Dautenhahn. A survey of socially interactive robots. *Robotics and Autonomous Systems*, 42:143–166, 2003.

J. Fritsch, M. Kleinehagenbrock, A. Haasch, S. Wrede, and G. Sagerer. A flexible infrastructure for the development of a robot companion with extensible HRI-capabilities. In *Proc. IEEE Int. Conf. on Robotics and Automation*, Barcelona, Spain, 2005. to appear.

J. Fritsch, M. Kleinehagenbrock, S. Lang, T. Plötz, G. A. Fink, and G. Sagerer. Multi-modal anchoring for human-robot-interaction. *Robotics and Autonomous Systems*, 43(2–3):133–147, 2003.

B. Graf, M. Hans, and R. D. Schraft. Care-O-bot II—Development of a next generation robotic home assistant. *Autonomous Robots*, 16(2):193–205, 2004.

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 *Int. Workshop on Advances in Service Robotics*, pages 27–32, Stuttgart, Germany, 2004.

M. Hirose, Y. Haikawa, T. Takenaka, and K. Hirai. Development of humanoid robot ASIMO. In *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, 2001.

H. Hüttenrauch and K. Severinson Eklundh. Fetch-and-carry with CERO: Observations from a long-term user study with a service robot. In *Proc. IEEE Int. Workshop on Robot-Human Interactive Communication (ROMAN)*, pages 158–163. IEEE Press, 2002.

IST2004. Information Society Technologies Event, November 2004. `http://europa.eu.int/information_society/istevent/2004/index_en.htm`.

G. Kim, W. Chung, M. Kim, and C. Lee. Tripodal Schematic Design of the Control Architecture for the Service Robot PSR. In *Proc. IEEE Int. Conf. on Robotics and Automation*, volume 2, pages 2792–2797, Taipei, Taiwan, 2003.

M. Kleinehagenbrock, J. Fritsch, and G. Sagerer. Supporting Advanced Interaction Capabilities on a Mobile Robot with a Flexible Control System. In *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, volume 3, pages 3649–3655, Sendai, Japan, 2004.

S. Lang, M. Kleinehagenbrock, S. Hohenner, J. Fritsch, G. A. Fink, and G. Sagerer. Providing the Basis for Human-Robot-Interaction: A Multi-Modal Attention System for a Mobile Robot. In *Proc. Int. Conf. on Multimodal Interfaces*, pages 28–35, Vancouver, Canada, 2003. ACM.

S. Li, M. Kleinehagenbrock, J. Fritsch, B. Wrede, and G. Sagerer. "BIRON, let me show you something": Evaluating the interaction with a robot companion. In W. Thissen, P. Wieringa, M. Pantic, and M. Ludema, editors, *Proc. IEEE Int. Conf. on Systems, Man, and Cybernetics, Special Session on Human-Robot Interaction*, pages 2827–2834, The Hague, The Netherlands, 2004. IEEE.

M. Pickering and S. Garrod. Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, 27(2):169–190, 2004.

B. Robins, K. Dautenhahn, R. te Boekhorst, and A. Billard. Robots as assistive technology - does appearance matter? In *Proc. IEEE Int. Workshop on Robot-Human Interactive Communication (ROMAN)*, pages 277–282, Kurashiki, Okayama Japan, 2004.

T. Salter, R. te Boekhoerst, and K. Dautenhahn. Detecting and analysing children's play styles with autonomous mobile robots: A case study comparing observational data with sensor readings. In *Proc. Int. Conf. on Intelligent Autonomous Systems*, pages 61–70, Amsterdam, 2004.

S. Schaal and C. Atkeson. Robot juggling: An implementation of memory-based learning. *Control Systems Magazine*, 14, 1994.

K. Severinson-Eklundh, H. Hüttenrauch, and A. Green. Social and collaborative aspects of interaction with a service robot. *Robotics and Autonomous Systems, Special Issue on Socially Interactive Robots*, 42(3–4), 2003.

J. Steil, F. Röhling, R. Haschke, and H. Ritter. Situated robot learning for multi-modal instruction and imitation of grasping. *Robotics and Autonomous Systems*, Special Issue on "Robot Learning by Demonstration"(47):129–141, 2004.

S. Thrun, M. Beetz, M. Bennewitz, W. Burgard, A. B. Cremers, F. Dellaert, D. Fox, D. Hähnel, C. Rosenberg, N. Roy, J. Schulte, and D. Schulz. Probabilistic algorithms and the interactive museum tour-guide robot Minerva. *Int. Journal of Robotics Research, Special Issue on Field and Service Robotics*, 19(11):972–999, 2000.

I. Toptsis, S. Li, B. Wrede, and G. A. Fink. A multimodal dialog system for a mobile robot. In *Proc. Int. Conf. on Spoken Language Processing*, volume 1, pages 273–276, Jeju, Korea, 2004.

Human Interactive Robot for Psychological Enrichment and Therapy

Takanori Shibata[*1, 2], Kazuyoshi Wada[*1], Tomoko Saito[*1], Kazuo Tanie[*3]
*[1]Intelligent Systems Research Institute, National Institute Advance Science and Technology (AIST)
1-1-1 Umezono, Tsukuba, Ibaraki 305-8568, JAPAN
*[2]PERESTO, JST
*[3]National Institute Advance Science and Technology (AIST)
{shibata-takanori, k-wada, tomo-saito, tanie.k}@aist.go.jp

Abstract

"Human interactive robots for psychological enrichment" are a type of service robots that provide a service by interacting with humans while stimulating their minds. Different from the industrial robots, accuracy or speed is not always of prime very importance. Their function or purpose is not simply entertainment, but also to render assistance, to guide, to provide therapy, to educate, to enable communication and so on. The market for human interactive robots designed for psychological enrichment is expected to grow rapidly and to become more wide-spread. This paper explains human-robot interactions in terms of the relationship between humans and robots, in terms of the duration of these interactions and in terms of design issues affecting human interactive robots for psychological enrichment. Then, examples of robot assisted activity using a human interactive robot are described.

1. Introduction

There are two categories of robots that are commonly recognized in the robotics industry; industrial robots and service robots (World Robotics, 2003). Industrial robots have been used widely in manufacturing factories since the early 1960s. Typical tasks for industrial robots are welding, assembly, painting, packaging and palletizing in automotive manufacturing and other industries. Industrial robots work very fast and accurately at their tasks, though they have to be taught by a human operator and their environment has to be specially prepared so that they can accomplish their tasks. Most industrial robots are considered as a potential danger to humans, so people are kept isolated from them.

The market for industrial robots grew rapidly during the 1970s and 1980s, with a peak demand in 1991. However, due to the subsequent recession in the world economy, the market for industrial robots has been slow or stagnated over the last decade. The price of industrial robots plummeted during the 1990s, whiles at the same time their performance, measured both in terms of mechanical and electronic characteristics, was improving continuously (World Robotics, 2003). The price of a typical industrial robot in 2000 was 43% less than it was in 1990. If advances in quality are taken into account, the adjusted price in 2000 was 80% less than it would have been in 1990. This means that value of industrial robots has decreased in real terms, even though they have undergone considerable technical advances.

On the other hand, service robots are new developments in the robotics industry, and include many different kinds of robot. These can be classified into two sub-categories; service robots for professional use and service robots for personal and private use (World Robotics, 2003). Service robots for professional use include cleaning robots, sewer robots, inspection robots, demolition robots, underwater robots, medical robots, robots for disabled persons such as assistive robots and wheelchair robots, courier robots, guide robots, refueling robots at gas stations, fire- and bomb-fighting robots, construction robots, agricultural robots and so on. Service robots for personal and private use include domestic (home) robots for vacuum cleaning, lawn-mowing and so on, as well as entertainment robots, educational robots and on the like. These service robots have been developed to interact with human beings.

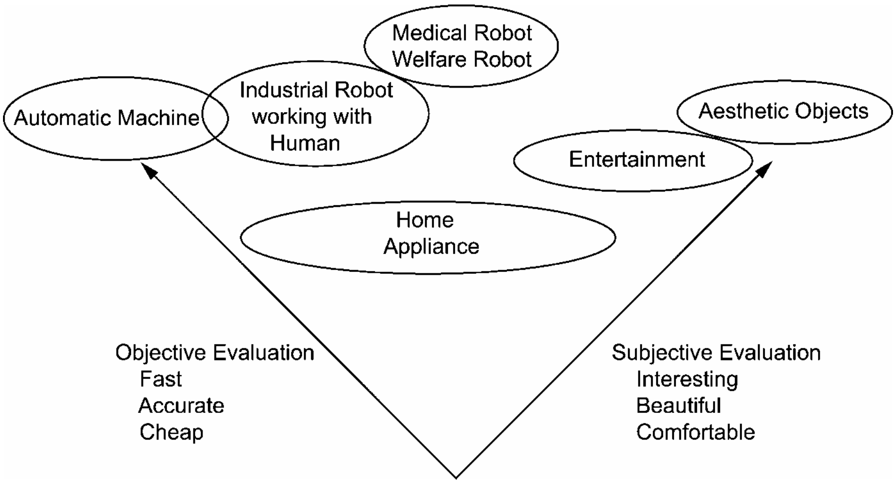


Figure 1: Objective Measures and Subjective Measures to Evaluate Artifacts

Service robots have much more interaction with human beings than industrial robots. They are evaluated not only in terms of objective measures such as speed and accuracy, but also in terms of subjective measures for interacting humans, such as joy and comfort. Service robots for entertainment are clear examples of importance of a subjective evaluation for their values (Fig. 1).

"Human-interactive robots for psychological enrichment" are a type of service robots that provide a service by interacting with humans while stimulating their minds, and we therefore tend to assign high subjective values to them. It is not necessary for these robots to be exclusive, but they should be as affordable as other new luxury products (Tucker, 1995; Silverstein et al., 2003). In addition, accuracy or speed is not always of prime very importance. Their function or purpose is not simply entertainment, but also to render assistance, to guide, to provide therapy, to educate, to enable communication and so on. The market for human interactive robots designed for psychological enrichment is expected to grow rapidly and to become more wide-spread.

In Chapter 2, human-robot interactions are explained in terms of the relationship between humans and robots, in terms of the duration of these interactions and in terms of design issues affecting human interactive robots for psychological enrichment. In Chapter 3, examples of robot assisted activity using human interactive robots are described. Chapter 4 summarizes the overview.

2. Human Robot Interaction

2.1. Relationship between Humans and Robots

There are four categories of human interactive robots for psychological enrichment in terms of their

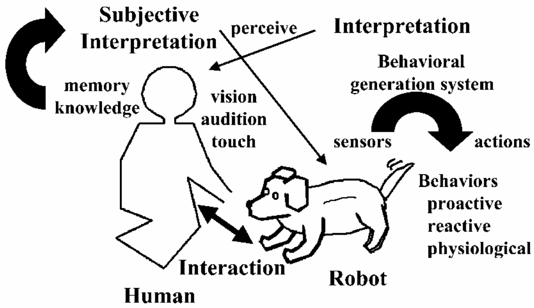


Figure 2 Illustration of Human Robot Interaction

relationship with humans: (World Robotics, 2003) performance robots, (Tucker, 1995) tele-operated performance robots, (Silverstein, et al., 2003) operation, building, programming and control robots, and interactive autonomous robots.

2.1.1. Performance robot

Performance robots are able to perform movements that express meanings to humans, mostly for fun. Performance robots have a long history, as explained in the previous chapter. Mechanical puppets had already been developed in Switzerland in the 18th Century that could play an organ or write pictures and letters. Karakuri dolls were developed to perform dances, magic and so on in Japan during the same era. Recently, a lot of performance robots have been used at exhibitions, in museums and in amusement parks such as Disney Land and Universal Studios. The replica of King Kong and the robotic dinosaurs on the Jurassic Park Ride at Universal Studios are famous examples. The Spring Show at the Bellagio Hotel and Casino in Las Vegas, USA, is another interesting example of performance robots. These robots were developed by SARCOS (http://www.sarcos.com/entprod.html). The field of animatronics yields other interesting examples.

Animal type robots have been used in many movies, such as "Deep Blue Sea," "Perfect Storm," and "Anaconda" (http://www.edgefx.com/). In Japan, robotic fishes can be seen swimming in an aquarium (Terada, 2000; Terada et al., 2004). Recent humanoid robots such as Honda's ASIMO and Sony's QRIO can be included in this category (Hirai, 1998; Kuroki et al., 2002). One performance robot can amuse a sizeable audience at any time. However, their movements will probably be preprogrammed and mostly repetitive, and so they are not usually very interactive with humans. A high degree of complexity is important in performance robots in order to keep humans amused.

2.1.2. Tele-operated performance robots

Tele-operated performance robots are controlled remotely by a hidden operator. Their movements can appear reactive to their audience or to the humans who interact with them because the operator senses their current actions and sends commands to the robot to simulate reactive behavior. At exhibitions or amusement parks for example, human-type robots are used as tele-operated performance robots. Ford used a humanoid robot developed by SARCOS at all auto show exhibitions in 1995, where an operator wearing a sensor-suit controlled the robot.

2.1.3. Operating, Building, Programming and Controlling a Robot

Operating, building, programming and controlling robots give fun and joy to humans. The human can watch the performance of the robot that he or she is operating. A simple example of this is the "UFO catcher" at amusement centers, in which the user's hand controls an X-Y stage to capture an object. Building and programming a robot is also included in this category. Contests between robots such as Micro-mouse, RoboCup, and RoboOne are popular examples (Kitano et al., 1997). RoboOne (http://www.robo-one.com/) is a new robot game where two operators remotely control two humanoid robots to fight each other by wrestling. LEGO Mindstorms and I-Blocks are other examples of human interactive robots. Because building and programming a robot can stimulate children's creativity, this activity combines entertainment with education, and is often referred to as "edutainment" (Papert, 1993; Druin et al., 2000; Lund et al., 1998; Lund et al., 2004).

2.1.4. Interactive autonomous robots

Interactive autonomous robots interact with humans in the physical world. There are verbal and non-verbal communications depending on the functions of the robots (Kanda et al., 2004; Fukuda et al., 2004). Fig. 2 shows an example of an interaction between a human (right) and an interactive autonomous robot (left) which has the appearance of a dog. An interactive autonomous robot behaves autonomously using various kinds of sensors and actuators and can react to stimulation by its environment, including interacting with a human. The human perceives the behavior of the interactive autonomous robot by using his senses. Interaction with the robot produces mental stimulation in the human. He then interprets the meaning of the robot's behavior and evaluates the robot with reference to his own knowledge and experiences. *A priori* knowledge has a significant influence on the interpretation and evaluation. When a human interacts with a robot over an extended period, he gradually learns about the robot. The acquired knowledge of the human then has a great influence on his interpretation and evaluation to the robot. If the robot has a learning function, it can also learn about the human. Then, there is change in the relationship between them. Autonomy and intelligence are key technologies in this category. Contrary to robots in other categories, the interactions between the human and the robot are mostly personal.

Recent research has identified additional roles for interactive autonomous robots other than entertainment (Mayor et al., 2002; Bischoff et al., 2002; Bischoff et al., 2004; Dario et al., 1998; Hans et al., 2002; Rogalla et al., 2002; Wosch, 2002; Fujie et al., 1998; Baum et al., 1984; Gammonley et al., 1991; Shibata et al., 1996; Shibata et al., 2001(a), 2001(b); Wada et al., 2002(a), 2002(b); Saito et al., 2002(a), 2002(b); Werry et al., 1999; Yokoyama, 2002; Libin et al., 2002; Wada et al., 2004(a), 2004(b); Libin et al., 2004; Fujita, 2004; Fujita et al., 1997; Yamamoto et al., 2002; Ozaki et al., 1998; Miyake et al., 2002; Haga, 2002; Onishi, 2002; Brooks et al., 1998; Hashimoto et al, 1998; Hara et al., 1998; Breazeal, 2002; Fukuda et al., 2002; Kanda et al., 2002; Watanabe et al., 2002; Fujita, 2002; Nakata et al., 2002; Menzel et al., 2000). Human interactive robots for guiding people at museums and exhibitions can communicate with visitors while providing a source of fun (Mayor et al., 2002; Bischoff et al., 2002; Bischoff et al., 2004). Human interactive robots are also used in hospitals, in institutions for the elderly and in homes for the elderly (Dario et al., 1998; Hans et al., 2002; Rogalla et al., 2002; Wosch, 2002; Fujie et al., 1998; Baum et al., 1984; Gammonley et al., 1991; Shibata et al., 1996; Shibata et al., 2001(a), 2001(b); Wada et al., 2002(a), 2002(b); Saito et al., 2002(a), 2002(b); Werry et al., 1999; Yokoyama, 2002; Libin et al., 2002; Wada et al., 2004(a), 2004(b); Libin et al., 2004; Fujita, 2004) Some of them help people by providing physical assistance, while others can help to heal the human mind. Robot-assisted therapy at hospitals and robot-assisted activity at institutions for the elderly are good examples (Baum et al., 1984; Fujita, 2004).

2.2. Duration of Interaction

Methods of human-robot interaction can be classified into two categories in terms of duration of interaction: short-term interactions and long-term interactions (Shibata et al., 2001(c)).

2.2.1. Short-term interaction

When a human interacts with a robot during a demonstration at an exhibition, a museum or a similar event, he acquires his first impressions of the robot in a very short time-scale. The appearance of the robot has a large influence on subjective interpretation of the behavior of the robot and in subjective evaluations of the short term interaction. For example, in the case of a human-type robot most people expect similar behavior and similar reactions to certain stimulation by the subject.

2.2.2. Long-term interaction

A human can interact with a robot over a prolonged period or even live together if the robot shares his home or is stationed in a hospital, in a nursing home, in a school or so on. The human interacting with the robot gradually acquires some knowledge on the robot by his learning ability. If the robot always displayed the same reaction or behavior during these interactions, the human would soon become bored with the robot and would quickly discontinue his relation with it. Therefore, it is important that the robot has some learning function to avoid the human becoming bored by the interaction. At the same time, in order to maintain the relation between human and robot, the robot should be robust and durable for long-term use. In addition, the robot has to be safe and easy to maintain by the human.

2.3. Design Issues of Human Interactive Robots for Psychological Enrichment

Human interactive robots for psychological enrichment are a new industry that has arisen from simple electronic dogs and cats. These robots can serve as both a technological playground and (potentially) as a platform for consumer electronics. Human interactive robots offer a proving ground where a diversity of electrical, mechanical and computer engineers can test, develop and apply their latest technologies. Leading technologies are emerging here that will later transfer to other application areas.

In practical applications, human interactive robots for psychological enrichment should have functions that combine an ecological balance with the purpose of the robot and also have an eye to cost (Petroski, 1996)(Pfiefer et al., 1999).

2.3.1. Appearance

The physical appearance of a robot has a significant influence on the subjective interpretation and evaluation of the robot's behavior by an interacting human, especially during short-term interactions. Humans tend to display some bias to a robot that is associated with its appearance (Shibata et al., 2001(c); Pfiefer et al., 1999; Shibata et al., 1997(a), 1997(b); Tashima et al., 1998; Shibata et al., 1999(a), 1999(b); Shibata et al., 2000). There are four principal categories of appearance: human type, familiar animal type, unfamiliar animal type and imaginary animals/new character type. However, the distinctions between them are not always clear and some categories can be combined to avoid bias by humans (Kanda et al., 2004).

a) Human Type:

The appearance of such a robot is similar to a human (Hirai, 1998; Kuroki et al., 2002; Lund, 2004; Kanda et al., 2004; Fukuda et al., 2004; Mayor et al., 2002; Bischoff et al., 2002; Bischoff et al., 2004; Dario et al., 1998; Brooks et al., 1998; Hashimoto et al, 1998; Hara et al.,1998; Breazeal, 2002; Fukuda et al., 2002; Kanda et al., 2002; Watanabe et al., 2002). Some robots have the upper torso of a human body on a mobile robot. The behavior of the human-type robots can be derived from humans, since humans can then easily understand facial expressions, gestures and so on. However, when the humans interacting with the robot compare the robot's behavior with that of humans, the humans tend to be severe in evaluating the robot.

b) Familiar Animal Type:

Familiar animals include such creatures as dogs and cats that are common as pets, so the designers of these robots can easily transfer the behavior of the modeled animal (Terada, 2000; Terada et al., 2004; Libin et al., 2004; Fujita, 2004; Haga, 2002; Onishi, 2002; Shibata et al., 1997(a), 1997(b); Tashima et al., 1998; Shibata et al., 1999(a), 1999(b); Shibata et al., 2000). However, some people have a bias towards a particular type of pet and they might apply this bias to a robot that uses this type of animal as a model. In addition, people compare the robot with the animal on which it is modeled, and they tend to be severe in their evaluation of such a robot, as in the case of human type (Fujita, 2004; Brooks et al., 1998).

c) Unfamiliar Animal Type:
Unfamiliar animals include such creatures as seals, penguins, bears and whales (Terada, 2000; Terada et al., 2004; Yamamoto et al., 2002; Shibata et al., 2001(c); Shibata et al., 2000). Most people know something about the unfamiliar animals, but they are not totally familiar with them in detail and have probably rarely interacted with them before. Therefore, people can accept robots whose appearance is modeled on an unfamiliar animal more easily (Brooks et al., 1998).

d) New Character/Imaginary Animal Type:
If people have an existing bias towards a new character or an imaginary animal such as a cartoon character, this bias may be applied to the evaluation of the robot. If the bias is positive, the value of the robot is improved, regardless of the quality and functions of the robot. In terms of scientific research, it is difficult to deal with characters that many people exhibit some bias about from the beginning. However, if people do not have any preconceptions about a new character or an imaginary animal, the designer of the robot can avoid its appearance being an influencing factor (Fujita, 2002; Nakata et al., 2002).

2.4. Hardware and Software

Human interactive robots are designed to perceive their environment, especially an interacting human, so the use of sensors is very important. Various kinds of sensors can be applied to a robot to mirror human senses (Blauert, 1997; Braitenberg, 1984; Brooks, 1989). Unlike industrial robots, accuracy is not always important in some interactive applications. However, some sensors require a lot of computational power and consume a lot of energy, so the ecological balance between the performance of the sensors and the purpose of the robot is of great importance. Durability is another factor that needs to be considered.

Actuators are keys to the behavior of robots. Small, powerful, light, durable actuators are desirable. Because this type of robot interacts with humans, the sound (or mechanical noise) generated by the actuators has to be carefully considered.

As for their structure, appearance as well as size and weight need to be carefully considered, depending on the specific application. As mentioned above, we tend to display bias in interactions between humans and robots. Thus, sensors, actuators and structure affect the way we interact and communicate with these robots.

The battery that powers the robot and its associated charger have an influence on the life of robot, its appearance and the way that it interacts with humans. Humans expect robots to be able to interact with them for some period of time and the robot has to continue behaving normally throughout this time. If the robot can be recharged automatically, then this is no problem. Otherwise, the method of charging has to be easy for the human to carry out, because charging is like caring for the robot. Batteries and actuators tend to cause a lot of heat.

In terms of computational capability, the processors and the network have to be correctly specified and well designed. Most interactive robots feature distributed computation. Energy consumption and heating problems have to be considered carefully for long-term use.

There are several ways of implementing the control architecture of the robot. A robot can be given some prior knowledge in a top-down approach and/or it can embody reactive behavior with behavior-based control (Braitenberg, 1984; Brooks, 1989; Brooks, 1999; Fukuda et al., 1994). In order to establish a friendly relationship with humans, functions that enable adaptation, learning and even evolution are keys (Picard, 1997; Carter, 1998; Pashler, 1999; Andreassi, 2000; Trappl et al., 2002; Holland, 1992). The intelligence of a robot emerges through interaction and is visible to an interacting human.

Safety should be considered from the viewpoints of both hardware and software.

3. An Example of Robot Assisted Activity Using a Human Interactive Robot

3.1. Mental Commit Robot

Mental commit robots are not intended to offer people physical work or service (Shibata et al., 1996; Shibata et al., 2001(a), 2001(b); Wada et al., 2002(a), 2002(b); Saito et al., 2002(a), 2002(b); Shibata et al., 2001(c); Shibata et al., 1997(a), 1997(b); Tashima et al., 1998; Shibata et al., 1999(a), 1999(b); Shibata et al., 2000; Mitsui et al., 2002). Their function is to engender mental effects, such as pleasure and relaxation, in their role as personal robots. These robots act independently with purpose and with ‘motives’ while receiving stimulation from the environment, as with living organisms. Actions that manifest themselves during interactions with people can be interpreted as though the robots have hearts and feelings.

3.2. Previous Process

A basic psychological experiment was conducted on the subjective interpretation and evaluation of robot behavior following interactions between robots and people (Shibata et al., 2001(c)). This showed the importance of appropriately stimulating the human senses and extracting associations. Sensor systems, such as visual, aural and tactile senses for robots, were studied and developed. A plane tactile sensor using an air bag was developed to cover the robot in order to enhance bodily contact between people and robots. This can detect position and force when people touch the robot, and at the same time, it allows people to feel softness. Dog, cat and seal robots were developed using these sensors.

3.3. Robot Assisted Therapy and Activity

Interaction with animals has long been known to be emotionally beneficial to people. The effects of animals on humans have been applied to medical treatment. Especially in the United States, animal-assisted therapy and activities (AAT&AAA) are becoming widely used in hospitals and nursing homes (Baum et al., 1984; Gammonley et al., 1991). AAT has clear goals set out in therapy programs designed by doctors, nurses or social workers, in cooperation with volunteers. In contrast, AAA refers to patients interacting with animals without particular therapeutic goals, and depends on volunteers. AAT and AAA are expected to have 3 effects:

(1) Psychological effect (e.g. relaxation, motivation)
(2) Physiological effect (e.g. improvement of vital signs)
(3) Social effect (e.g. stimulation of communication among inpatients and caregivers)

However, most hospitals and nursing homes, especially in Japan, do not accept animals, even though they admit the positive effects of AAT and AAA. They are afraid of negative effects of animals on human beings, such as allergy, infection, bites, and scratches.

Recently, several research groups have tried robot assisted therapy and activity (RAT&RAA). Dautenhahn has used mobile robots and robotic dolls for therapy of autistic children (Werry et al., 1999). For example, robot-assisted activity that uses commercialized animal type robots (such as AIBO, NeCoRo, etc.) has been tried (Yokoyama, 2002; Libin et al., 2002;

Hashimoto et al, 1998; Fujita, 2004). Yokoyama used AIBO in a pediatrics ward, and observed the interaction between children and pointed out that the initial stimulus received from AIBO was strong. However, the long term stability was quite weak, compared with living animals. In other words, when patients meet AIBO for the first time, they are interested in it for a short while. However, relaxation effects such as those obtained from petting a real dog are never achieved with AIBO.

We have proposed Robot-Assisted Therapy and Activity since 1996 (Shibata et al., 1996; Shibata et al., 2001(a), 2001(b); Wada et al., 2002(a), 2002(b); Saito et al., 2002(a), 2002(b)). Major goals of this research are follows:

(1) Investigation of psycho-physiological influences of Human-Robot interaction, including long-term interaction
(2) Development of design theory for therapeutic robots
(3) Development of methodology of RAT & RAA suitable for the subjects

The seal robot named Paro have been designed for therapy (Fig.3), and used at a pediatric ward of university hospital (Shibata et al., 2001(a)). The children's ages were from 2 to 15 years, some of them having immunity problems. During *11 days* observation, the children's moods improved on interaction with Paro, encouraging the children to communicate with each other and caregivers. In one striking instance, a young autistic patient recovered his appetite and his speech abilities during the weeks when Paro was at the hospital. In another case, a long-term inpatient who felt pain when she moved her body, arms, and legs, and could not move from her bed. However, when Paro was given to her, she smiled and was willing to stroke Paro. A nurse said that Paro had a rehabilitative function as well as a mental effect.

In the robot assisted activity for elderly people in this paper, a mental commit seal robot known as "Paro" was also used (Fig. 3).

3.4. Seal Robot "Paro"

The appearance was designed using a baby harp seal as a model, and its surface was covered with pure white fur. A newly-developed plane tactile sensor (Shibata, 2004(a)) was inserted between the hard inner skeleton and the fur to create a soft, natural feel and to permit the measurement of human contact with the robot. Whiskers are touch sensors, too. The robot is equipped with the four primary senses; sight (light sensor), audition (determination of sound source direction and speech recognition), balance and the above-stated tactile sense. Its moving parts are as follows: vertical and horizontal neck movements, front and rear paddle movements and independent movement of each eyelid, which is important for creating facial expressions. The robot operates by using the 3 elements of its internal states, sensory information from its sensors and its own diurnal rhythm (morning, daytime and night) to carry out various activities during its interaction with people.

The studies have been conducted using questionnaires given out at exhibitions held in six countries; Japan, U.K., Sweden, Italy, Korea and Brunei, in order to investigate how people evaluate the robot. The results showed that the seal robot widely accepted beyond the culture (Shibata et al., 2002(a), 2002(b), Shibata et al., 2003(a), 2003(b); Shibata et al., 2004(b), 2004(c), Shibata et al., 2004(d)).


Figure 3: Seal Robot: Paro

3.5. Robot Assisted Activity for Elderly People

Seal Robots, Paro, has been used at a day service centre for five weeks, and at a health service facility for the aged for more than a year. The day service center is an institution that aims to decrease nursing load for a family by caring for elderly people during the daytime (9:00-15:30). On the other hand, the health service facility is an institution that aims to rehabilitate elderly people during their stay in the facility.

In both institutions there was little communication and the atmosphere was gloomy. In addition, caregivers felt difficulty in communication with their charges because of a lack of common topics of discussion.

3.5.1. Method of Interaction

At the day service centre, Paro was given elderly people for about 20 minutes, for three days per week over five weeks. People staying the health service facility were given Paro for about one hour, on two days per week from Aug. 2003. The robot was placed on the center of a table, with the patients arranged around it.

Before starting the robot assisted activity, the purposes and procedure was explained to the elderly people to receive their approval. All of the subjects were women in both experiments. There were 23 subjects aged between 73 to 93 years old at the day center, and 14 subjects aged between 77 to 98 years old at the health service facility.

3.5.2. Methods of Evaluation

In order to investigate the effects on the elderly people before and after interaction with Paro, the following three types of data and additional information were collected.

(1) Face scale (Figure 4)
(2) Geriatric Depression Scale (GDS)
(3) Urinary tests Comments of nursing staff

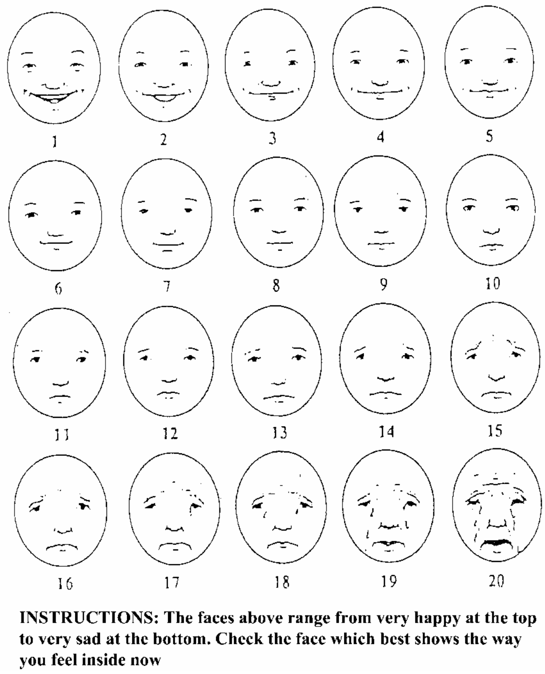


Figure 4: Face Scale

The original Face Scale (Lorish et al., 1986) contains 20 drawings of a single face, arranged in rows, with each face depicting a slightly different mood state. They are arranged in decreasing order of mood and numbered from 1 to 20, with 1 representing the most positive mood and 20 representing the most negative. However, sometimes the subjects are confused by the original face scale because it contains too many similar images. Thus, the scale was simplified by using seven images #1, 4, 7, 10, 13, 16, and 19 from the original set. The original face scale was used at the day service center, and the simplified one used at the facility.

The original GDS (Yesavage, 1988) is a 30-item instrument developed from 100 popular questions commonly used to diagnose depression. A 15-item short version has also been validated. In this research, we used the short version that was translated in Japanese by Muraoka, et al. The scale is in a yes/no format. Each answer counts one point; scores greater than 5 indicate probable depression.

Regarding urinary tests, we examined the change in stress reaction of elderly by measuring urine 17 - Ketosteroid sulfates (17-KS-S) and 17-hydroxycorticosteroids (17-OHCS) values before and after the introduction of Paro. The 17-KS-S value, indicating the restorative degree to the stress, has a high value in healthy individuals (Nishikaze et al., 1995). The 17-OHCS value, indicating the stress load degree, rises at the stress (Selye, 1970; Furuya et al., 1998), and ratio of 17-KS-S/17-OHCS indicates an inclusive living organisms reaction (Furuya et al., 1998).

Moreover, mental impoverishment of caregivers was investigated by using Burnout scale. Burnout is a syndrome where nurses lose all concern or emotional feelings, for the persons they work with and come to treat them in detached or even dehumanized manner. This occurs in nurses who have to care for too many people with continual emotional stress (Maslach, 1976). The Burnout Scale is a questionnaire that consists of 21 items. These items represent three factors such as body, emotions and mental impoverishment. Each item is evaluated over seven stages. If total average score of the items is 2.9 or less, people are mentally and physically healthy, and mentally stable. If the score is 3.0-3.9, the symptoms of Burnout are present. People are judged to fall into the Burnout category if the score is 4.0 or more.

3.5.3. Results of Evaluation

a) Day service center

Figure 6 indicates the average face value (low score – positive mood, high score – negative mood) of 12 people at the day service center. Average scores before interaction varied from about 5.3 to 3.0. However, scores after interaction were constant at about 3.0 for five weeks. Moreover, the sixth week, when Paro had been removed, was higher than the score after interaction with Paro. Thus, interaction with Paro improved the mood state of the subjects, and its effect was unchanged throughout during the five weeks of interaction.

Table I shows a result of urinary test. The participant's 17-KS-S values and ratios of 17-KS-S/17-OHCS were increased after introduction of Paro. Therefore, we consider that RAA improved the ability to in the elderly to recover from stress.

Regarding the comments and observations of the caregivers, interaction with Paro made the elderly people more active and communicative, both with each other and caregivers (Figure 5). In an interesting instance, an elderly woman who rarely talked with others began communicating after

Figure 5: Interaction between Elderly People and Seal Robot at a day service center

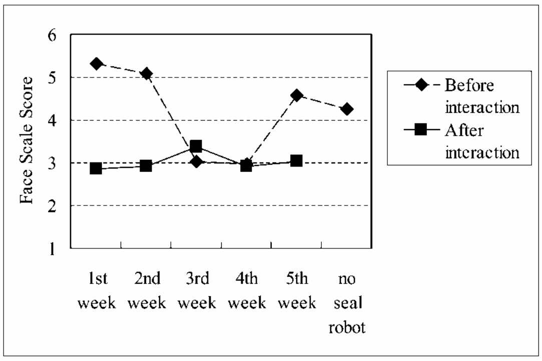


Figure 6: Change of Average Face Scale Scores of 12 subjects over 6 Weeks at the day service center (Score: 1=best mood, 20=worst mood)

Table I: Average values of hormones in urine of 7 subjects before and after introduction of Paro to the day service center

	Before	After	
17-OHCS	8.35±2.87	9.17±3.33	ns
17-KS-S	1.25±0.88	2.41±2.23	*
17-KS-S/17-OHCS	0.14±0.07	0.34±0.45	ns

n=7 Average ± SD

Wilcoxon signed rank test * $p<0.05$

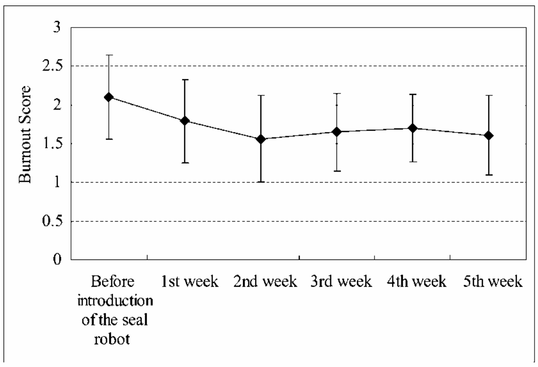


Figure 7: Change of Average Burnout score of the 6 caregivers for 6 weeks at the day service center

Figure 8: Interaction between Elderly people and a Seal Robot at a health service facility

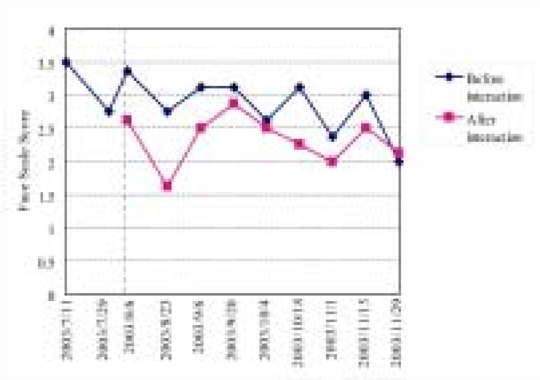


Figure 9: Change of Average Face Scale Scores of 8 Subjects for 5 Months at the health service facility (Score: 1=best mood, 7=worst mood)

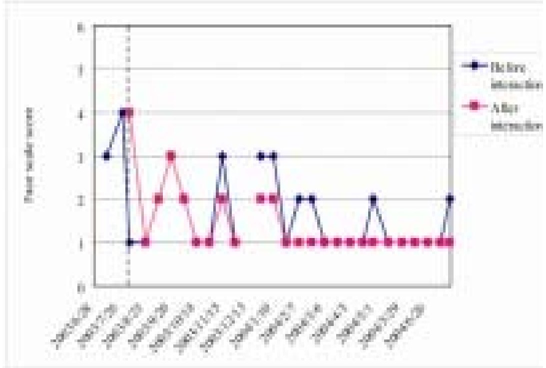


Figure 10: Change of Face Scale Scores of a Subject for one year (Score: 1=best mood, 7=worst mood)

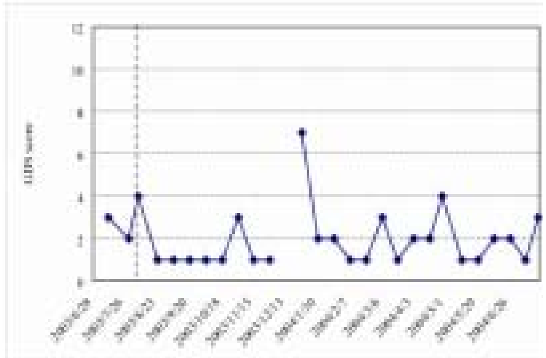


Figure 11: Change of GDS Scores of a Subject for one year (Score: 5 < probable depression)

interacting with Paro. In addition, Paro had an influence on people with dementia. A woman, who had refused to help herself and was frequently forgetful, often laughed and became brighter than usual after playing with Paro. Another elderly woman, who previously wanted to go back home soon, kept staying at the day service center to play with Paro, and looked happy.

Figure 7 shows average Burnout score of the caregivers. Average Burnout score of a week before introduction of Paro was the highest, and then the average score decreased until second week of after the introduction, and kept the small score until the last week. As a statistic analysis, we applied Friedman's test to the Burnout score. We obtained statistically significant changes that the score decreased ($p < 0.05$). As a result, mental impoverishment of the caregivers decreased through RAA.

b) Health service facility

Face scale data were obtained from 8 subjects. The average scores before interaction varied from 3.3 to 2.0 over a 5 month period (Fig.9). However, scores after interaction were almost always lower than those before interaction in each week (except Nov. 29). In particular, a statistically significant difference* was noted in Nov. 15 (Wilcoxon's test: *$p < 0.05$). Therefore, the effects of Paro were unchanging, even though 4 months elapsed.

A case study: Hanako (pseudonym), aged 89, was sociable and comparatively independent. On the first day of the interaction with Paro, she looked a little nervous of the experiment. However, she soon came to like Paro. She treated Paro like her child or grandchild. Her face scale scores after interaction were always lower than before interaction after the first day (Fig.10). Unfortunately, she was hospitalized during Dec. 10 to 26, 2003. When she met Paro for the first time after leaving hospital, she said to Paro "I was lonely, Paro. I wanted to see you again." Her GDS score then improved (Fig.11). To the present, she has continued to join the activity and willingly interacted with Paro.

Caregivers commented that interaction with Paro made the people laugh and become more active. For example, their facial expression changed, softened, and brightened. On the day of activity, they looked forward to Paro, sitting down in their seats before starting interaction. Some people who usually stayed in their room came out and willingly joined the activity. In addition, Paro encouraged the people to communicate, both with each other and caregivers, by becoming their common topic of conversation. Thus, the general atmosphere became brighter.

The elderly people came to love the Paros very much and gave them new names of “Maru” and “Maro”. 3 months after the initial introduction, we added one more Paro to the facility because many others of the elderly had voluntarily joined in the activity. The new Paro was given the name “Hana-chan” by the elderly. Moreover, the Paros have been widely accepted by caregivers, making a home for Paros in the facility.

Generally speaking, people often lose interest in things such as toys, after interacting with them several times. However, regarding interaction with Paro, the elderly people did not lose interest, and its effect on them showed up through one year. In addition, no breakdown and accident occurred by now. Paro fulfill its durability and safety of the robot, which are very important when it interacts with human beings for long-term.

More details of the results of the experiments at the day service center and the health service facility are explained elsewhere. (Saito et al., 2002; Wada et al., 2002, 2004)

4. Conclusions

In this paper we present an overview of human interactive robots for psychological enrichment. Human-robot interactions are explained in terms of the relationship between humans and robots, in terms of the duration of interactions and in terms of design issues affecting human interactive robots for psychological enrichment.

The results of experiments of robot assisted activity for elderly people showed that, a human interactive robot has the potential to enrich people psychologically, physiologically, and socially.

Human interactive robots have a very different character from industrial robots. Since human interactive robots are evaluated by humans mostly in terms of subjective measures, these robots have the potential to engender subjective values in humans.

Acknowledgements

We would like to thank the staff members of "Ousuikai Hanamuro Day Service Center" and “Toyoura” health service facility for their cooperation to our experiment.

References

J. L. Andreassi, Psychophysiology: Human Behavior and Physiological response, Fourth Edition, Lawrence Erlbaum Associates, 2000

M. M. Baum, N. Bergstrom, N. F. Langston, L. Thoma, Physiological Effects of Human/Companion Animal Bonding, Nursing Research, Vol. 33. No. 3, pp.126-129, 1984

R. Bischoff and V. Graefe, Dependable Multimodal Communication and Inteaction with Robotic Assistants, Proc. of the IEEE Int’l Workshop on Robot and Human Interactive Communication, pp. 300-305, 2002

R. Bischoff and V. Graefe, HERMES – a Versatile Personal Robotic Assistant, Special Issue on Human Interactive Robots for Psychological Enrichment, Proceedings of the IEEE, 2004 (to appear)

J. Blauert, Spacial Hearing; The Psychophysics of Human Sound Localization (Revised Edition), The MIT Press, 1997

V. Braitenberg, Vihicles; Experiments in Synthetic Psychology, The MIT Press, 1984

C. L. Breazeal, Designing Sociable Robots, The MIT Press, 2002

R. A. Brooks, A Robust Layered Control System for a Mobile Robot, IEEE Jour. Of Robotics and Automation, pp. 14-23, 1989

R. A. Brooks, et al., The Cog Project: Building a Humanoid Robot, Proc. of the IARP First Int’l Conf. on Humanoid and Human Friendly Robotics, I-3, pp. 1-36, 1998

R. A. Brooks, Cambrian Intelligence; the Early History of the New AI, The MIT Press, 1999

R. Carter, Mapping The Mind, Univ. of California Press, 1998

P. Dario, et al., New Challenges in the Design of Personal Robots, Proc. of the IARP First Int’l Conf. on Humanoid and Human Friendly Robotics, VI-2, pp. 1-6, 1998

A. Druin and J. Hendler, eds., Robots for Kids; Exploring New Technologies for Learning, Academic Press, 2000

M. Fujie, et al., Daily Life Support System for Elderly – Power-Assisted Walking Support and Walk Rehabilitation, Proc. of the IARP First Int’l Conf. on Humanoid and Human Friendly Robotics, VI-2, pp. 1-4, 1998

M. Fujita and K. Kageyama, An Open Architecture for a Robot Entertainment, Proc. of the First Int’l Conf. on Autonomous Agent, 1997

Y. Fujita, Personal Robot PaPeRo, Jour. of Robotics and Mechatronics, Vol. 14, No. 1, pp. 60-63, 2002

M. Fujita, On Activating Human Communications with Pet-type Robot AIBO, Special Issue on Human Interactive Robots for Psychological Enrichment, Proceedings of the IEEE, 2004 (to appear)

T. Fukuda and T. Shibata, Fuzzy-Neuro-GA based Intelligent Robotics, in Comutational Intelligence, IEEE Press (1994)

T. Fukuda, et al., Generalized Facial Expression of Character Face Based on Deformation Model for Human-Robot Communication, Proc. of the IEEE Int’l Workshop on Robot and Human Interactive Communication, pp. 331-336, 2002

T. Fukuda, M. J. Jung, M. Nakashima, F. Arai and Y. Hasegawa, Facial Expressive Robotic Head System for Human-Robot Communication and Its Application in Home Environment, Special Issue on Human Interactive Robots for Psychological Enrichment, Proceedings of the IEEE, 2004 (to appear)

E. Furuya, et al., 17-KS-Sulfate as a Biomarker in Psychosocial Stress, Clinical pathology, Vol. 46, No.6, pp.529-537, 1998.

J. Gammonley, J. Yates, Pet Projects Animal Assisted Therapy in Nursing Homes, Journal of Gerontological Nursing, Vol.17, No.1, pp.12-15, 1991

Y. Haga, WonderBorg and BN-1, Jour. of Robotics and Mechatronics, Vol. 14, No. 1, 68-72, 2002

M. Hans, et al., Robotic Home Assistant Care-O-bot: Past – Present – Future, Proc. of the IEEE Int’l Workshop on Robot and Human Interactive Communication, pp. 380-385, 2002

F. Hara, et al., Personality Characterization of Animate Face Robot through Interactive Communication with Human, Proc. of the IARP First Int’l Conf. on Humanoid and Human Friendly Robotics, IV-1, pp. 1-10, 1998

S. Hashimoto et al., Humanoid Robots in Waseda University – Hadaly-2 and WABIAN, Proc. of the IARP First Int’l Conf. on Humanoid and Human Friendly Robotics, I-2, pp. 1-10, 1998

K. Hirai, Humamoid Robot and Its Applications, Proc. of the IARP First Int’l Conf. on Humanoid and Human Friendly Robotics, V-1, pp. 1-4, 1998

J. H. Holland, Adaptation in Natural and Artificial Systems, The MIT Press, 1992

T. Kanda, et al., Development and Evaluation of an Interactive Humanoid Robot Robovie, Proc. of the IEEE Int’l Conf. on Robotics and Automation, 2002

T. Kanda, H. Ishiguro, M. Imai and T. Ono, Development and Evaluation of Interactive Humanoid Robots, Special Issue on Human Interactive Robots for Psychological Enrichment, Proceedings of the IEEE, 2004 (to appear)

H. Kitano, et al., The Robocup ’97 Synthetic Agents Challenge, Proc. of the First Int’l Workshop on RoboCup, 1997

Y. Kuroki, T. Ishida, J. Yamaguchi, M. Fujita, T. Doi, A Small Biped Entertainment Robot, Jour. of Robotics and Mechatronics, Vol. 14, No. 1, pp. 7-11, 2002

E. Libin and A. Libin, Robotherapy: Definition, Assessment, and Case Study, Proc. of the 8th Int’l Conf. on Vertual Systems and Multimedia, pp. 906-915, 2002

A. Libin and E. Libin, Person-Robot Interaction From the Robopsychologists Point of View: The Robotic Psychology and Robotherapy Approach, Special Issue on Human Interactive Robots for Psychological Enrichment, Proceedings of the IEEE, 2004 (to appear)

C. D. Lorish, R. Maisiak, The Face Scale: A Brief, Nonverbal Method for Assessing Patient Mood, Arthritis and Rheumatism, Vol.29, No.7, pp.906-909, 1986.

H.H. Lund and O. Miglino, Evolving and Breeding Robots, Proc. of the First European Workshop on Evolutionary Robotics, Springer-Verlag, 1998

H. H. Lund, Modern Artificial Intelligence for Human-Robot Interaction, Special Issue on Human Interactive Robots for Psychological Enrichment, Proceedings of the IEEE, 2004 (to appear)

L. Mayor, et al., Improving the Expressiveness of Mobile Robots, Proc. of the IEEE Int’l Workshop on Robot and Human Interactive Communication, pp. 325-330, 2002

C. Maslach, Burned-out, Human Behavior, Vol.5, No.9, pp.16-22, 1976.

P. Menzel and F. D’Aluisio, Robo Sapience: Evolution of a New Species, The MIT Press, 2000

T. Mitsui, T. Shibata, K. Wada, and K. Tanie, Psychophysiological Effects by Interaction with Mental Commit Robot, Vol. 14, No. 1, pp. 20-26, 2002

N. Miyake, T. Shiina, M. Oshiro and Y. Matsuda, Interactive Simulation Ride, Jour. of Robotics and Mechatronics, pp. 64-67, 2002

T. Nakata, T. Mori, T. Sato, Analysis of Impression of Robot Bodily Expression, Jour. of Robotics and Mechatronics, Vol. 14, No. 1, pp. 27-36, 2002

O. Nishikaze, et al., Distortion of Adaptation (Wear & Tear and Repair & Recovery)-Urine 17- KS-Sulfates and Psychosocial Atressin Humans-Job Stress Res, Vol.3, pp. 55-64, 1995.

T. Onishi, POO-CHI, Jour. of Robotics and Mechatronics, Vol. 14, No. 1, pp. 76, 77, 2002

F. Ozaki, et al., Beach Ball Volley Playing Robot, Proc. Of the IARP First Int’l Conf. on Humanoid and Human Friendly Robotics, VI-1, pp. 1-9, 1998

S. Papert, The Childrens Machine; Rethinking School in the Age of the Computer, Basic Books, 1993

H. E. Pashler, The Psychology of Attention, The MIT Press, 1999

H. Petroski, Invention by Design, Harvard University Press, 1996.

R. Pfiefer and C. Scheier, Understanding Intelligence, The MIT Press, 1999

R. Picard, Affective Computing, MIT Press, 1997

O. Rogalla, et al., Using Gesture and Speech Control for Commanding s Robot Assistant, Proc. of the IEEE Int’l Workshop on Robot and Human Interactive Communication, pp. 454-459, 2002

T. Saito, T. Shibata, K. Wada, K. Tanie, Examination of Change of Stress Reaction by Urinary Tests of Elderly before and after Introduction of Mental Commit Robot to an Elderly Institution, Proc. of AROB, 2002(a)

T. Saito, T. Shibata, K. Wada, K. Tanie, Change of Stress Reaction by Introduction of Mental Commit Robot to a Health Services Facility for the Aged, Proc. of Joint 1st Int’l Conf. on SCIS and ISIS, paper number 23Q1-5, in CD-ROM Proc., 2002 (b)

Selye H.: Stress and aging, Journal of American Geriatric Society, Vol. 18, pp. 669-676, 1970.

T. Shibata, et al., Emotional Robot for Intelligent System - Artificial Emotional Creature Project, Proc. of 5th IEEE Int'l Workshop on ROMAN, pp.466-471, 1996

T. Shibata and R. Irie, Artificial Emotional Creature for Human-Robot Interaction - A New Direction for Intelligent System, Proc. of the IEEE/ASME Int'l Conf. on AIM'97, paper number 47 and 6 pages in CD-ROM Proc., 1997.(a)

T. Shibata, et al., Artificial Emotional Creature for Human-Machine Interaction, Proc. of the IEEE Int'l Conf. on SMC, pp.2269-2274, 1997 (b)

T. Shibata, T. Tashima, and K. Tanie, Emergence of Emotional Behavior through Physical Interaction between Human and Robot, Procs. of the 1999 IEEE Int'l Conf. on Robotics and Automation, 1999 (a)

T. Shibata, T. Tashima, K. Tanie, Subjective Interpretation of Emotional Behavior through Physical Interaction between Human and Robot, Procs. of Systems, Man and Cybernetics, pp.1024-1029, 1999 (b)

T. Shibata, K. Tanie, Influence of A-Priori Knowledge in Subjective Interpretation and Evaluation by Short-Term Interaction with Mental Commit Robot, Proc. of the IEEE Int'l Conf. on Intelligent Robot and Systems, 2000

T. Shibata, et al., Mental Commit Robot and its Application to Therapy of Children, Proc. of the IEEE/ASME Int'l Conf. on AIM'01, paper number 182 and 6 pages in CD-ROM Proc., 2001 (a)

T. Shibata, K. Wada, T. Saito, K. Tanie, Robot Assisted Activity for Senior People at Day Service Center, Proc. of Int’l Conf. on Information Technology in Mechatronics, pp.71-76, 2001 (b)

T. Shibata and K. Tanie, Emergence of Affective Behaviors through Physical Interaction between Human and Mental Commit Robot, Jour. of Robotics and Mechatronics, Vol. 13, No. 5, pp. 505-516, 2001 (c)

T. Shibata, T. Mitsui, K. Wada, and K. Tanie, Subjective Evaluation of Seal Robot: Paro -Tabulation and Analysis of Questionnaire Results, Jour. of Robotics and Mechatronics, Vol. 14, No. 1, pp. 13-19, 2002 (a)

T. Shibata, K. Wada, and K. Tanie, Tabulation and Analysis of Questionnaire Results of Subjective Evaluation of Seal Robot at Science Museum in London, Proc. of the 2002 IEEE Int. Workshop on ROMAN, pp.23-28, 2002 (b)

T. Shibata, K. Wada, and K. Tanie, Subjective Evaluation of Seal Robot at the National Museum of Science and Technology in Stockholm, Proc. of the 2003 IEEE Int. Workshop on ROMAN, 2003 (a)

T. Shibata, K. Wada, and K. Tanie, Subjective Evaluation of Seal Robot at the Japan Cultural Institute in Rome, Proc. of the ICCAS, 2003 (b)

T. Shibata, Ubiquitous Surface Tactile Sensor, 2004 1st IEEE Technical Exhibition Based Conf. on Robotics and Automation Proc. pp. 5, 6, 2004 (a)

T. Shibata, K. Wada and K. Tanie, Tabulation and Analysis of Questionnaire Results of Subjective Evaluation of Seal Robot in Japan, U.K., Sweden and Italy, Proc. of the 2004 IEEE Int. Conf. on Robotics and Automation, pp.1387-1392, 2004 (b)

T. Shibata, K. Wada, K. Tanie, W. K. Chung and Y. Youm, Subjective Evaluation of Seal Robot in Gyeongju, Korea, Submitted to IEEE IECON, 2004 (c)

T. Shibata, K. Wada and K. Tanie, Subjective Evaluation of Seal Robot in Brunei, Proc. of the 2004 IEEE Int. Workshop on ROMAN, 2004 (d)

M. J. Silverstein, N. Fiske, J. Butman, Trading Up: The New American Luxury, Penguin, 2003

T. Tashima, S. Saito, M. Osumi, T. Kudo and T. Shibata, Interactive Pet Robot with Emotion Model, Proc. of the 16th Annual Conf. of the RSJ, Vol. 1, pp.11-12, 1998

Y. Terada, A Trial for Animatronic System Including Aquatic Robots, Jour. of Robotics Society of Japan, Vol. 18, No. 2, pp. 195-197, 2000

Y. Terada and I. Yamamoto, An Animatronic System Including Lifelike Robotic Fish, Special Issue on Human Interactive Robots for Psychological Enrichment, Proceedings of the IEEE, 2004 (to appear)

R. B. Tucker, Win The Value Revolution, Career Press, 1995

R. Trappl, P. Petta, and S. Payr, Emotions in Humans and Artifacts, MIT Press, 2002

K. Wada, T. Shibata, T. Saito, K. Tanie, Robot Assisted Activity for Elderly People and Nurses at a Day Service Center, Proc. of the IEEE Int'l Conf. on Robotics and Automation , 2002 (a)

K. Wada, T. Shibata, T. Saito, K. Tanie, Psychological and Social Effects to Elderly People by Robot Assisted Activity at a Health Services Facility for the Aged, Proc. of Joint 1st Int'l Conf. on SCIS and ISIS, paper number 23Q1-3, in CD-ROM Proc., 2002 (b)

K. Wada, T. Shibata, T. Saito and K. Tanie, Effects of Robot Assisted Activity for Elderly People and Nurses at a Day Service Center, Special Issue on Human Interactive Robots for Psychological Enrichment, Proceedings of the IEEE, 2004 (to appear) (a)

K. Wada, T. Shibata, T. Saito and K. Tanie, Psychological and Social Effects in Long-Term Experiment of Robot Assisted Activity to Elderly People at a Health Service Center fot the Aged, Proc. of IEEE/RSJ IROS 2004, pp. 3068-3073, 2004 (b)

T. Watanabe, et al., InterActor: Speech-Driven Embodied Interactive Actor, Proc. of the IEEE Int'l Workshop on Robot and Human Interactive Communication, pp. 430-435, 2002

I. Werry and K. Dautenhahn, Applying Mobile Robot Technology to the Rehabilitation of Autistic Children, Proc. of 7th Int. Symp. on Intelligent Robotic Systems, pp.265-272, 1999.

The International Federation of Robotics and United Nations Economic Commission for Europe, World Robotics 2003 – Statistics, Market Analysis, Forecasts, Case Studies and Profitability of Robot Investment, 2003

T. Wosch, Robot Motion Control for Assitance Tasks, Proc. of the IEEE Int'l Workshop on Robot and Human Interactive Communication, pp. 524-529, 2002

H. Yamamoto, H. Miyazaki, T. Tsuzuki and Y. Kojima, A Spoken Dialogue Robot, Named Wonder, to Aid Senior Citizens Who Living Alone with Communication, Jour. of Robotics and Mechatronics, Vol. 14, No. 1, pp. 54-59, 2002

J. A. Yesavage, Geriatric Depression Scale, Journal of Psychopharmacology Bulletin, Vol.24, No.4, 1988

A. Yokoyama, The Possibility of the Psychiatric Treatment with a Robot as an Intervention -From the Viewpoint of Animal Therapy-, Proc. of Joint 1st Int'l Conf. on SCIS&ISIS, paper number 23Q1-1, in CD-ROM Proc., 2002.

Practical and Methodological Challenges in Designing and Conducting Human-Robot Interaction Studies

Michael L. Walters, Sarah Woods, Kheng Lee Koay, Kerstin Dautenhahn
{m.l.walters, S.N.Woods, K.L.Koay, K.Dautenhahn}@herts.ac.uk
Adaptive Systems Research Group
School of Computer Science
University of Hertfordshire
College Lane, Hatfield, AL10 9AB

Abstract
Human-robot interaction is a rapidly growing research area which more and more roboticists and computer scientists are moving into. Publications on work resulting from such studies rarely consider in detail the practical and methodological problems encountered. This paper aims to highlight and critically discuss such problems involved in conducting human-robot interaction studies. We provide some examples by discussing our experiences of running two trials that involved humans and robots physically interacting in a common space. Our discussion emphasises the need to take safety requirements into account, and minimise the risk of physical harm to human subjects. Ethical considerations are considered, which are often within a formal or legal framework depending on the host country or institution. We also discuss future improvements for features of our trials and make suggestions as to how to overcome the challenges we encountered. We hope that the lessons learnt will be used to improve future human-robot interaction trials.

1 Introduction

In the course of our research for the COGNIRON Project [2005], we are primarily interested in the research area of Human-Robot Interaction (HRI), in particular with regard to socially interactive robots. An excellent overview of socially interactive robots (robots designed to interact with humans in a social way) is provided in Fong et al. [2003]. As we are primarily studying the human perspective of human-robot interaction, human scaled robots in live trials within a human orientated environment were required.

Fig. 1: Children playing games with the University of Hertfordshire PeopleBot™ at the London Science Museum event in October 2004.

Other researchers that have conducted similar human centred trials with human sized robots include Dario et al. [2001], Severinson-Eklundh et al. [2003], Kanda et al. [2004] and Hinds et al. [2004].

To date, we have conducted two human-robot trials with human scaled PeopleBot™ robots. One trial involved a single robot interacting with groups of children in a game scenario. The trials took advantage of a software evaluation event at the University of Hertfordshire, hosted by the Virtual ICT Empathic Characters (VICTEC) project [VICTEC, 2003]. The other trial involved individual adults interacting with a robot in various contexts and situations, within a simulated domestic (living-room) environment. We have also participated in other displays and demonstrations which have involved robots interacting in the same physical space as one or more humans. In particular, we successfully ran interactive games for groups of up to 40 children at a time, at a major public event at the Science Museum in London [BBC Science News, 2004]. The PeopleBot™ robots have also been demonstrated on several occasions during open days at the University of Hertfordshire. This paper will present some of the methods we have developed and critically discuss the various trials and events we have been involved with to date.

2 Planning, Legal and Safety

Before running a trial involving humans and robots physically interacting, certain legal and ethical issues must be satisfied. At this stage it is good practice, and in the UK a legal requirement under the Management of Health and Safety at Work Regulations 1999, to carry out a risk assessment for all work activities involving employees or members of the public [Crown Copyright, 2003]. These first activities are considered here.

2.1 Legal and Ethics Approval

Many institutions, including the University of Hertfordshire [UPR AS/A/2, 2004], require that an Ethics Committee must give approval for all experiments and trials involving human subjects. Usually, this approval is gained by submitting a (written) description of the trials or experiments to be performed to the committee. The Ethics committee will then consider the proposal, and may modify, request further clarification, ask for a substantial rewrite, or even reject the proposal outright on ethical grounds. In general, the Ethics committee will make possible objections on the following grounds:

Privacy – If video, photographic or records of personal details of the subjects are being made and kept, the committee will be concerned that proper informed consent is given by subjects, any personal records are securely stored and will not be misused in any way. If personal data is to be held on a computer database, then the legal requirements of the Privacy and Electronic Communications Regulations [Crown Copyright, 2003] must be adhered to. If any public use of the video or photographs is to be made for conferences or publicity purposes, then participants must give explicit permission.

Protection of minors and vulnerable adults – In the UK it is a legal requirement (Protection of Children and Vulnerable Adults Order, 2003) that anyone who works with children or vulnerable adults must have their criminal record checked. In the UK, anyone under 18 is classed as a child in this context, and the term vulnerable adult includes the infirm or elderly in a care situation. Regulations in many other countries in Europe are less strict, but if experiments or trials are planned to involve children or vulnerable adults, then any legal implications or requirements must be considered. For example, not gaining the appropriate checking of criminal records could lead to a situation where subjects who are keen to participate in a study need to be turned down. Given the general problem in recruiting a sufficiently large sample of human subjects, this could potentially cause problems.

Mental or emotional stress and humiliation – The trials should not give rise to undue mental or emotional stress, with possible long-term repercussions. Where an experimental situation is actually designed to put a subject under stress intentionally, it may not be possible to avoid stressing the subject. The Ethics Committee will want to be satisfied that if any mental or emotional stress is suffered by subjects, it is justified and that no after effects will be suffered by subjects. In our own studies we were interested in how subjects 'spontaneously', or 'naturally' behaved towards robots, so we had to carefully design the scenarios in order to be on the one hand controlled enough to be scientifically valuable, but on the other hand open enough to allow for relaxed human-robot interactions. It is advised to include a statement in the consent form which points out that the subject can interrupt and leave at any stage during the trial for whatever reasons, if he or she wishes to.

Physical harm – Practically all experiments that involve humans moving will involve some degree of risk. Therefore, any human-robot trials or experiments will pose some physical risk for the subjects. The Ethics Committee will want to be satisfied that the proposal has considered any potential physical risks involved. The subjects' safety is covered in more detail below.

2.2 Safety

Robot-Human Collision Risk - For trials involving humans and robots, the obvious immediate risk is the robot colliding with a human subject, or vice versa. The robots that we used in our trials are specifically marketed for the purpose of human-robot interaction studies. In order to alleviate the risk of the robot colliding with human subjects, two strategies were adopted:

Overriding anti-collision behaviour – The PeopleBot™ robots we use can have several behaviours running at the same time as any top-level program. This is a natural consequence of the PeopleBot™ operating system which follows the principles of the subsumption architecture expounded originally by Rodney Brooks [1991]. Many other commercially available robot systems have similar programming facilities. We always had basic collision avoidance behaviours running at a higher priority than any task level program. This means that no matter what the task level program commands the robot to do, if a collision with an object is imminent, the underlying anti-collision behaviour cuts in. Depending on the form of the hazard and the particular safety behaviour implemented, the robot will either stop or turn away from the collision hazard. The lower priority task level programs include both those that provide for direct or semi-autonomous remote control by Wizard of Oz (WOZ) operators [Maulsby et al. 1998] and also fully autonomous programs. We have found that the sonar sensors used by the PeopleBot™ are very sensitive to the presence of humans. However, some common household objects, especially low coffee tables, are not so readily picked up by the sensors. By judicious placing of objects that are readily sensed, such as boxes, footballs, cushions etc, it is possible to create a trial environment where it is literally impossible for the robot to collide with any object. For example, we adopted this strategy to avoid the robot bumping into the table where the person was sitting (see fig.2).

Fig. 2: A subject sitting at the desk, showing a box placed under the table to create a target for the robot's sonar sensors.

Monitoring by the WOZ operators – Even while the robot is running a fully autonomous program, a WOZ operator (see section 3.1) monitors discreetly what is happening. The robot's underlying safety behaviours include the overriding ability for the WOZ operator to stop the robot immediately by remote wireless link if it is perceived that the robot poses a risk to a human at any time. There is also a large red emergency stop button on the robot, which is hardwired, providing an independent failsafe method to stop the robot. Simply pressing the button cuts the power to all the robot's motors. This is simple enough for non experts to operate, and will work even if the robots control software crashes or fails to respond. Anyone who is physically close enough (i.e. in perceived danger) to the robot can access the button.

In our trials, only during the software development process of the program has it been necessary for WOZ operator or others to initiate a stop; mainly to avoid the robot damaging itself rather than actually posing a threat to humans in the vicinity. During our human-robot interaction trials, the underlying safety behaviour has proved to be both robust and reliable in detecting and avoiding collisions with both children and adults. The actual robot programs have been heavily tested in the physical situations for all the trials we have run. This is necessary as knowing how the robot will respond in all physical circumstances is critical for the safety of the participants in any trial.

For the risk case of a human colliding with the robot, there is little action that can be taken by the robot to avoid a human. The robot moves and reacts relatively slowly, compared to the speeds achievable by a human. Therefore, it is up to the human to avoid colliding with the robot. Luckily, most humans are experts at avoiding collisions and we have found that none of our subjects has actually collided with the robot. In some of the trials with children it has been necessary to advise the children to be gentle or to move more carefully or slowly when near to the robot. We found that children will mostly take notice if the robot actually issues these warnings using the robot's own speech synthesis system.

Other Possible Risks to Participants - Our robot was fitted with a lifting arm, which had a small probability of causing injury to humans. The arm itself was made of coloured cardboard made to look solid, so it looked more dangerous than it actually was. Our main concern about the arm was if the 'finger' was accidentally pointed into a human's face or eyes. This risk was minimised by keeping the arm well below face level even when lifted. Other possible risks to participants that must be considered are those that would be present in any domestic, work or experimental situation. These include things such as irregular or loose floor coverings, trailing cables, objects with sharp or protruding edges and corners, risk of tripping or slipping, etc.

In our trial involving children, small prizes were given during and at the end of each session. We were advised against providing food (i.e., sweets) as prizes, as some children may have had allergies or diabetes which could be aggravated by unplanned food intake. We also never left subjects alone with a robot without monitoring the situation.

3 Experimental Implementation

When running a human-robot interaction trial, the question that must be addressed is how to implement the proposed robot functions and behaviour. There are two main methods for developing suitable robot features, functions and behaviour for trials where we are primarily interested in the human-centred perspective towards the robot or its function.

3.1 Wizard of Oz Methods

It is usually relatively quick to create a scenario and run the robot under direct WOZ operator control. This is a technique that is widely used in HRI studies as it provides a very flexible way to implement complex robot behaviour within a quick time-scale (Robins et al. 2004 and Green et al. 2004). The main advantage is that it saves considerable time over programming a robot to carry out complex interactions fully autonomously. However, we have found that it is very tiring for the WOZ operators to control every aspect of the robot's behaviour, especially in multi-modal interactions and scenarios. It usually requires two operators, one for controlling movement and one for speech, in order to maintain reasonable response times during a trial. It is also difficult to maintain consistency between individual trial sessions. Practise effects are apparent as the operators become better at controlling the robot at the particular task scenario through the course of a series of trials. Practise effects can be minimised by thoroughly piloting the proposed scenario before carrying out 'live' trials.

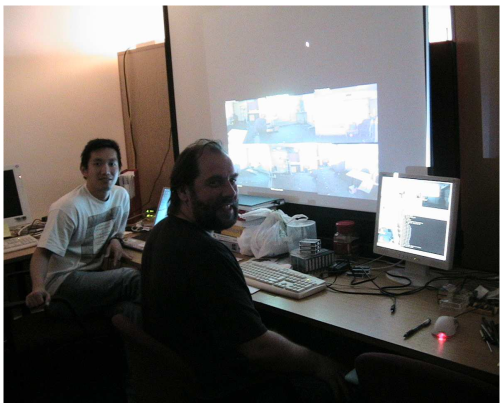


Fig. 3: The Wizard of Oz operators and control room area for human-robot interaction trials at the University of Hertfordshire in 2004.

3.2 Autonomous Robot Control

The other robot control method is to pre-program the robot to run all functions autonomously. Obviously this method overcomes the problems of operator tiredness and consistency, but implementing complex autonomous behaviour is very time-consuming. However, if trials are testing complex human-robot social behaviours, or implementing desired future robot capabilities, it will not be technically feasible at present to program a robot to act fully autonomously. In accordance with the COGNIRON project aims, we are studying scenarios that go "beyond robotics". For this we have to project into the future in assuming a robot companion already exists that can serve as a useful assistant for a variety of tasks in people's homes. Realistically, such a robot does not yet exist.

The PeopleBot™ robots have a sophisticated behaviour based programming API called ARIA [ActivMedia Robotics, 2005]. This provides facilities to develop task control programs, which can be integrated into the ARIA control system. The actual task control program can be assigned a priority, which is lower than the previously mentioned safety behaviours (see section 2.2). Therefore, fundamental safety and survival behaviour, such as collision avoidance, emergency stop etc. will always take precedence over the actual task commands.

In practice, we have found that a mixture of autonomous behaviours and functions, and direct WOZ control provides the most effective means of generating the desired robot's part of the HRI. The basic technique is to pre-program the robot's movements, behaviours or sequences of movements, as individual sequences, gestures or actions that can be initiated by the WOZ operator. In this way the WOZ operator is able to exercise judgement in initiating an appropriate action for a particular situation, but is not concerned with the minute details of carrying that action out. The operator then is able to monitor the action for potential hazard situations and either stop the robot or switch to a more appropriate behaviour. Because the robot is actually generating the individual movements and actions autonomously, better consistency is ensured. Also, the temporal behaviour of a robot under WOZ or autonomous control is likely to differ significantly, so whenever possible and safe, autonomous behaviour is advantageous over remote-controlled behaviour.

Robot program development & *pilot studies*- When developing robot programs, which will be used to implement a HRI trial scenario, it is important to allow enough time to thoroughly practise the programs and scenarios thoroughly before the actual trials take place. Pilot studies should be conducted with a variety of humans, as it is easy for the programmer or operator to make implicit or explicit assumptions about the way that humans will behave in response to a given trial situation. Of course, humans all exhibit unique behaviour and can do unexpected things which may cause the robot program to fail.

The first trials we ran involved interactive game sessions with groups of children. These required the children to play two short games with the robot, a *Rotation game* and a *Wander game*. The game programs ran mostly autonomously, except for starting the respective game programs, and also at the end of each round where a winning child was selected manually by remote control. When developing the interactive game programs for the Science Museum visit, the games ran totally autonomously for the whole of each game session. The Science museum game program was more complex than the previous child group games programs as sensor interpretation was involved. However, because the Science Museum robot game program was fully autonomous, the pre-testing phase had to be much longer. The extra time was needed to empirically find out optimum action and response timings and durations, sensor levels and cues, and refining the program so that it worked properly with all the human test subjects.

For the single adult HRI trials, there were time limitations on setting up and implementing the experiment. The robot behaviour was implemented almost entirely by direct WOZ control (with overriding safety behaviour active). There was also limited time available for practicing the scenarios, which were to be implemented for the study. The only autonomous behaviours used for this study were the wandering behaviour, used for acclimatising the subject to the robot's presence, and the arm lift height, which was used to set the arm to the correct height for picking up special pallets which contained items that would be fetched by the robot at various times during the trial (fig 4).

Fig. 4: The robot, fitted with a hook-like end-effector, was able to fetch small items in special pallets.

The WOZ operators were out of direct sight of robot and subject, and observed the scenarios via network video cameras placed around the room. The images from these were delayed by approximately 0.5 sec. There was also a direct, but restricted, video view from the robot camera which did not have any discernable delay. These factors made providing timely responses (comparable to human responses) to the subject very difficult for the WOZ operators. However, it can be argued that, in the near future at least, this is likely to be true of all robots, and this was a realistic simulation of likely future robot performance.

4 Video Recording

It is desirable to make a complete video record of the trials. Video footage is one of the primary means of gaining results for later analysis and validation of results. They can be used to validate data obtained by other means, e.g. from direct measurement, questionnaire responses, or recorded sensor data. Good video footage can provide time stamped data that can be used, processed and compared with future studies. However, in addition to the obvious advantages of video data, there are some drawbacks that researchers should be aware of at the outset of the design phase. Analysing video footage is an extremely time consuming process and requires thorough training in the application of the scoring procedures, which can be complex. Observations made from video footage are subjective and the observer may portray their own perceptions and attitudes into the date. For this reason, it is essential that a full reliability analysis of video data is carried out involving independent rating and coding by observers who were not involved with the study, and did not meet any of the participants.

4.1 Video camera types

We used two types of video cameras for recording our trials; tripod mounted DV camcorders, and network cameras. The DV camcorders record onto mini DV tape, which must then be downloaded onto a computer hard disk before further analysis can be performed. The network cameras have the advantage that they record directly to a computer's hard disk, so there is no tedious downloading later on. They do require some synchronising, converting and combining, but this can be done automatically in batches overnight. We have found that the DV Camcorders provide a better quality picture than the network cameras, with a synchronised soundtrack. While high quality video may not be strictly necessary for analysis purposes, it does allow high quality still pictures to frame-grabbed from the video recordings, which are invaluable for later writing up, papers and reports. It is also easy to create short videos to incorporate into presentations and demonstrations using standard video editing software.

It is advisable to use at least two camera systems for recording trials or experiments. If one camera fails, than there will another stream of video data available. It should be borne in mind that if a network camera fails, it may also lead to all the network cameras being bought down. Therefore, at least one camera should be a freestanding camcorder type, which stores the video data on (mini DV) tape.

Note, a similar backup strategy is also advisable as far as the robotic platforms are concerned. In our case, we had a second PeopleBot™ in place, in the event that one robot broke down. Having only one robot available for the trials is very risky, since it could mean that a trial had to be abandoned if a robot fails. Re-recruiting subjects and properly preparing the experimental room is a very time-consuming activity, unless a permanent setup is available. This was not the case in our trials, where rooms were only temporarily available for a given and fixed duration (two weeks for the study involving children, 2 months for the adult study). Afterwards the setups had to be disassembled and the rooms had to be transformed back into seminar or conference rooms. This also meant that any phases of the trials could not be repeated. Therefore, it was essential to get it right first time despite the limited preparation time. This is a situation common to a University environment with central room allocations and usually few permanent large laboratory spaces suitable for studies with large human-sized robots.

4.2 Camera Placement

The placement of the cameras should be such that the whole trial area is covered by one or two views. For our first trial, we used two cameras placed in opposite corners of the room, both facing towards the centre of the room. As a result we recorded two views of the centre of the room, but missed out on what was happening at the edges of the room. A better way to position the cameras would have been to point the cameras to the right (or left) of room centre, with only a small view overlap in the centre of the room. This way, the two views also include the outer edges of the room. (See fig.5)

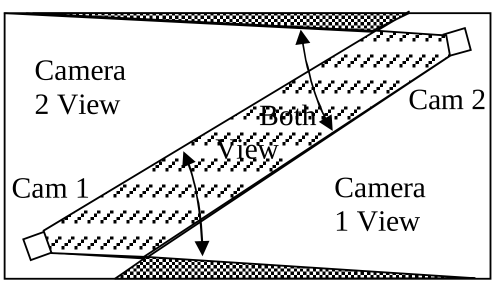


Fig. 5: Diagram showing correct placement of cameras to maximise coverage of room.

It is best to use two cameras to cover the entire area as shown in the diagram, with additional cameras to obtain detailed views of specific areas of interest. For example; when it is known that subjects will have to sit at a certain desk, which is in a fixed position, it is worth setting up an individual camera just to record that position in detail. Setting the correct height of the cameras is important to obtain a good view of the subjects face.

4.3 Distance Measurements

One main aspect of our trials was focused on examining the spatial distances between the robot and human subjects. Video images can be useful in estimating these distances. In both our child group and single adult trials, markings were made on the floor with masking tape to provide a method to estimate the position of the robot and the human subjects within the trial areas. However, these markings were visible to the subjects, and may possibly have influenced the positioning of the human subject during the course of parts of the experimental scenarios.

In the context of a study described by Green et al. [2004], a method was used that involved overlaying a grid of 0.5m squares onto still images of the floor of their trial area for individual frames from their video recordings. This method would allow the positions of the robot and subject to be estimated with a high degree of accuracy if it can be adapted for live or recorded video data. It would provide a semi-transparent grid metric overlaid onto the floor of the live or recorded video from the cameras. The possibility of visible floor markings affecting the positions taken by the subject would not happen. For future trials we will want to use such a 'virtual grid' on the floor of the recorded video data. We are currently evaluating suitable video editing software.

5 Subject's Comfort Level

For the adult trials, we experimented with a method of monitoring how comfortable the subject was while the trials were actually running. We developed a hand held comfort level monitoring device (developed by the first author) which consisted of a small box that could be easily held in one hand (see fig. 6). On one edge of the box was a slider control, which could be moved by using either a thumb or finger of the hand holding the device. The slider scale was marked with a happy face, to indicate the subject was comfortable with the robot's behaviour, and a sad face, to indicate discomfort with the robot's behaviour.

Fig. 6: Photograph of Hand Held Comfort Level Monitoring Device

The device used a 2.4GHz radio signal data link to send numbers representing the slider position to a

PC mounted receiver, which recorded the slider position approximately 10 times per second.. The data was time stamped and saved in a file for later synchronisation and analysis in conjunction with the video material. The data downloaded from the hand held subject comfort level device was saved and plotted on a series of charts. However, unexpectedly, the raw data was heavily corrupted by static from the network cameras used to make video recordings of the session. It has been possible to digitally clean up and recover a useful set of data. A sample of the raw data and the cleaned up version is shown in the figs. 7 and 8.

Many of the comfort level movements correspond to video sequences where the subject can be seen moving the slider on the comfort level device. This confirmed that the filtered files were producing a reliable indication of the comfort level perceived by the subject. For future trials, it is intended to incorporate error checking and data verification into the RF data transfer link to the recording PC in order to reduce problems with static.

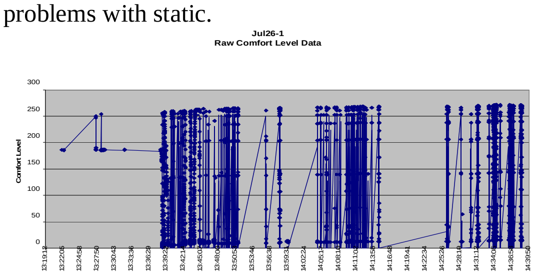


Fig. 5: Raw Data as Received from Handheld Comfort Level Monitoring Device

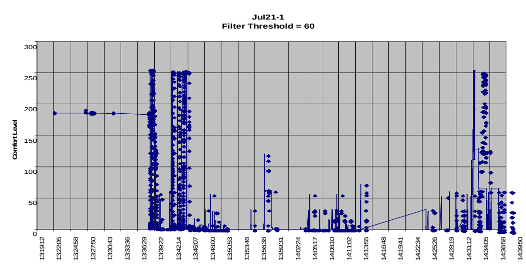


Fig. 6: Digitally Filtered Data from Handheld Comfort Level Monitoring Device

6 Questionnaires

For both interactive trials, subjects were asked to complete questionnaires. For the child-robot interactions only five minutes at the beginning, and five minutes at the end of each session were available. Due to limited time, only basic information was obtained, such as gender, age, approval of computers and robots, and how they liked the interactive session. For the adult study, the questionnaires were much more comprehensive. The time taken for the session typically ranged from 40 minutes to 1 hour. Up to half the time was spent completing questionnaires. The questionnaires covered the subjects' personality traits, demographics, technical experience, opinions towards a future robot companion, how they felt about the two contrasting robot 'personalities' exhibited by the robot during the interaction scenarios, what they liked or disliked about the robot interactions, and how it could be improved, etc.

6.1 Questionnaire Design

Questionnaire design requires training and experience. There are a number of different considerations researchers should take into account before embarking on designing a questionnaire. Firstly, the notion of whether a questionnaire is the best form of data collection should be addressed. For instance, in some situations an interview format might be preferred (e.g. if conducting robot interactions studies with young children that have low reading abilities). A questionnaire is usually completed by the participant alone. This does not allow the researcher to probe for further information they feel may be relevant to the experiment or verify participant responses. However, the advantage of using questionnaires is that they are usually fast to administer, and can be completed confidentially by the participant.

The development of a questionnaire goes through a series of different cycles. Questions that should be considered are:

- Is the questionnaire I am going to use a valid measure (i.e. does it measure what I really want it to measure)
- Is it reliable (i.e. do I get the same pattern of findings if the questionnaire is administered a few weeks later?),
- Have I used value-laden or suggestive questioning (e.g. "Do you think this robot is humanlike?"), compared to neutrally phrased questions (e.g. "What kind of appearance do you think this robot has?")?,
- Do I want to use a highly structured questionnaire or a semi-structured questionnaire, for example where subjects can express their attitudes towards a particular aspect of the robot interaction in more detail?

Some questionnaires are easier to design than others. For example, a questionnaire that enquires about subject demographics must include items that enquire about age and gender. However, even when considering something as simple as age, the researcher must decide whether to use age categories or simply get the subject to write their age in.

The complexity of questionnaire design occurs when a new research domain is being explored, and human-robot interaction is a perfect example of this. There is no such thing as a perfect questionnaire, but careful team planning and pilot testing can ensure that you have the best possible measure. To carry out a pilot test for a questionnaire, the researcher must recruit independent subjects with the same demographics that they hope to include in the real experiment. Sometimes, it is not easy to get volunteers to participate in a pilot test, but obviously the more responses you get, the more certain you can be of what necessary changes need to be made. It is good practise to carry out the pilot study with approximately 5-10 subjects although this depends on the number of conditions etc in the experiment. In addition to asking the pilot subjects to complete the trial questionnaires, it is recommended to ask them directly whether they found any aspects particularly unclear, complicated or irrelevant etc. One could also ask the subjects whether they would change anything about the overall structure or format, and whether there were important questions that you omitted.

A further issue relates to the type of data you will have to analyse. It is important at the design and pilot testing phase to consider the statistical frameworks that you wish to use, as the questions need to be asked in order to fit their requirements as well as the research goals. For example, continuous scales for questionnaire responses lead to very different analytical frameworks compared to categorical (e.g. yes/no) response formats (i.e. interval versus nominal/ordinal data). Although this process can seem time consuming at the outset, it is certainly worth it, as it is impossible to make changes while the trials are running. An error in the questionnaires could possibly invalidate one or more questions, or in the worst case, the whole questionnaire. As highlighted above, no questionnaire is perfect and we discovered this for ourselves in the adult robot-interaction study. Below we give an example of a possible problematic question and a suggested solution:

Example question
Q. Would you like the robot to approach you at a speed that is?

1) Fast
2) Slow
3) Neither fast nor slow

The above question is phrased in an unspecific way, resulting in, whatever answer is given, little quantifiable information about the preferred approach speed. Due to this lack of a reference point, in practice, most subjects are likely to choose answer 3), as most people want the robot to approach at a speed which is 'just right'. An improved way of asking the question could be:

Suggested improvement to question
Q. Did the robot approach you during the trials at a speed that you consider to be?

1) Too fast
2) Too slow
3) About right

Hopefully, results obtained from this improved question would relate a subject's preferred robot approach speed relatively to the actual speed employed by the robot in a trial. If finer graduations of preferred robot approach speeds are desired, then the trial context and situation must be more closely controlled, with multiple discrete stages, with the robot approaching at different speeds at each stage.

Questionnaire Completion - In our trials it was necessary for some of the questionnaires to be completed in the robot trial area. The subject completed the first questionnaires while the robot was wandering around the trial areas in order to acclimatise the subject to the robot's presence. The two post scenario questionnaires were also administered in the trial area, straight after the respective scenarios, while they were fresh in the subject's memory. We were not able to gain access to the trial area to turn the video cameras off during this time, as we wanted to preserve the illusion that the subject and supervisor were on their own with the robot during the trial. However, there were several other questionnaires and forms, which could have been, administered elsewhere. This would have reduced the amount of video tape used per session. Also the WOZ operators had to sit perfectly still and quiet for the duration of these questionnaires. However, a drawback of administering the questionnaires outside the experimental room is that it changes the context, and might distract the subject etc. Such factors might influence the questionnaire results. Thus, there is a difficult trade-off between savings in recording video tape and other data during the trials, and providing a 'natural' and undisturbed experimental environment.

The environmental context is an important consideration for human-robot interaction studies as questionnaire and interview responses, and observational data will vary depending on the experimental set-up. For example, it would not appear to be problematic to complete a participant demographics questionnaire in the experimental room, which in this case was the simulated living room containing the robot. However, when administering a questionnaire that relates to robot behaviour, appearance, personality or the role of future robot design, the robot and room set-up could influence subject responses. For example, in both the child and adult studies subjects completed a questionnaire at the end of the robot interaction scenarios about their perceptions towards a future robot companion. If the intention is that they consider the robot interaction and robot appearance they have just interacted with in the responses (as it was the case in our study), then this is acceptable. However, the researchers must be aware that subject experiences with the robots in the simulated living room are likely to have influenced their responses in some way.

For trials run in 2004 at the Royal Institute of Technology, Sweden, the WOZ and camera operators were in view of the subject while user trials were taking place [Green et al. 2004]. However, the focus of their study was mainly on human-robot dialogue and understanding, command and control of the robot, which may not have been affected by the presence of other people. We have found that when other people are present, then subjects will tend to interact with those other people, as well as the robot. For our single adult interactions, we wanted to observe the subjects reactions as they interacted only with the robot. Thus, while the experimenter in the adult study stayed in the same room as the subject, she deliberately *withdrew* herself from the experiment by sitting in a chair in a corner and reading a newspaper. Moreover, she did not initiate any communication or interaction with the subjects, apart from situations when she had to explain the experiment or the questionnaires to the subject, or when she had to respond to a verbal query from the subject. We opted for this approach since the study targeted a '*robot in the home'* scenario, where it would be likely that a person and robot would spend a considerable amount of time alone together in the environment.

7 Design and Methodological Considerations

At the outset of designing any study there are a number of crucial design and methodological considerations.

First, the research team must decide what the sample composition will be including, individuals, groups, children, adults, students, or strangers from the street. This is important as the interpretation of results will be influenced by the nature of the sample. For example in the current study, we observed quite distinct differences in the interaction styles between groups of children who were familiar with each other, and individual adults who were alone in the room with an experimenter who did interact in the experiment. Also, as with many other studies, the current adult sample were self-selected and were all based at the university (either as staff or students), which could result in a positive or negative bias in the results. It is very difficult to recruit completely randomised samples and there is always a certain amount of self-selection bias in all studies of this design.

Second, the environmental context should be considered, in the sense of whether a laboratory set-up is used or a more naturalistic field study. Different results are likely to emerge depending on the environment chosen. The adult human-robot interaction study involved a simulated living-room situation within a conference room at the University. Although we tried to ensure it was as realistic as possible, subjects still knew it was not a real living room and were likely to have felt monitored by the situation. Ideally it would be best to carry out future robot-human interaction studies in peoples' homes or work places in order to capture more naturalistic responses and attitudes towards the interactions. However, there are advantages for carrying out laboratory based studies as it allows the researchers greater control and manipulation of potential confounding variables. This cannot be done in the naturalistic field, so it is certainly common practise to begin new research protocols in laboratory set-ups.

Cultural differences are also important if the researchers are hoping for widespread generalisation of the findings. However, this is often impractical, highly expensive and time-consuming.

The overall design of experiments is extremely important in terms of whether between-subject groups (independent measures design) or within-subject groups (repeated measures designs) are used. There are advantages and disadvantages associated with both. Between-subject designs involve different subjects participating in different conditions, whereas within-subject designs mean that the same set of subjects take part in a series of different conditions. Between-subject designs are less susceptible to practice and fatigues effects and are useful when it is impossible for an individual to participate in all experimental conditions. Disadvantages include the expense in terms of time and effort to recruit sufficient participant numbers and insensitivity to experimental conditions. Within-subject designs are desirable when there are sensitive manipulations to experimental conditions. As long as the procedures are counterbalanced, biased data responses should be avoided.

A final consideration should be whether the researchers feel the results are informative based on information recorded at one time point. Human-robot interaction involves habituation effects of some kind and it would be highly useful for researchers to be able to follow-up the same sample of subjects over an extended period of time at regular intervals, to determine whether for example, they become more interested/less interested in the robot, more positive/negative towards the robot and so forth.

Human-robot interaction studies are still a relatively new domain of research and are likely to have a high explorative content during initial studies. It took the science of human psychology many years to build up a solid base of methods, techniques and experience, and this process is still going on at the present. The field of human-robot interactions is still in its infancy and carrying out these initial explorative studies implies that there are not likely to be any concrete hypotheses claiming to predict the direction of findings. This would be impossible at the outset of studies if there are not many previous research findings to base predictions on. The nature of exploratory studies means that there are likely to be many different research questions to be addressed and in any one study, it is simply impossible to consider all possible variables that might influence the findings. However, once exploratory studies have been conducted it should allow the researchers to direct and elucidate more concrete and refined research hypotheses for future, more highly controlled studies.

8 Summary and Conclusions

We have discussed our experiences of running two trials that involved humans and robots physically interacting, and have highlighted the problems encountered.

1. When designing and implementing a trial that involves human and robots interacting physically within the same area, the main priority is the human subject's safety. Physical risk cannot be eliminated altogether, but can be minimised to an acceptable level.
2. There are ethical considerations to be considered. Different countries have differing legal requirements, which must be complied with. The host institution may also have additional requirements, often within a formal policy.
3. Practical ways are suggested in which robots can be programmed or controlled in order to provide intrinsically safe behaviour while carrying out human-robot interaction sessions. This complements work in robotics on developing safe robot motion and navigation planners by other partners within the COGNIRON project and elsewhere [Roy and Thrun, 2002]
4. The advantages of different types of video cameras are discussed, and we suggest that if using network based video cameras, it is wise to use at least one videotape-based camera as a backup in case of network problems, and vice versa. We also suggest some (obvious) ways to optimise camera placement and maximise coverage.
5. Similarly, we suggest it is good practice to have a backup robot available.
6. Sufficient time should be allocated to setup the experimental room and test all equipment and experimental procedures in situ. For example, our study used Radio Frequency (RF) based equipment to monitor and record the comfort level of the human subjects throughout the adult trial. We found that there was interference coming from sources that were only apparent when all the trial equipment was operating simultaneously.
7. Some points to consider when designing questionnaires are made. Completing questionnaires away from the trial area may conserve resources but influence the questionnaire results.
8. A careful consideration of methodological and design issues regarding the preparation of any user study will fundamentally impact any results and conclusions that might be gained.

It is vital that sufficient time is allowed for piloting and testing any planned trials properly in order to identify deficiencies and make improvements before the trials start properly. Full scale pilot studies will expose problems that are not apparent when running individual tests on the experimental equipment and methods. In our own studies the problems we did encounter were not serious enough to damage or invalidate major parts of the trials. We have highlighted other features of our trials we can improve upon, and made suggestions as to how to overcome the problems we have encountered. The lessons learned can be used to improve future trials involving human-robot interaction.

Acknowledgements

The work described in this paper was conducted within the EU Integrated Project COGNIRON ("The Cognitive Robot Companion") and was funded by the European Commission Division FP6-IST Future and Emerging Technologies under Contract FP6-002020. Many thanks to Christina Kaouri, René te Boekhorst, Chrystopher Nehaniv, Iain Werry, David

Lee, Bob Guscott and Markus Finke for their help in designing, implementing and carrying out the trials.

References

ActivMedia Robotics (2005). ActivMedia Robotics Interface for Applications (ARIA). http://robots.activmedia.com/.

BBC Science news (2004). Interactive games run by the COGNIRON team from the University of Hertfordshire at the Science Museum, London.http://news.bbc.co.uk/2/hi/technology/396 269stm 2004.

Brooks, R.A. (1991). Intelligence without representation. *Artificial Intelligence* Vol 47, pp. 139–159.

COGNIRON (2005). Website available at www.cogniron.org. 2005

Crown copyright, (2003). An introduction to health and safety. HSE Books .http://www.hse.gov.uk/pubns/indg259.pdf. ISBN 0 7176 2685 7. 2003

Crown copyright, (1998). Data Protection Act 1998. http://www.hmso.gov.uk/acts/acts1998/ 19980029.htm#aofs. ISBN 0 10 542998 8. 1998.

P. Dario, E. Guglielmelli, C. Laschi (2001). Humanoids and Personal Robots: Design and Experiments. *Journal of Robotic Systems* 18 (12), pp. 673–690.

T. Fong,, I. Nourbakhsh, K. Dautenhahn (2003). A survey of socially interactive robots. *Robotics and Autonomous Systems*, Vol. 42, pp. 143-166.

A. Green, H. Hűttenrauch, K Severinson Eklundh (2004). Applying the Wizard of Oz Framework to Cooperative Service discovery and Configuration. Proc. *IEEE Ro-man 2004,* 13th IEEE International Workshop on Robot and Human Interactive, Oka-yama, Japan, IEEE Press.

P. Hinds, T. Roberts, H. Jones (2004). Whose Job Is It Anyway? A Study of Human-Robot Interaction in a Collaborative Task.. *Human-Computer Interaction*, Vol. 19, pp.151-18.

T. Kanda, T Hirano, D Eaton (2004). Interactive Robots as Social Partners and Peer Tutors for Children: A Field Trial. *Human-computer Interaction*, Vol. 19, pp. 61-84.

D. Maulsby, S. Greenberg, R Mander (1993). Prototyping an intelligent agent through Wizard of Oz. *Proc.ACM SIGCHI Conference on Human Factors in Computing Systems, Amsterdam*, The Netherlands, ACM Press, pp. 277-284.

B. Robins, K. Dautenhahn, J. Dubowski (2004). Investigating Autistic Children's Attitudes Towards Strangers with the Theatrical Robot - A New Experimental Paradigm in Human-Robot Interaction Studies . *Proc. IEEE Ro-man 2004,* 13th IEEE International Workshop on Robot and Human Interactive. Oka-yama Japan, IEEE Press, pp. 557-562

N. Roy, S. Thrun. Motion planning through policy search (2002). *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems*, EPFL, Switzerland, pp. 2419–2425.

UPR AS/A/2 (2004). University of Hertfordshire Ethics Policy. http://wisdom.herts.ac.uk/research/Indexes2/usefullinks2.htm. 2004.

VICTEC (2003). VICTEC Project website available at http://www.victec.org, 2003.

Ontological and Anthropological Dimensions of Social Robotics

Jutta Weber [*]
[*]Institute for Philosophy of Science
University of Vienna
jutta.weber@univie.ac.at

Abstract

From a philosophical viewpoint ontological and anthropological dimensions of concepts of sociality and social intelligence in robotics are discussed. Diverse ontological options of social interaction as static or dynamic are analysed with regard to different theoretical approaches in sociology and the socio-behavioural sciences.

1 Introduction

Recent research on social robots is focussing on the creation of interactive systems that are able to recognise others, interpret gestures and verbal expressions, which recognize and express emotions and that are capable of social learning. A central question concerning social robotics is how "building such technologies shapes our self-understanding, and how these technologies impact society" (Breazeal 2002).

To understand the implications of these developments it is important to analyse central concepts of social robotics like the social, sociality, human nature and human-style interactions. Main questions are: What concepts of sociality are translated into action by social robotics? How is social behaviour conceptualised, shaped, or instantiated in software implementation processes? And what kind of social behaviours do we want to shape and implement into artefacts?

2 Some Clarification: 'Ontology' and 'Anthropology'

In the following I will use the term ontology in a philosophically but not in the sense of a branch of metaphysics which defines the nature of existence or the categorical structure of reality. The term ‘ontology’ here refers to the meta-theoretical core of a theory. Contemporary philosophy of science agrees that there is no theory without meta-theoretical principles or orienting strategies. These principles or strategies contain syntactical structures as well as ontological options. Ontological options lay down what set of things, entities, events or systems are regarded as existing (Lowe 1995). Central semantics are also regarded as part of the ontological options of a theory (Ritsert 2003). Following this understanding of ontology, 'anthropology' can be regarded as part of the ontological options of a theory and not as an essentialist and pregiven definition of human nature. Anthropology is defined in the sense of a set of human properties and behaviour which is taken for granted in the frame of a theory.

3 Sociality, Social Intelligence, and Social Relations

The growing interest in the social factor in robotics is related to the idea of a biologically-grounded, evolutionary origin of intelligence. The Social Intelligence Hypothesis - also called Machiavellian intelligence hypothesis - states that primate intelligence evolved to handle social problems (Jolly 1966; for discussion see Kummer et al. 1997). Social behaviour is said, not only to be grounded in the reflection of mental states and their usage in social interaction, but as necessary to predict the behaviour of others and change one´s own behaviour in relation to these predictions.

Kerstin Dautenhahn and Thomas Christaller describe the function of social interaction in the sense of double contingency as that which "enables one to establish and effectively handle highly complex social relationships and, at the same time, this kind of 'inner eye' […] allows a cognitive feedback, which is necessary for all sorts of abstract problem solving" (Dautenhahn and Christaller 1997) According to this argument intelligent behaviour has a social off-spring and an embodied basis (ibid.; see also Duffy 2003) and helps humans - and it shall help robots - to survive in a complex and unpredictable world (Breazeal 2003).

This definition of social interaction developed in the sense of reflection of one´s own and anticipation of the behaviour of others, which was developed mainly in behavioural sciences like primatology, ethology and psychology, is quite similar to that of 'double contingency' in sociological approaches of system theory (Luhmann 1984; Parsons 1968) or interactionism. (Mead 1938, for critical discussion see Lindemann 2002).

The socio-behaviourist and these sociological concept of sociality share a quite formal understanding of the social, while other theories like critical theory, ethnomethodology, or Marxism developed a more contextual and material understanding of the social. As there is no generally acknowledged understanding of the social in social theory the decision for a more formal concept of the social can be regarded as part of the ontological option of a theory.

4 Dynamic Social Knowledge and Social Mechanisms

The sociological theorem of double contingency in system theory (Parsons, Luhmann) or interactionism is (implicitly) build on an anthropology that understands the relation of humans and their environment as open and flexible (Lindemann 2002), as a product of culture and it is grounded in a constructivist epistemology (Weber 1999).

The argument of the Machiavellian intelligence hypothesis is based on an anthropology that understands human nature as the product of a biological and contingent process: evolution. The epistemological frame stands in the tradition of naturalism (Danto 1967).

Both approaches share a formal understanding of social interaction which leaves plenty of room for different or maybe diverse kinds of interpretation of the 'nature' of social interaction.

In some approaches of social robotics human nature is regarded as flexible and open as it is embedded in time and space. For example, Dautenhahn and Christaller (1997) "do not regard 'social expertise' as a set of special social rules (propositions), which are stored, retrieved and applied to the social world. Instead, social knowledge is dynamically reconstructed while remembering past events and adapting knowledge about old situations to new ones (and vice versa). (…) we hypothesize that social intelligence might also be a general principle in the evolution of artificial intelligence, not necessarily restricted to a biological substrate." (Dautenhahn / Christaller 1997)

Here we find an anthropological option of an open and flexible human nature and the understanding of social knowledge as a very complex and dynamic product embedded into a historical frame, which is regarded as the product of evolution but can emerge (because of its dynamic nature?) also under different conditions.

While this interpretation of social knowledge stresses the dynamic and flexible process of social interaction, we also find more static and behaviourist interpretations of social behaviours - especially in the discussion on emotional intelligence which interpret social action more in terms of social mechanisms.

5 Emotional Intelligence

Social interaction in the sense of double contingency affords the understanding of the emotions of the alter ego (Duffy 2004). Emotional intelligence is understood as an important part of social intelligence (Canamero 1997) and is defined by Daniel Goleman (1997) as "the ability to monitor one´s own and others´ emotions, to discriminate among them, and to use the information to guide one´s thinking and actions".

In discussions on emotional intelligence - mostly with regard to psychology and ethology - social interaction is interpreted in terms of pregiven social mechanisms, like for example a few (fixed) basic emotions (see Breazeal 2003), 'moral sentiments' or social norms (Petta / Staller 2001). The latter are said to fulfil very particular functions to improve the adaptability of the individual towards the demands of his or her social life (Ekman 1992).

The understanding of sociality is reframed and made operational (for computational modelling) by defining the function of emotion in social interaction in terms of costs and benefits of the individual: "? there must be a material gain from having these emotions, otherwise they would not have evolved. (?) emotional predispositions have long-term material advantages: An honest partner with the predisposition to feel guilt will be sought as a partner in future interactions. The predisposition to get outraged will deter others from cheating." (Staller / Petta 2001) This interpretation of emotional predispositions is due to a less dynamic and more functional understanding of social interaction.

6 Sociality and Individualism

While most approaches in social robotics agree that social intelligence was developed out of the necessity to survive in a dynamic, unpredictable environment, some stress the dynamics of social knowledge, while others draw on the importance of fixed sets of rules and social norms for social interaction. These diverse interpretations are made possible by the formal character of the interpretation of social interaction in the sense of 'double contingeny', of the ability to predict the behaviour of others and change one´s own behaviour in relation to these predictions.
On the one side we find more functional approaches which understand society as the accumulation of individuals and social interaction as the negotiation of personal values: "Most behavioural and social sciences assume human sociality is a by-product of individualism. Briefly put, individuals are fundamentally self-interested; 'social' refers to the exchange of costs and benefits in the pursuit of outcomes of purely personal value, and "society" is the aggregate of individuals in pursuit of their respective self-interests." (Carporeal 1995)
Sociological approaches in system theory (Luhmann) or interactionism (Mead) more often defines sociality as something that is realized in the behaviour of the alter ego and as the outcome of a contingent and historical process of interpretation. According to this society is understood as a relation of socialized individuals that is regulated through culture and societal institutions (Lindemann 2002).
While many socio-behaviourist approaches take for granted that social behaviour is a general achievement of primates (and it is only abstract problem solving, which is a human-only property), system theory and interactionism regard humans as the only social actors (Lindemann 2002).
Only in recent time there are new approaches - especially in the field of science and technology studies - that make a claim for a "symmetrical anthropology" (Latour 1993; see also Haraway 1989) in which humans, animals as well as machines are regarded as social actors. (for discussion see Albertsen and Diken 2003)

7 Socio-Behaviourist Sciences and the Computational Modelling of Social Intelligence

There are historical reasons for the dominance of socio-behaviourist approaches (mostly in the anglo-american tradition) in artificial intelligence (see Chrisley / Ziemke 2002), but there might be also pragmatic ones.
One reason is the dominance of psychology, ethology and primatology which fits especially to approaches of Artificial Life and biologically-inspired robotics, while Luhmannïs system theory or Mead´s interactionism is oriented primarily towards sociology. The socio-behaviourist tradition regards not only humans, but also organisms in general as capable of social intelligence which is much more attractive for social robotics that wants to model social interaction in artificial systems.
While both 'traditions' share a more formal understanding of social interaction that enables naturalist, biological ontological groundings as well as constructivist, cultural ones with an dynamic understanding of the social, we nevertheless find many socio-behaviourist conceptions which offer a quite functional and less dynamic understanding of social interaction that makes the implementation of concrete social behaviours into artefacts much easier.
Social interaction is understood in these approaches in the sense of social mechanisms and norms thereby using quite static models of social behaviours: For example, "(s)tereotypical communication cues provide obvious mechanisms for communication between robots and people." (Duffy 2003, 188) Other relevant standardizations used in social robotics are stereotypical models of 'basic' emotions, distinct personality traits (see also Fong et al. 155), gender and class stereotypes (Moldt / von Scheve 2002) etc. These norms, stereotypes and standardizations make social intelligence (easier) operational for the computational modelling of social intelligence (Salovey and Meyer 1990).

8 On the Compatability of Ontological Options

On the one hand the formal description of social interaction as 'double contingency', as the prediction of the behaviour of others and adaption of one´s own behaviour leaves plenty of room for dynamic as well as static understandings of social interaction with divergent epistemological framings. On the other hand it is an open question how an embodied and situated understanding of social intelligence which regards organisms in general as social actors, can be used coherently with functional psychological concepts of emotion, personality and social mechanisms. If social intelligence is regarded as the outcome of situated, embodied social interaction one would expect to regard robots as an own kind (Duffy 2004) developing their own way of sociality. This would leave it open whether artificial systems will be able to develop the potential for abstract problem solving. Therefore imitating the social interaction of humans might neither be helpful for the development of human-robot interaction and probably also not very desirable (Billard 2004).
In any case, the analysis of ontological options of concepts of sociality might be helpful to think of the compatibility of diverse approaches and design

methods and the outcome of their combination. As there is no agreement on a concept of 'the' social neither in sociology or psychology (similar to the discussion on the concept of life in Artificial Life) - it would be interesting to take more sociological approaches in general into account, which were mostly neglected up to now. It could be helpful to compare not only the different effects of the implementation of dynamic and static concepts of sociality but also of formal and contextual ones.

References

Niels Albertsen and Bülent Diken, What is the Social? (draft), published by the Department of Sociology, Lancaster University, retrieved: June 7th, 2003, from http://www.comp.lancs.ac.uk/sociology/soc033bd.html (last access 7.6.2003).

Aude G. Billard, Imitation, Language and other Natural Means of Interactions we build for our Machines: Do we really want machines to resemble us that much? Position Paper for the Workshop *Dimesions of Sociality. Shaping Relationsships with Machines*, Vienna, 19-20th November, 2004

Cynthia Breazeal, Designing Sociable Robots. Cambridge, MA: MIT Press 2002

Cynthia Breazeal, Emotion and Sociable Humanoid Robots. In: *International Journal of Human-Computer Studies*, Volume 59, Issue 1-2, 119-155, 2003

Lola Canamero, Modeling Motivations and Emotions as a basis for intelligent behaviour, in: Lewis W. Johnson (ed.): Proceedings of the International Conference on Autonomous Agents. Agents '97. New York: ACM Press, 148-155, 1997.

Linnda R. Caporael, Sociality: Coordinating Bodies, Minds and Groups, *Psycoloquy 6(01), Group-selection 1, 1995;* retrieved: September 30, 2004, from http://www.psycprints.ecs.soton.ac.uk/archive/00000448

Ron Chrisley and Tom Ziemke, Embodiment. Encyclopedia of Cognitive Science, London: Macmillan Publishers, 1102-1108, 2002.

Arthur C. Danto, Naturalism. In: Paul Edwards (ed.): The Encyclopedia of Philosophy. New York / London: Routledge 1967, 446-447.

Kerstin Dautenhahn and Thomas Christaller, Remembering, rehearsal and empathy - towards a social and embodied cognitive psychology for artifacts. In: Two Sciences of Mind. Readings in cognitive science and consciousness. John Benjamins, pp. 257-282.

Brian R. Duffy, Anthropomorphism and the Social Robot. In: *Robotics and Autonomous Systems*, 42, 177-190, 2003

Brian R. Duffy, The Social Robot Paradox. Position Paper for the Workshop *Dimensions of Sociality. Shaping Relationsships with Machines*, Vienna, 19-20th November 2004.

Paul Ekman, 'Are there basic emotions?' *Psychological Review* 99 (3), 550-553, 1992.

Terrence Fong and Illah Nourbakhsh and Kerstin Dautenhahn, A survey of socially interactive robots. In: *Robotics and Autonomous Systems* 42, 143-166, 2003.

Donna Haraway, Primate Visions. Gender, Race, and Nature in the World of Modern Science. New York / London: Routledge 1989.

Alison Jolly, Lemur social behaviour and primate intelligence. *Science*, 153:501-506, 1996.

Hans Kummer and Lorraine Daston and Gerd Gigerenzer and Joan B. Silk, The social intelligence hypothesis. In: Peter Weingart and Sandra D. Mitchell and Peter J. Richerson and Sabine Maasen (eds.): Human by Nature: between biology and social sciences. Hillsdale, NJ: Lawrence Erlbaum 1997, 157-179 .

Bruno Latour, We Have Never Been Modern. Hertfordshire: Harvester Wheatsheaf 1993.

Gesa Lindemann, Die Grenzen des Sozialen. Zur sozio-technischen Konstruktion von Leben und Tod in der Intensivmedizin. München, Wilhelm Fink 2002.

Jonathan E. Lowe, E.: Ontology. In: Ted Honderich (ed.): The Oxford Companion to Philosophy. Oxford / New York: Oxford University Press 1995, 634-635.

Niklas Luhmann, Soziale Systeme. Grundriss einer allgemeinen Theorie. Frankfurt a.M.: Suhrkamp 1984.

George H. Mead, The Philosophy of the Act. Chicago, London: University of Chicago Press 1938

Daniel Moldt and Christian von Scheve, Attribution and Adaptation: The Case of Social Norms and Emotion in Human-Agent Interaction? in:

Marsh et al. (eds.), Proceedings of The Philosophy and Design of Socially Adept Technologies, workshop held in conjunction with CHI'02, 20.4.02, Minneapolis/Minnesota, USA, 39-41, 2002

Paolo Petta and Alexander Staller, Introducing Emotions into the Computational Study of Social Norms: A First Evaluation. In: *Journal of Artificial Societies and Social Simulation,* vol. 4, no. 1., 2001

Juergen Ritsert: Einfuehrung in die Logik der Sozialwissenschaften. Muenster: Westfaelisches Dampfboot 2003

Peter Salovey, and John D. Mayer, Emotional intelligence. *Imagination, Cognition, and Personality*, 9, 185-211, 1990.

Jutta Weber, Contested Meanings: Nature in the Age of Technoscience. In: Juergen Mittelstrass (ed.): Die Zukunft des Wissens. XVIII. Deutscher Kongress fuer Philosophie. Konstanz, UVK Universitaets-Verlag Konstanz 1999, 466-473

Child and Adults' Perspectives on Robot Appearance

Sarah Woods, Kerstin Dautenhahn
University of Hertfordshire, School of Computer Science, Adaptive Systems Research Group
Hatfield, AL10 9AB, UK
s.n.woods, k.dautenhahn@herts.ac.uk

Joerg Schulz
University of Hertfordshire
Psychology Department
Hatfield, AL10 9AB, UK
j.schulz@herts.ac.uk

Abstract

This study explored children's and adults' attitudes towards different types of robots. A large sample of children viewed different robot images and completed a questionnaire that enquired about different robot physical attributes, personality and emotion characteristics. A few adults independently rated the overall appearance of different robot images. Results indicated high levels of agreement for classifications of robot appearance between children and adults, but children only differentiated between certain robot personality characteristics (e.g. aggressiveness) and emotions (e.g. anger) in relation to how adults rated the robots' appearances. Agreement among children for particular robots in terms of personality and emotion attributes varied. Previously, we found evidence for the Uncanny Valley based on children's ratings of robot appearance. However, based on the adults' rating of robot appearances, we did not find evidence of the Uncanny Valley in terms of how children perceived emotions and personality of the robots. Results are discussed in light of future design implications for children's robots.

1 Introduction

Robots are being used within increasingly diverse areas and many research projects study robots that can directly interact with humans [1]. Robot-human interaction encapsulates a wide spectrum of factors that need consideration including perception, cognitive and social capabilities of the robot and the matching of the robot interaction with the target group [2].

Because of the potential benefit of using robots that are able to interact with humans, research is beginning to consider robot-human interaction outside of the laboratory. Service robots are used within a variety of settings such as to deliver hospital meals, operate factory machinery, and clean factory floors, which does involve some shared human environments. However, the amount of human-robot interaction with these service robots is still minimal requiring little social behaviour [3]. Robots said to be able to engage in more extensive social interaction with humans among others include AIBO [4], Kismet [5], Feelix [6], and Pearl [7]. It is suggested that social robots should be able to exhibit some "human social" characteristics such as emotions, recognition of other agents and exhibiting personality characteristics [8] both in terms of physical appearance and behavioural competencies.

An important consideration for the designers of robots involves the target population whether it is children, adolescents, adults, or elderly as the attitudes and opinions of these groups towards robot interactions are likely to be quite different. For example Scopelliti et al. [9] revealed differences between young and elderly populations towards the idea of having a robot in the home with young people scoring highly positive and older people expressing more negativity and anxiety towards the idea of having a robot assistant in the home.

Related to this is the issue of matching the appearance and behaviour of the robot to the desired population. Goetz, Kiesler & Powers [10] [11] revealed that people expect a robot to look and act appropriately for different tasks. For example, a robot that performs in a playful manner is preferred for a fun carefree game, but a more "serious" robot is preferred for a serious health related exercise regime.

Kanda & Ishiguro [12] offer a novel approach and aim at developing a social robot for children where the robot (Robovie) can read human relationships from children's physical behaviour. This example highlights the importance of Robovie being designed appropriately for young children. It seems that if a robot cannot comply with the user's expectations, they will be disappointed and unengaged with the robot. For example, if a robot closely resembles a human in appearance but then does not behave like one, expectations are being violated and there is the danger of the human-robot interaction breaking down. It could even lead to feelings of revulsion against the robot as in the 'Uncanny Valley' proposed by Mori [13].

One domain of robotics research which remains scarcely explored but is beginning to emerge is the involvement of psychologists and extensive use of methods and techniques commonly used in psychology research in assisting the design and evaluation of robots for different target groups [14-18]. Using a psychology approach allows the exploration of different evaluation techniques [19], to enquire about robot perceptions and issues such as anxiety towards robots [18], and to study the ascription of moral development towards robots [15]. It is our position that the input of psychologists could assist in the design of socially interactive robots by examining what social skills are desirable for robots, what the most suitable appearance is for robots in different roles and for different target groups, and assisting in the design of robots with personality, empathy and cognition.

This paper is part of a research project which is exploring children's perceptions and attitudes towards different robotic designs paying special consideration to both the physical properties and social, behavioural aspects of robots. Previous work related to this project has reported that children are able to clearly distinguish between emotions and personality when judging different types of robots [19]. For example, children judged humanlike robots as aggressive but human-machine like robots as friendly. The work proposed possible design implications for children's robots such as considering a combination of robotic features (e.g. facial features, body shape, gender, forms of movement) rather than focusing on certain features in isolation (e.g. just the face).

The current paper considers the findings of Woods, Dautenhahn & Schulz [19] from a different perspective and examines the implications more deeply. We investigate whether adults and children agree on ratings of overall robot appearance and what children's perceptions of robot personality and emotions are in relation to adult views of overall robot appearance. Furthermore, we examine whether children among themselves agree about their perceptions of robot personality and emotion attributes.

The specific research questions that we were interested in were:

1. Do children differentiate robots in terms of personality and emotions based on adult ratings of robot appearance?
2. How do children perceive robot comprehension/understanding in relation to adult ratings of robot appearance?
3. To what extent do adults and children agree when classifying robot appearance?
4. To what extent do children among themselves agree on their ratings of robot appearance, personality and feelings?

The remainder of the paper is structured as follows. After introducing the experimental method, we address each research question in separate sections of the results section. The last section summarizes and concludes the paper.

2 Method

2.1 Design & Participants

159 children (male: N: 82 (52%) and girls: N: 77 (48%) aged 9-11 (years 5 & 6) participated in the study (M age = 10.19 years, SD: 0.55) which used a questionnaire based design and quantitative statistical techniques. Children viewed 5 robot images, completing the robotics questionnaire for each image. Five adults from the Adaptive Systems Research Group at the University of Hertfordshire also participated in this study in terms of devising the coding scheme for the robots and providing ratings of robot appearance.

2.2 Instruments

Robot Pictures

A coding schedule was developed to categorise 40 robot images according to the following criteria: a) movement, b) shape, c) overall appearance (e.g. human, machine, animal, human-machine, animal-machine, animal-human), d) facial features, e) gender, f) functionality (e.g. toy, friend, machine). Based upon the age and cognitive abilities of the children who took part in the study, 8 groups containing 5 robot images were formed containing different robot classifications derived from the coding schedule, (total N: 40 robot images).

Robot Pictures Questionnaire: 'What do you think?'

A questionnaire was designed to enquire about children's perceptions of different robot attributes. Section one referred to questions about robot appearance (e.g. what does this robot use to move around? What shape is the robot's body?). Section two asked questions about robot personality, rated according to a 5-point Likert scale and included questions about friendliness, aggressiveness, whether the robot appeared shy, and whether the robot appeared bossy. An example question was: Do you think this robot is (or could be) aggressive? The content and structure of the questionnaire was checked with a head teacher at one of the schools to ensure that it was age appropriate.

2.3 Procedure

The Robot Pictures Questionnaire was completed by groups of between 4-8 children from a number of primary schools. Children were seated in such a way that they would be able to answer the questionnaires confidentially without distraction from other children. A set of 5 robot images were distributed to each child. Each child completed 5 copies of the Robot Pictures Questionnaire for each of the images. In the lab, 5 adults independently rated the overall appearance (e.g. human, machine, animal, human-machine, animal-machine, animal-human) for the 40 robot images.[1]

3 Results

3.1 Children's perceptions of robot personality and emotion in relation to adult ratings of robot appearance

One-way analysis of variance was carried out to examine whether there were any significant differences between children's perceptions of robot personality attributes and emotions in relation to adult ratings of robot appearance. Significant differences were revealed for robot friendliness and overall appearance ($F = 5.84$, (5, 795), $p < .001$), robot aggressiveness and overall appearance ($F = 4.40$, (5, 795), $p < .001$) and robot anger and overall appearance ($F = 3.27$, (5, 795), $p = .006$). Post-hoc analyses revealed that human-machine looking robots were rated by children as being significantly more friendly than pure machine looking robots (human-machine $\underline{X} = 3.66$, machine $\underline{X} = 3.13$) and human-animal looking robots (human-animal $\underline{X} = 2.60$). For robot aggressiveness and overall appearance, post-hoc analyses revealed significant differences between pure machine looking robots and human-machine looking robots (machine $\underline{X} = 2.88$, human-machine $\underline{X} = 2.36$). Pure machine looking robots were rated by children as being the most aggressive according to adult ratings of robot appearance. Post-hoc comparisons highlighted significant differences between pure machine looking robots, human-machine looking robots and robot angriness (machine $\underline{X} = 2.88$, human-machine $\underline{X} = 2.42$) with machine looking robots being rated by children as significantly more angry compared to human-machine looking robots. No significant differences were found between children's ratings of robot shyness, bossiness, happiness, sadness and fright with respect to adult ratings of overall robot appearance (See Figure 1 for mean values of children's perceptions of robot personality attributes and Figure 2 for mean values of children's perceptions of robot emotions in relation to adult ratings).

These results indicate that children's views of robot personality and emotions were quite distinguishable according to adult ratings of different robot appearances. For example, children perceived human-machine robots rated by adults as being friendlier and less angry than pure machinelike robots.

Children's views of robot personality and emotions in relation to different robot appearances were quite different compared to adult ratings [18]. Results of children's views of overall robot appearance in relation to robot personality and emotions provided support for the Uncanny Valley with pure humanlike robots being rated by children as the most aggressive, and a mix of human-machine robots as the most friendly. In contrast, when relating children's perceptions of the robots with adult ratings of robot appearance we do not find any evidence for the Uncanny Valley. Children perceived humanlike robots ('human-like' in terms of how adults rated their appearance) as being the most friendly and least aggressive. This could suggest that children use different criteria of the robots external features in rating human-like and human-machine like robots compared to adults.

3.2 Children's perceptions of robot comprehension and adult ratings of robot appearance

Chi-square analysis in the form of cross tabulations revealed a significant association between children's views of a robot being able to understand them and adult ratings of robot appearance ($X = 122.45$, df = 5 (795), $p = 0.000$). Children stated that human-like robots were most likely to understand them (87%), followed by human-machine looking robots (76%). Only 32% of children felt that a machine-like robot would understand them if they tried to talk to it (See Figure 3.) This result indicates that children and adults may have similar perceptions

[1] We are aware that this is an unrepresentative sample for an adult population, but it seemed suitable for this preliminary study, since we wanted to link children's perceptions to potential robot designer views, and members of the research group have been involved in robot design (though not in a commercial context).

about what types of robot appearances are linked to robots being able to communicate[2].

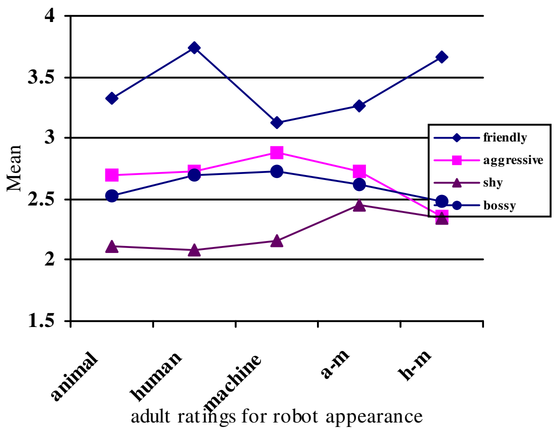


Figure 1: Adult ratings of overall robot appearance and children's ratings of robot characteristics[3]

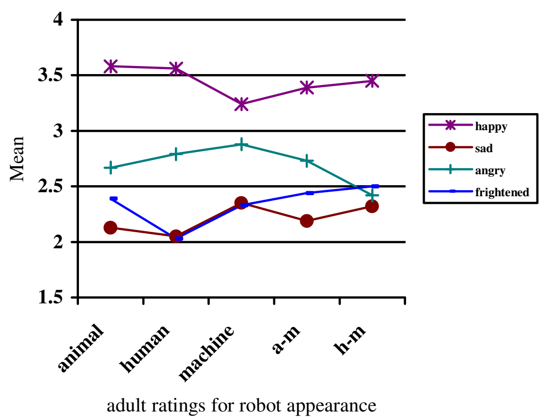


Figure 2: Adult ratings of robot appearance and children's perceptions of robot emotions

[2] Note, the term 'understanding' is a very generic and ambiguous concept, ratings might not necessarily be linked to communication skills. More research is needed to refine this work.

[3] Significance at P < .001 level for friendliness and aggressiveness in Figure 1. Shy and bossy were non-significant. (A-M = animal-machine, H-M = human-machine).

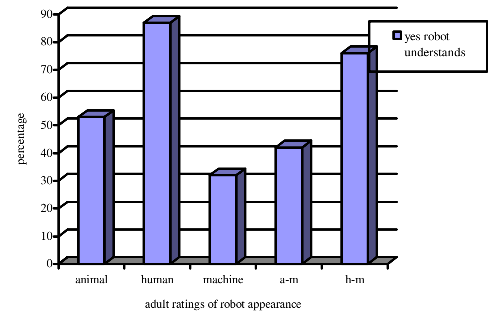


Figure 3: The association between children's perceptions of a robot understanding them and adult ratings of robot appearance.

3.3 Agreement between adult and child classifications of robot appearance

As adult and children's views are likely to be from very different cognitive and social perspectives we were interested in examining the degree of agreement between adults' and children's ratings of robot appearance. Table 1 illustrates the percentage levels of agreement between children and adults with corresponding Kappa Coefficients. The lowest percentage agreement between children and adults was for pure human-like robots, and the highest level of agreement was for machine-like and human-machine looking robots, although all Kappa coefficients were highly significant indicating high levels of agreement between children and adults for robot appearance.

Robot Appearance	% level of agreement between children and adults	Kappa Coefficient
Animal	66.7	0.77 ($p < .001$)
Machine	77.8	0.79 ($p < .001$)
Human	40.0	0.54 ($p < .001$)
Animal-machine	75.0	0.48 ($p < .001$)
Human-machine	77.8	0.66 ($p <.001$)

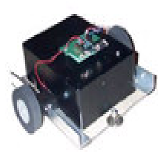
Table 1: Percentage level of agreement between children and adults for the appearance of robots.

3.4 Children's agreement for robot personality and emotions

Kendall's W coefficient of concordance statistic was carried out for each of the 40 robots to examine the level of agreement between children's opinions of robot personality (friendliness, aggressiveness, bossiness, shyness) and robot emotions (happiness, sadness, anger and fright). Overall agreement across all robots was quite low although significant. The levels of concordance between children for particular robots varied considerably. The lowest level of agreement between children was for robot id number 28, and the highest level of agreement was for robot id number 59. Other robots that high levels of agreement were robots with id number 90 and 100 (see images below). See table 2 for Kendall's W coefficient statistic.

Robot Id	Low/high concor-dance	Kendall's W (KW) for robot per-sonality	Kendall's W (KW) for robot emotions
Overall (40 robot images)	Moderate	KW = 0.10 ($p < 0.001$)	KW = 0.12 ($p < .001$)
Robot 28	Low	KW = 0.02 ($p = 0.67$)	KW = 0.03 ($p = .60$)
Robot 59	High	KW = 0.45 ($p < 0.001$)	KW = 0.43 ($p < .001$)
Robot 100	Moderate	KW = 0.48 ($p < 0.001$)	KW = 0.21 ($p = .01$)
Robot 90	High	KW = 0.52 ($p < 0.001$)	KW = 0.53 ($p < .001$)

Table 2: Children's agreement for robot personality & emotions for different robots (Kendall's W values and corresponding significance levels).



Robot image no. 28: lowest levels of child agreement

Robot image no. 59: highest levels of child agreement

Robot image no. 100: high level of child agreement

Robot image no. 90: high level of child agreement

A one-way analysis of variance was carried out between children's ratings of overall robot appearance and the Kendall's W concordance rating. No significant differences were uncovered ($F = 0.49$, (4, 40), $p = 0.75$) indicating no differences between children's levels of agreement on robot personality and emotions and the different types of robot appearance. Thus, agreement across children was no better for human-like robots than for machine-like robots. Results of a Mann Whitney U test did not reveal any significant median differences but it was clear from an error plot that those robots rated as being human-like in appearance had less spread in variance even though it was still quite low on consistency. Other robot appearances had much more variance.

4 Discussion and Conclusions

Firstly, the current study explored the levels of agreement between adults and children for overall robot appearance. Secondly, we considered children's ability to differentiate between different robot personality and emotions according to adult ratings of robot appearance, and finally the level of agreement among children for robot personality and emotions was examined.

A summary of the main results indicated that:

- Children only differentiated between certain robot behaviours (aggressiveness and friendliness) and emotions (anger) in relation to the overall robot appearance ratings by adults.
- Overall, children and adults demonstrated high levels of agreement for classifications of robot appearance, particularly for machine-like and human-machine robots.
- In contrast to previous findings suggesting evidence for the Uncanny Valley based on children's ratings of robot appearance, this finding was not replicated with respect to adult ratings of robot appearance.
- An exploration of children's levels of agreement for particular robots and robot personality and emotion revealed varying degrees of agreement.
- No differences were found between children's levels of agreement on robot personality and emotions and the different types of robot appearance (i.e. agreement across children was no better for human-like robots, as rated by children, than machine-like robots).

The finding that agreement between children and adults for classifications of overall robot appearance was generally high is a positive result for the future design of robots. As adults and children have different social and cognitive views of the world [20] we expected less agreement.

The levels of agreement demonstrated by children for robot understandability in relation to adult ratings of overall robot appearance was a positive finding as it suggests that robots can be designed in such a way that children are able to differentiate this dimension[4]. However, closer exploration is required to determine what the exact features are that allow children to distinguish in their minds whether a robot is able to understand them or not (e.g. is it the fact that the robot has a mouth or general human form?)

Results from this study point to the notion that children were only able to differentiate between certain robot personalities and emotions in relation to adult ratings of robot appearance. This could be attributed to a number of reasons including the fact that children aged 9-10 have a limited understanding of applying emotions and certain personalities to robots. The results showed that children differentiated between robot friendliness, aggressiveness and anger and robot appearance but not for more subtle personalities and emotions such as bossiness, shyness, fright. However, it was somewhat surprising that children did not distinguish between sadness and happiness as children should have a clear understanding of these two basic emotions at aged 9-10. Another explanation could be that robots are not yet able to convey subtle emotions and personalities such as shyness and fright, therefore making it hard for the user to recognise the possibility of such personalities. This question is worth further exploration for designers as it would certainly be a desirable feature for robots to be able to perform and exhibit subtle personalities and emotions. In future, this needs to be explored with some live interactions between children and robots and cannot be fully answered by the current study as only static images of robots were used which is a limitation of the present study.

It is important to consider children's overall agreement towards the personality and emotions of different robots as designers' intentions are usually to convey a particular type of personality in line with a particular robot. For example, AIBO has been designed to be a toy or a pet and designers wanted to convey it as being a friendly non-aggressive robot. From a designer point of view, one would hope that there would be high agreement between children for this robot being friendly, happy and non-aggressive. It would be disappointing if some children viewed it as a sad, aggressive and angry robot. The findings of the current study revealed that for particular robots there was high agreement among children towards the robots personality and emotions but for others agreement was extremely low. It was somewhat surprising that agreement across children for robot personalities and emotions were not affected by the overall appearance of the robot as rated by the children. One might have expected for example that children would demonstrate higher agreement for human-like robots compared to a mixture of human-machine like robots. This is worthy of further study as designers in the future could well want to design different robot appearances that have definite personalities and emotion patterns. To assist with the future design of robots, designers should perhaps compare the appearance of those robots that lead to highly consistent views with those that were inconsistent. This result highlights the importance of adults to include children in the design phase of robots, that are meant for a child target audience, from the outset of the planning stage to ensure that children's views are accurately captured [21].

The previous finding that children's perceptions of robot personality and emotions, according to their ratings of human-like appearance, fell into the Uncanny Valley could not be confirmed in relation to adult ratings of robot appearance. This is an interesting finding and emphasises the importance of considering children's views of particular robot appearances in addition to adults. The Uncanny Valley theory proposed by Mori [13] posited that as a robot increases in humanness, there is a point where the robot is not 100% similar to a human and the balance becomes uncomfortable or even repulsive. Children clearly felt uncomfortable with their views of pure humanlike appearances (according to how they judged 'human-like'), but did not experience this discomfort based on adult ratings of humanlike robot appearance. Note, while a large sample of children was used in the present study, only few adults participated in the study. A larger adult sample size would clearly be desirable for future work.

Overall, the study emphasises the importance of designers considering the input of children's ideas and views about robots before, during and after the design and construction of new robots specifically pitched for children. In order to overcome the limitations of the current study, future studies should consider children's attitudes using live child-robot interactions and should pay closer attention to the finer details of robot appearance that are necessary to communicate different personalities and emotions. Future studies could also consider comparing adult and children's views of robot personality constructs and emotions and how these relate to the appearance of robots. Finally, while in the present study adults with a robotics related background were considered 'potential robot designers', future studies involving professional robot designers are necessary

[4] Note, not all the robots included in our study were specifically designed for a child target audience.

in order to investigate in more depth the relationship between children's views of robots designed by adults.

Acknowledgements

We would like to thank the following schools for participating in the above study: Commonswood School, Welwyn Garden City, Hertfordshire, UK, Applecroft school, Welwyn Garden City, Hertfordshire, UK, Cunningham School, St Albans, Hertfordshire, UK, and High Beeches School, St Albans, Hertfordshire, UK. .

References

[1] C. Bartneck, "Rapid prototyping for interactive robots," Proc. IAS-8, Amsterdam, 2004.

[2] C. Breazeal, "Social interaction in HRI: The robot view," *IEEE Transactions on Systems, Man and Cybernetics: Part C*, vol. 34, pp. 181-186, 2004.

[3] D. Wilkes, R. Alford, R. Pack, R. Rogers, R. Peters, and K. Kawamura, "Toward socially intelligent service robots," *Applied Artificial Intelligence Journal*, vol. 12, pp. 729-766, 1997.

[4] Sony, "Sony Entertainment Robot Europe http://www.aibo-europe.com/," 2004.

[5] C. L. Breazeal, *Designing sociable robots*. Massachusetts: The MIT Press, 2002.

[6] L. Canamero, "Playing the emotion game with Feelix: What can a LEGO robot tell us about emotion?," in *Socially intelligent agents: Creating relationships with computers and robots*, K. Dautenhahn, A. Bond, L. Canamero, and B. Edmonds, Eds. Massachusetts, USA: Kluwer Academic Publishers, 2002, pp. 69-76.

[7] M. Montemerlo, J. Pineau, N. Roy, S. Thrun, and V. Varma, "Experiences with a mobile robotic guide for the elderly," presented at AAAI National Conference on Artificial Intelligence, Edmonton, Alberta, Canada, 2002.

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

[9] M. Scopelliti, M. V. Giuliani, A. M. D'Amico, and F. Fornara, "If I had a robot at home.... Peoples' representation of domestic robots," in *Designing a more inclusive world*, S. Keates, J. Clarkson, P. Langdon, and P. Robinson, Eds. Cambridge, UK: Springer, 2004, pp. 257-266.

[10] J. Goetz and S. Kiesler, "Cooperation with a robotic assistant," Proc. CHI'02 Conference on Human Factors in Computing Systems, New York, USA, 2002.

[11] J. Goetz, S. Kiesler, and A. Powers, "Matching robot appearance and behaviour to tasks to improve human-robot cooperation," Proc. RO-MAN, Milbrae, CA, 2003.

[12] T. Kanda and H. Ishiguro, "Reading human relationships from their interaction with an interactive humanoid robot," *Lecture Notes in Computer Science*, vol. 3029, pp. 402-412, 2004.

[13] K. Dautenhahn, "Design spaces and niche spaces of believable social robots," Proc. IEEE Int. Workshop on Robots and Human Interactive Communication; RoMan, Berlin, Germany, 2002.

[14] B. Friedman, P. H. Kahn (Jr.), and J. Hagman, "Hardware companions? What online AIBO discussion forums reveal about the human-robotic relationship," *Digital Sociability*, vol. 5, pp. 273-280, 2003.

[15] P. H. Kahn (Jr.), B. Friedman, D. R. Perez-Granados, and N. G. Freier, "Robotic pets in the lives of preschool children," Proc. CHI, 2004.

[16] Z. Khan, "Attitude towards intelligent service robots," NADA KTH, Stockholm 1998.

[17] C. DiSalvo, F. Gemperie, J. Forlizzi, and S. Kiesler, "All robots are not created equal: The design and perception of humanoid robot heads," Proc. DIS2002, London, 2002.

[18] T. Nomura, T. Kanda, and T. Suzuki, "Experimental investigation into influence of negative attitudes toward robots on human-robot interaction," Proc. SID 2004, Universiteit Twente, Enschede, The Netherlands, 2004.

[19] S. Woods, K. Dautenhahn, and J. Schulz, "The design space of robots: Investigating children's views," Proc. Ro-Man, Kurashiki, Japan, 2004.

[20] J. Piaget, "Part I: Cognitive development in children: Piaget: Development and learning," *Journal of Research in Science Teaching*, vol. 40, pp. S8-S18, 2003.

[21] A. Druin, "The role of children in the design of new technology," *Behaviour and Information Technology*, vol. 21, pp. 1-25, 2002.

The necessity of enforcing multidisciplinary research and development of embodied Socially Intelligent Agents

Julie Hillan
University of Washington
Department of Technical Communication, College of Engineering
14 Loew Hall, Seattle, WA 98105
julie4@u.washington.edu

Abstract

Today, robots can be provided with gender-specific identities and simulated personalities, emotional responses and feelings, and are able to interact with other agents and establish relationships, making them Socially Intelligent Agents (SIA). Engineers have designed robots that are formed of human-like muscles and skin, and can see, talk, feel and taste. Fashionable home robotics – companion toys -- from Aibo® to Robosapien®, have already become a heavily-marketed reality and fixtures in popular Western and Eastern cultures. A rich multicultural history of folktales, literature and film has popularized certain lay expectations about Human-Robot Interaction (HRI), where relationships range from companion-companion to owner-laborer to victim-nemesis. The diverse academic interpretive communities of social scientists, engineers, artists and philosophers that develop SIA need to collaboratively explore the ethical and social implications of home-companion robot deployment during these early stages of widespread home-use of (even simplistic) embodied Socially Intelligent Agents.

1 Introduction

Artificially intelligent agents imbued with social attributes have long held the public's fascination and have been represented in many films, books and other media. Popular art has helped define the ranges of appropriate human social behavior toward and with robots, in particular. The following excerpt is from an interview with filmmaker Greg Pak, writer and director of *Robot Stories* (Generation5, 2004):

> Generation5: Do you ever think people will come to love a machine – even a machine that isn't anthropomorphic?
>
> Greg Pak: Absolutely. People already love their AIBOs -- I read an interview with a researcher who said that even when people were told ahead of time that their AIBOs were not really sentient and could not really think and learn and feel, people still attached emotional value to them. It makes sense -- kids fall in love with stuffed animals all the time. Adults love their cars, their computers, their guns, their iPods... Human-machine love will run literally out of control when machines can actually think and feel.

It may be unusual to begin an academic paper on embodied Socially Intelligent Agents (SIA) with a quote from a filmmaker. However, Artificial Intelligence (AI) in art typically deals with how humans and embodied AI will develop relationships and interact. The artistic and emotional aspects of human-AI interaction demand to be fully incorporated into academic curriculum and product development and should not be relegated to a lesser or even "separate but equal" status in AI education and research. Art may be considered an outward expression of an individual's interpretation of emotion; therefore, the very act of designing SIAs which display emotion are an artistic and creative process in addition to feats of engineering.

A recently released toy humanoid robot, Robosapien, has become so popular that it has sold out in stores around the world. Besides Robosapien's lower price-tag compared to other widely-released robot toys (approximately $100 USD), this robot is remarkable for its human-like characteristics and foibles, including dancing, whistling and snoring. As a perfect example of art imitating life, the robot's inventor, Mark Tilden, stated in a recent article that he built the seven-motor Robosapien in his own image. Tilden said:

> He is exactly me…even the dance steps, it is me. Now there are 1,4-million of them out there. I'm probably the most prolific father in the history of mankind (Mail & Guardian, 2004).

According to John Jordan, a principal at the Information Technology (IT) consulting company Cap Gemini, "Humans are very good at attributing emotions to things that are not people…Many, many moral questions will arise" (Kageyama, 2004).

Are moral and ethical considerations of human-SIA interaction being considered as part of regular university curriculum or during corporate product development? Perhaps the answer is "sometimes," but it is not consistently practiced. It is necessary to improve communication across the multidisciplinary academic and private sector information silos and to make sincere efforts for separate discourse communities to be collaborative during the current era of SIA technology development because of the complex issues surrounding Human-Robot Interaction (HRI),

1.1 From un-embodied to social

Classical AI researchers have modeled intelligent interfaces -- Web-based, robotic, or otherwise -- on the mathematical processing of infinite logical calculations meant to mimic aspects of Human-Human Interaction (HHI); current AI research explores interfaces that go so far as to replicate humanity's fundamental motivations and emotional states via Human-Computer Interaction (HCI). However, past attempts to define and create artificial personalities rarely draws on the existing body of available psychological literature on human personality. This is indicative of the roots of AI research, which began in the programming community. There is a widespread acknowledgement that the future of AI will involve replicating human traits in computers and embodied artificially intelligent agents such as robots, including the integration of simulated emotional reactions. As the products of AI, robotics and big industry are released to home users in increasingly more accessible and anthropomorphized packages, it is time for separate disciplines to collaborate in development and investigative research involving the human factors.

Many articles review the history of AI, but they are commonly focused on the impact of one discipline to the field as a whole. As a recommendation for broadening how AI is taught, researched and developed, this paper reviews contributions that many academic subjects have made to today's AI community.

2 Computer Science

The technical side of Artificial Intelligence development is well documented. This paper outlines a few significant highlights to illustrate where other disciplines will complement traditional AI research.

In the late 1940's and early 1950's the links between human intelligence and machines was observed and investigated. As Bedi points out in *That's How AI Evolved* (2000), Norbert Wiener introduced the term *cybernetics*, which he used to mean the communication between human and machines. Weiner was one of the first Americans to make observations on the principle of feedback theory. A popular example of feedback theory is the thermostat. Thermostats control the temperature of an environment by gathering the actual temperature of an environment, comparing it to the desired temperature, and responding by turning the heat up or down accordingly. What is significant about Wiener's research into feedback loops is that he posited that all intelligent behavior was the result of feedback mechanisms which could possibly be simulated by machines. This theory influenced a great deal of early AI development. According to Crevier (1999), the first version of a new program called the General Problem Solver (GPS), which is an extension of Wiener's feedback principle, was introduced in 1957; GPS was capable of solving a large amount of common sense problems.

While more programs were being produced, John McCarthy was busy in 1958 developing a major breakthrough in AI history; he announced his new development, the LISP language, which is still used today. LISP stands for LISt Processing, and was soon adopted as the language of choice among most AI developers.

In 1963, MIT received a 2.2 million dollar grant from the United States government to be used in researching Machine-Aided Cognition (what is now commonly referred to as Artificial Intelligence). The grant by the Department of Defense's Advanced Research projects Agency (DARPA) was endowed in an attempt to ensure that the United States would stay ahead of the Soviet Union in technological advancements. The project served to increase the pace of development in AI research by creating a pool of knowledge an international body of computer scientists.

The MIT researchers, headed by Marvin Minsky, demonstrated that when confined to a small subject matter, computer programs could solve spatial problems and logic problems. Other programs that appeared during the late 1960's were STUDENT, which could solve algebra story problems, and SIR that could understand simple English sentences. The result of these programs was a refinement in language comprehension and logic.

During the 1970's many new methods in the development of AI were tested, notably Minsky's frames theory. During this time David Marr also proposed new theories about machine vision; for example, how it would be possible to distinguish an image based on the shading of an image, basic information on shapes, color, edges, and texture (Crevier, 1999).

However, it was in late 1955 that Newell and Card's creation of Logic Theorist, considered by many to be the first AI program, set a long-standing AI development paradigm. Logic Theorist, representing each problem as a tree model, would attempt to solve it by selecting the branch that would most likely result in the correct conclusion. The impact that Logic Theorist made on AI has made it a crucial stepping-stone in the direction of the field.

1.1 Logic

> *One day Alice came to a fork in the road and saw a Cheshire cat in a tree. "Which road do I take?" she asked.*
>
> *"Where do you want to go?" was his response.*
>
> *"I don't know," Alice answered.*
>
> *"Then", said the cat, "it doesn't matter" (Carroll & Woollcott, 1990).*

Intertwined in theory with much early AI programming is the pervasive AI reliance on syllogistic logic. The *Internet Encyclodpedia of Philosophy* (2004) credits Aristotle with inventing this reasoning system. In a nutshell, syllogisms present a deductive form of logic where certain things are stated, and then something different than what is stated is true because of previous assertions.

Mathematician Charles L. Dodgson, who is better known as Lewis Carroll, popularized this logic in his literary examples of satire and verbal wit, which are rich in mathematical and syllogistic humor. Throughout this paper, several of Dodgson's quotes are used to illustrate the human aspect of logic in a lighthearted, yet meaningful, way.

Dodgson also wrote two books of syllogisms; his work provides many interesting examples, such as this one:

> No interesting poems are unpopular among people of real taste
>
> No modern poetry is free from affectation
>
> All your poems are on the subject of soap-bubbles
>
> No affected poetry is popular among people of real taste
>
> No ancient poetry is on the subject of soap-bubbles

The conclusion to this logic structure is then all of the poems the signified author has written are poor.

The advantage to syllogistic logic is that it is something that computers can do well. However, as Clay Shirky points out in his article *The Semantic Web, Syllogism, and Worldview* (2003), much AI research still follows a syllogistic path, when in fact, this reasoning cannot always hold true. Deductive reasoning has been and still is a dominant theme in AI research and development. But, human language, meaning, situations and context are more than mathematical, stilted formulas. Syllogisms are a significant part of AI, but are not a panacea for truly human-like intelligence by themselves.

3 Communications

1.1 Semiotics

> *"When I use a word," Humpty Dumpty said in rather a scornful tone. "It means just what I choose it to means, neither more or less."*
>
> *"The question is," said Alice, "whether you can make words mean so many different things."*
>
> *"The question is," said Humpty Dumpty, "which is to be master - that's all" (Carroll & Woollcott, 1990).*

The role of Communications in Artificial Intelligence is the bridge to the human user factor; this paper summarizes the significance of messages, meaning, context and medium. Historically, research in Artificial Intelligence has been primarily based on text- and speech-based interaction due to the focus on Natural Language Processing (NLP). Continued study of the development of communication in humans will lead to theoretical and practical advances in the construction of computer systems capable of robust communication.

French intellectual Roland Barthes (1968) introduced some terms to the fields of Communication and Literature that can be applied to human-SIA interaction, and in particular, interaction theories surrounding linguistics.

Discourse, according to Barthes, is any interfacing between a subject and another thing that provides information. For example, watching a film or reading a book actively involves a viewer/user involved with creating the film or book imagery in their mind. The viewer/user puts a personal mark upon the film or book content, and the film/book becomes the viewers. Then, the film or book adds or subtracts from the ideas and concepts that the viewer had created.

According to the *Encylopedia of Semiotics* (1998):

> Semiotics represents one of the main attempts -- perhaps the most enduring one -- at conceiving a transdisciplinary framework through which interfaces can be constructed between distinct domains of inquiry. Other endeavors, such as the unified science movement of the 1930s or cybernetics and general systems theory in the 1950s and 1960s, met with only limited success. By contrast, semiotics remains a credible blueprint for bridging the gaps between disciplines and across cultures, most likely because of its own intellectual diversity and pluridisciplinary history, as well as its remarkable capacity for critical reflexivity.

Consider, then, computer semiotics as one potential platform for opening interdisciplinary discourse among SIA development communities.

1.2 Context and Situatedness

> *"Then you should say what you mean," the March Hare went on.*
>
> *"I do," Alice hastily replied; "at least I mean what I say, that's the same thing, you know."*
>
> *"Not the same thing a bit!" said the Hatter. "Why, you might just as well say that 'I see what I eat' is the same thing as 'I eat what I see!' (Carroll* & Woollcott *l, 1990)*

Lucy Suchman's work also spans several academic fields, including communication and sociology. In 1987, she described *situated action* as "actions that are always being taken in the context of concrete circumstances." In Suchman's *Plans and Situated Actions* (1987), she argued that the planning model of interaction favored by the majority of AI researchers does not take sufficient account of the situatedness of most human social behavior. If intelligent machines can communicate effectively with humans in a wide range of situations for a wide range of purposes (teaching, advising, persuading and interacting with), then they will need to accurately take into account actual and possible motivational and emotional states. In other words, embodied AI agents need to determine human contextual possibilities and also learn from their own past and apply their own situational responses.

1.3 Medium

Marshall McLuhan's studies focused on the media effects that permeate society and culture. One of McLuhan's most popularly known theories is what he termed the *tetrad* (McLuhan & Fiore, 1968). Via the *tetrad*, McLuhan applied four laws, structured as questions, to a wide spectrum of mass communication and technological endeavors, and thereby gave us a new tool for looking at our culture.

The four tetrad questions framed by McLuhan are:

1. What does it (the medium or technology) extend?
2. What does it make obsolete?
3. What is retrieved?
4. What does the technology reverse into if it is over-extended?

Because the tetrad was developed to uncover hidden consequences of new technologies, they are an excellent set of questions that may be used as guidelines in many aspects of SIA development, from requirements analysis to ethical considerations. His theories are admirable examples of the fusing of communications, psychology, ethics, and even futurism.

Based on some McLuhanisitc tradition, Clifford Nass and Byron Reeves (1996) theorize that people equate media with real life in a fundamentally social and natural way, and may not even realize that they are doing so. Reeves and Nass have applied experimental techniques such as brainwave monitoring, video, interview, observation and questionnaires to measure human response to media in many forms. Interestingly, their research has been complicated by the fact that their test subjects did not realize that they were responding in a social and natural way to the media, and so did not give valid answers in testing. This fact held true despite any differences in the test subjects, and variations in the media itself.

Nass and Reeves' research results underscores that for media designers a simple way in which to improve their products is to make them more natural to use. Psychological evaluation tools measure response to the media and so evaluate its affect, and by using social science research in order to further media research, new paths can be opened. The most important issue Reeves and Nass unveil is the human tendency to confuse what is real with what is perceived to be real, sometimes glaringly different things. Then, the implications for human-embodied intelligent agent interaction are far-reaching.

The basic theory which Reeves and Nass propose is that the human brain isn't evolved enough to handle modern technology. Up to this point, humans could respond both socially and naturally to other humans, and have not yet developed a biological mechanism for dealing with non-sentient social responses.

There is also a recent academic movement that emphasizes researching the role of emotion in AI and human interaction to emotional AI. Rosalind W. Picard (1997) explains in her book, *Affective Computing*, that a critical part of human ability to see and perceive is not logical, but emotional. Therefore, for computers to have some of the advanced abilities engineers' desire, it may be necessary that they comprehend and, in some cases, feel emotions.

Imitating human-like emotion, countenance and response in intelligent, embodied computer interfaces is becoming accepted across academic communities as a valid, natural and needed branch of modern AI technology. The thesis behind including human-like emotion and personality in artificial agents is that people desire them to behave more like people, essentially so that people do not have to behave like artificially or unnaturally themselves when they interact with computers.

4 Psychology

> *As large as life, and twice as natural (Carroll* & Woollcott, *1990).*

The very term Artificial Intelligence is misleading since even human intelligence is not well defined. Intelligence is often characterized by the properties of thought that are not demonstrated by other organic sentient beings, such as language, long-term planning, symbolic manipulation, reasoning, and meta-cognition.

The most common criterion considered for evaluating whether a computer has achieved human intelligence is the Turing test, developed by British mathematician, logician, and computer pioneer Alan Turing (1950). In the Turing test, a person communicates via a text terminal with two hidden conversational partners: another human and a computer. If the person cannot distinguish between the human and the computer, then the computer would be considered to be behaving intelligently. However, numerous problems with this definition of Artificial Intelligence exist. For example, in practice, when this test is applied, humans are often mistaken for computers. The human-for-computer mistake highlights that message meaning and nuance can be significantly impeded by disembodiment, and that accurate or appropriate written communication interaction may not be the most successful test for true intelligence.

The Intelligence Quotient (IQ) test is often still considered the most accurate way to develop a metric for human intelligence. However, Daniel Goleman (1997) argues in *Emotional Intelligence: Why it can matter more than IQ*, that intelligence measurement should not be limited by an IQ definition. Goleman presents a case for Emotional Intelligence (EQ) actually being the strongest indicator of human achievement. He defines Emotional Intelligence as shades of "self-awareness, altruism, personal motivation, empathy, and the ability to love and be loved by friends, partners, and family members." Emotional Intelligence also encompasses a set of skills that includes impulse control, self-motivation, empathy and social competence in interpersonal relationships.

Goleman states that people who have high emotional intelligence are people who succeed in work as well as personal lives; these are the type of people who build successful careers and meaningful relationships.

Ubiquitous, autonomous affective computer agents that operate by recognizing images will eventually be able to seamlessly detect, learn, mimic, process and react to human facial expressions, gaze, body posture and temperature, heart rate and many other human criteria. If robotic agents are to be able to sense, respond to, or model these types of affective states, and also have rich and subtle linguistic abilities and a deep understanding of the structure of human minds while appearing human in countenance, they will also be replicating levels of EQ.

Computers may never truly experience human emotions, but if current theories are correct, even modest emotion-facsimile applications would change machines from purely reactionary automatons into persuasive actors in human society. If intelligent machines can communicate effectively with humans in a wide range of situations for a wide range of purposes (teaching, advising, appealing to and generally interacting with), then they will need to accurately take into account actual and possible motivational and emotional states.

The topic of human intelligence (defined) and personality is important for future Al research on self-aware and socially aware agents. Psychologists have been attempting to define personality, and identifying how particular people will respond to various personality types, for many decades.

Looking back in years to come, engineers will find that much of the current research methodology for developing SIAs is inadequate, especially development models involving explicitly labeled emotional states and special emotion-generating rules. People will respond to these generated/artificial personalities in the same way they would respond to similar human personalities. It is therefore necessary for AI to draw from technical and empirical disciplines and integrate the complex nuances of both early in the process of AI personality development.

5 Usability

Usability issues, which may fall under communication, psychology, Human-Computer Interaction or engineering depending upon the primary interpretative community one is affiliated with, are critical in the interaction of many AI systems where a human user works with the system to find and apply results and when the AI system also serves as the user interface.

Standards for HCI and usability have been developed under the supervision of the International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC). The standard ISO 13407 (1999) explains the elements required for user-centered design:

> This standard provides guidance on human-centered design activities throughout the life cycle of interactive computer-based systems. It is a tool for those managing design processes and provides guidance on sources of information and standards relevant to the human-centered approach. It describes human-centered design as a multidisciplinary activity, which incorporates human factors and ergonomics knowledge and techniques with the objective of enhancing effectiveness and efficiency, improving human working conditions, and counteracting possible adverse effects of use on human health, safety and performance.

The official definition as stated above explains that from a usability standpoint, human-centered design is a multidisciplinary activity. Therefore, a cross-discipline approach to SIA and companion robot development will be the most successful method of producing a truly user-friendly agent.

6 Ethics

> *You shouldn't anthropomorphize robots. They don't like it (Anonymous).*

Ethics tries to evaluate acts and behavior. Are there ethical boundaries on what computers should be programmed to do? In 1993, Roger Clarke, a fellow at the Australian National University, wrote an essay questioning how Asimov's *Laws of Robotics* (1942) applied to current information technology. Of interest here is his emphasizing the idea that AI agents may be used in ways not intended by their designers, and then the scope of interaction takes an unpredictable turn.

One of Clarke's conclusions was:

> Existing codes of ethics need to be re-examined in the light of developing technology. Codes generally fail to reflect the potential effects of computer-enhanced machines and the inadequacy of existing managerial, institutional and legal processes for coping with inherent risks.

Clarke’s quote illustrates very clearly the overwhelming need for ethics to be an important consideration in intelligence research and development.

Human-Computer Interaction (and all of the specialties mentioned here that touch on HCI) is not only the study of the end-use of human users working with technology – HCI also considers the design and requirements analysis of the product before it is developed and the interaction of engineer/creator and computer. The field of ethics, a broad academic discipline with a long history of theories about human behavior, does not have one divine theory that can be applied as the essential unifying framework across cultures or sub-cultures. How then can engineers, scientists and teachers begin to consistently apply ethical development of humanoid/robotic SIAs? What will happen if people use robots and SIAs --who will be able to interact almost seamlessly with humans and are capable of behaving as if they love and hate and interact in a meaningful way -- to substitute human relations?

7 Conclusion

> *The ideal engineer is a composite ... He is not a scientist, he is not a mathematician, he is not a sociologist or a writer; but he may use the knowledge and techniques of any or all of these disciplines in solving engineering problems (Dougherty, 1955).*

The scope of this paper could have gone on to cover many other areas of study that are contributing to the development of AI, including neurosciences, bioinformatics, evolutionary theory, biology and philosophy of mind. As all of these disciplines merge into truly effective AI, the need for overarching theories and unified approaches becomes more important.

Although scientists and researchers investigate the intricate hard science that is the framework of real intelligent agents, the majority of people acting as users who will encounter ubiquitous SIA will only react to the outward interface. As more embodied intelligent agents incorporated with social affects are developed, mass-marketed and placed into human-agent social context, the issues surrounding the role of emotions in Human-Computer Interaction need to be addressed more proactively, concurrently and iteratively with development -- not only after products are on the market.

Acknowledgements

Many thanks to Roger Grice (Rensselaer Polytechnic Institute) and Clifford I. Nass (Stanford University) for their countless helpful comments and discussions.

References

Asimov, I. (1942). *Runaround*. New York: Street & Smith Publications.

Barthes, R. *Elements of semiology*. (1968). New York: Hill & Wang.

Bates, J. (1994).*The role of emotion in believable agents*. Communications of the ACM, vol. 37, pp. 122-125.

Bedi, J. (2003). *That's how AI evolved*. The Tribune (India). Monday, Sept. 8.

Bouissac, P. (Ed.). (1998). *Encylopedia of semiotics.*New York: Oxford University Press.

Carroll, L. & Woollcott, A. (1990). *The Complete Works of Lewis Carroll*. Random House: New York.

Cassell, J. (2000). *More than just another pretty face: Embodied conversational interface agents*. Communications of the ACM (to appear). MIT Media Laboratory.

Clarke, R. *Asimov's laws of robotics for information technology*. IEEE Computer. 26,12 (December 1993) pp.53-61 and 27,1 (January 1994), pp.57-66.

Crevier, D. (1999). *AI: The tumultuous history of the search for Artificial Intelligence*. Basic Books: New York.

Dougherty, N.W. (1955).

Fogg, B.J. (2002). *Persuasive technology: Using computers to change what we think and do.* Morgan Kauffman: New York.

Generation5. *Greg Pak.* Retrieved March 19, 2004: http://www.generation5.org/content/2004/pak.asp

Goleman, D. (1997). *Emotional intelligence: Why it can matter more than IQ*. Bantam: New York.

International Organization for Standardization. *ISO 1347*. Retrieved April 18, 2004: http://www.iso.ch/iso/en/ISOOnline.openerpage

International Electrotechnical Commission. *ISO 1347*. Retrieved April 16, 2004: http://www.iso.ch/iso/en/ISOOnline.openerpage

Internet Encyclopedia of Philosophy. Aristotle (384-322 BCE.), Overview. Retrieved April 17, 2004: http://www.utm.edu/research/iep/a/aristotl.htm.

Kageyama, Y. In gadget-loving Japan, robots get hugs in therapy sessions. *The San Diego Union-Tribune.* April 10, 2004.

Mail & Guardian Online. *Belching, farting robot outsells 'em all.* Hong Kong. Retrieved November 1, 2004: http://www.mg.co.za/Content/l3.asp?cg=BreakingNews-Andinothernews&ao=124743

McLuhan, M. and Fiore, Q. (1968): *War and peace in the global village*. New York: Bantam

Picard, R. (2000). *Affective Computing*. MIT Press: Boston.

Reeves, B. & Nass, C. (1996). *The media equation: How people treat computers, television, and new media like real people and places*. Cambridge University Press: New York.

Shirky, C. *The Semantic Web, Syllogism, and Worldview*. Retrieved February 9, 2004: http://www.shirky.com/writings/semantic_syllogism.html. First published November 7, 2003 on the "Networks, Economics, and Culture" mailing list.

Suchman, L. A. (1987). *Plans and situated actions: The problem of human-machine interaction*. Cambridge University Press.

Turing, A.M. (1950). *Computing machinery and intelligence*. Mind, 59, 433-460.

Human-Robot Interaction Experiments: Lessons learned

Cory D. Kidd[*]
[*]MIT Media Lab
Cambridge, MA 02139
coryk@media.mit.edu

Cynthia Breazeal[†]
[†]MIT Media Lab
Cambridge, MA 02139
cynthiab@media.mit.edu

Abstract

This paper presents lessons learned in designing and performing human-robot interaction experiments based on our work. We present work in adapting techniques from other disciplines, developing new study designs, methods for results analysis, and presentation of the results. We discuss both successful and unsuccessful study methods in an attempt to pass on what we have learned to the HRI community.

1 Introduction

Studying human-robot interaction is a relatively new area of research that presents challenges not found in other fields. During the past three years, we have drawn on other work and learned a number of lessons that we believe will be useful for other researchers.

Understanding how people interact with robots, how people interpret human-like signals from robots, how the physical embodiment of a robot affects an interaction, and how robots are seen differently from other technologies that can be utilized in interactions are all questions that interest us as researchers.

2 Influences on HRI

Early work in HRI evaluation includes work at Carnegie Mellon University (Goetz and Kiesler, 2002) and the University of Washington (Kahn Jr. et al., 2002). Another influence is the work of Reeves and Nass, which they summarize in their 1996 book (Reeves and Nass, 1996). Many studies carried out in their work have analogues in questions that interest HRI researchers. We have drawn on other computing-related fields of work such as those on anthropomorphic interfaces (Bengtsson et al., 1999), understanding user perceptions in HCI, and the personification of interfaces (Koda and Maes, 1996).

Another important area to draw from is psychological research, from concepts and measurement scales to theories and experimental protocols. Finally, we have looked at the self report questionnaire-based methods developed in the communications research field as well as physiological measures.

3 Measurement types

3.1 Self-report measures

One of our preferred measures has been self-report questionnaires. The scales that have been used have come from a variety of sources: similar experiments, equivalent measures for human-human interaction, and scales we developed.

The advantage of using these measures is the ease of gathering and analyzing the data. The analysis of this data is straightforward, involving basic statistical techniques.

There are two main drawbacks to using these measures in HRI work. The first is that there are few scales that have been designed that can simply be taken and used in an experiment. Often, scales must be carefully designed to measure the aspect of an interaction that is important. The second difficulty is that the data is self-reported by subjects. There are known problems with this that are addressed elsewhere along with common solutions.

3.2 Physiological measures

In early work, we used physiological measures such as galvanic skin response. The advantage of measuring a physiological signal is that it is difficult for a person to consciously control autonomic activities. There are numerous difficulties in using physiological measures, however. A major problem is the gathering of reliable data from a sensor attached to a subject in a real-world HRI scenario. Another issue that must be confronted is that there are many confounds for most physiological signals.

3.3 Behavioral measures

Another type of measure we use is behavioral data. This includes data gathered from a subject's activities during an experiment, often using video tape or software logging. Examples we have used include time spent looking at the robot, mutually looking at the same object with the robot (Sidner et al., 2004) and time spent in free-form interaction with a robot (Kidd et al., 2005).

We initially chose behavioral measures because of their predominance in psychological and communications research studies of HRI. This work measures aspects of interactions that are important to HRI and the metrics can easily be adapted to HRI work.

Two difficulties in using these measures are that it is time-consuming to gather data and requires independent coders for the data.

4 Design recommendations

We recommend combining self-report questionnaires and behavioral measures to gather data in an efficient and robust manner. Questionnaires must be carefully adapted, designed, and tested before beginning an experiment. We have found it indispensable to gather and analyze data on test subjects before running a full experiment.

HRI experiments are unusual in that subjects are often excited about being able to interact with a "real" robot for the first time. To address this and other issues, we suggest the following protocol: (1) introduce the subject to the experiment and the robot, (2) let the subject attempt any portions of the interaction which may require assistance and allow the subject to become familiarized with the robot, (3) start a video camera to record the interaction, (4) allow the subject to complete the interaction, (5) administer a questionnaire to the subject, (6) complete a recorded interview with the subject, and (7) debrief the subject on the aims of the experiment.

This protocol allows the subject to become familiar with the experimental setup. Parts (1) and (2) can be extended to reduce novelty effects if that is a strong concern. Parts (3) through (6) allow the gathering of data.

5 Conclusions

We have found that HRI studies provide challenges not found in HCI work. Not only must the experimental protocol be well-developed as in those studies, but we must ensure that the entire robotic system (perception, control, and output) is prepared to "participate" in the study as well. In addition, we run our experiments with the robots under autonomous control, so there is no way for the human to step in and assist the robot, which means the system must be robust for many users before beginning the experiment.

We have presented some lessons learned from designing and conducting several HRI experiments that we hope will be useful to the HRI community. As this field matures, it is important to develop standardized methods of evaluating our work. We believe that a combination of drawing on established scientific fields and development of our own methods where appropriate is the best course to take.

References

Bjorn Bengtsson, Judee K. Burgoon, Carl Cederberg, Joseph A. Bonito, and Magnus Lundeberg. The impact of anthropomorphic interfaces on influence, understanding, and credibility. In *32nd Hawaii International Conference on System Sciences*, 1999.

Jennifer Goetz and Sara Kiesler. Cooperation with a robotic assistant. In *Conference on Human Factors In Computing Systems (CHI 2002)*, pages 578–579, Minneapolis, MN, USA, 2002.

Peter H. Kahn Jr., Batya Friedman, and Jennifer Hagman. "i care about him as a pal": Conceptions of robotic pets in online aibo discussion forums. In *CHI 2002*, pages 632–633, Minneapolis, MN, 2002. ACM.

Cory D. Kidd, Andrea Lockerd, Guy Hoffman, Matt Berlin, and Cynthia Breazeal. Effects of non-verbal communication in human-robot interaction. In *Submitted to IROS 2005*, Edmonton, Alberta, Canada, 2005.

Tomoko Koda and Pattie Maes. Agents with faces: The effect of personification. In *Fifth IEEE International Workshop on Robot and Human Communication*, pages 189–194. IEEE, 1996.

Byron Reeves and Clifford Nass. *The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places*. Cambridge University Press, Cambridge, England, 1996.

Candace Sidner, Christopher Lee, Cory Kidd, and Neal Lesh. Explorations in engagement for humans and robots. In *Humanoids 2004*, Santa Monica, CA, USA, 2004.

Ethical Issues in Human-Robot Interaction

Blay Whitby

University of Sussex
Brighton
blayw@sussex.ac.uk

Abstract

Scientists and engineers involved in Human-Robot Interaction design need to pay far more attention to the ethical dimensions of their work. This poster briefly presents some key issues.

It might seem, at first glance, that the design of robots and other intelligent systems which have more human-like methods of interacting with users is generally to be welcomed. However, there are a number of important ethical problems involved in such developments which require careful consideration. Given present progress and promises in robotic systems that interact with humans in complex and intimate fashion there is a need to discuss and clarify as many as possible of these problems now.
Since humans tend to adapt to technology to a far greater extent than technology adapts to humans (Turkle 1984), it is reasonable to expect that this process of adaptation will be especially noticeable in cases where robots are used in everyday and intimate settings such as the care of children and the elderly.

More human-like interaction with robots may seem a worthy goal for many technical reasons but it increases the number and scope of ethical problems. There seems to be little awareness of these potential hazards in current HCI research and development and still less in current research in Human-Robot Interaction. Current codes of practice give no significant guidance in this area. This is despite clear warnings having been offered (for example in Picard 1998 and Whitby 1996)

The principles of user-centred design – rarely followed in current software development - are generally based around the notion of creating tools for the user. In human-robot interaction, by contrast, we might sometimes better describe the goal as the creation of companions or carers for the user. This requires deliberate and properly informed attention to the ethical dimension of the interaction.

Designers can and do force their view of what constitutes an appropriate interaction on to users. In the field of IT in general there have been many mistakes in this area. Some writers (e.g. Norman 1999) argue that there is a systematic problem. Even if we do not grant the full force of Norman's arguments there would seem be cause for anxiety about this particular aspect of human-robot interaction.

There is an obvious role for ethicists in the design of Human-Robot Interaction. In other areas the introduction of artificial Intelligence and similar technologies has often resulted in a movement of power towards those at the top (see Whitby 1996). Unless Human-Robot Interaction designers are constantly reminded of the difficult relationship between matters of fact and matters of value and of the importance of empowering users, this pattern may be repeated with even more unfortunate consequences in the area of personalized robots.

References

Norman, D. (1999) The Invisible Computer, MIT Press, Cambridge, MA.
Picard, R. (1998) Affective Computing, MIT Press, Cambridge MA
Turkle, S. (1984) The Second Self, Computers and the Human Spirit, Granada, London
Whitby, B. (1996) Reflections on Artificial Intelligence, The legal, moral, and ethical dimensions, Intellect Books, Exeter.