Sonderforschungsbereich 378
Ressourcenadaptive kognitive Prozesse

KI-Labor am Lehrstuhl für Informatik IV

Leitung: Prof. Dr. Dr. h.c. mult. W. Wahlster

Universität des Saarlandes
FR 6.2 Informatik IV
Postfach 151150
D-66041 Saarbrücken
Germany

Tel.: +49 - 681 / 302 - 2363

BAIR

Memo Nr. 86

Modelling and Designing User Assistance in Intelligent Environments (MODIE 2006)

in Conjunction with the 8th International Conference on Human Computer Interaction with Mobile Devices and Services (MobileHCI 2006)

Thomas Pederson, Helder Pinto, Michael Schmitz,

Christoph Stahl, Lucia Terrenghi (Eds.)

September 2006

Sonderforschungsbereich 378
Ressourcenadaptive Kognitive Prozesse

ISSN 0944-7822

86

Einbandgestaltung, Druck und Verarbeitung: Universität des Saarlandes
Printed in Germany

ISSN 0944-7822

MODIE

Modelling and Designing User Assistance in Intelligent Environments

Workshop at Mobile HCI 2006

September 12, 2006, Espoo, Finland

Organizing Committee:

- Christoph Stahl (co-chair)
  Artificial Intelligence Group, Saarland University, Germany
  http://w5.cs.uni-sb.de/staff/show/stahl
  stahl@cs.uni-sb.de
- Thomas Pederson (co-chair)
  Department of Computing Science, Umeå University, Sweden
  http://www.cs.umu.se/~top/
  top@cs.umu.se
- Helder Pinto
  Telecomputing Engineering research group, University of Minho, Portugal
  http://piano.dsi.uminho.pt/~helder/index.html
  helder@dsi.uminho.pt
- Michael Schmitz
  Intelligent User Interfaces, DFKI GmbH, Germany
  http://www.dfki.de/~schmitz
  schmitz@cs.uni-sb.de
- Lucia Terrenghi
  Media Informatics Department, *University of Munich, Germany*
  https://wiki.medien.ifi.lmu.de/Main/LuciaTerrenghi
  lucia.terrenghi@ifi.lmu.de

An online copy of this document can be obtained from the workshop website:
http://www.ai.cs.uni-sb.de/workshops/modie06/

Background and Motivation

Ubicomp research continually develops novel interaction techniques, sensing technologies, and new ways of presenting personalized information to the user. Gradually, companies operating in environments such as airports, museums or even shopping malls are becoming aware of the potential benefits in letting such technologies assist their users and customers. Intelligent environments are predicted to aid their users in pursuing their activities, such as wayfinding or shopping, through the situated presentation of personalized information. However, due to the large design space that ranges from wearable computing to public displays, the conceptual and technological choices pose new challenges to the designer of such user-assistance systems.

The MODIE workshop aims towards models, principles and methodologies, which guide the designer of an intelligent environment in the early stages of the development process, such as task and requirements analysis and conceptual design:

- Which user activities and tasks require assistance?
- How to map an activity model into interactions with computing artifacts?
- How should the designer choose the best sensing and interaction technologies for a scenario?
- Which mobile or wearable personal devices can be employed?
- How should multiple users with concurrent activities be supported?
- How should the current state of the user assistance system be represented, especially when dealing with multiple tasks?

The first session of the workshop covers user-oriented design methods and provides some interesting case studies. The second session covers the modelling of user assistance, user activity and multimodal interaction. The third session specializes on the evaluation of system prototypes in virtual- or mixed-reality and investigates how intelligent environments can be tested before they are actually deployed.

Besides the presentations, the intention of the workshop is to discuss some of the following topics:

- How to unify the complementary concepts of public and personal devices in Intelligent Environments (IEs)
- How to model user activity (terminology, structure, notation) for the design of IEs
- Suggest a terminology for Intelligent Environments
- How tools can support the modelling and designing of user assistance in IEs
- What are the problems of applying traditional software engineering methodologies
- Are there principles that can be generalized for the design of all IEs

The MODIE Workshop will be held in conjunction with the 8th International Conference on Human Computer Interaction with Mobile Devices and Services, MobileHCI 2006, which will foster a discussion on how the contrasting paradigms of mobile/personal and ubiquitous computing can converge into user assistance.

Finally, we wish to thank our co-organizers for their reviewing support and fruitful discussions, which helped to raise interesting research questions and to define the scope of this workshop. Our thanks also go to the MobileHCI workshop chair, Giulio Jacucci, and of course to the participants for their contributions.

Espoo, September 12th 2006,

Christoph Stahl and Thomas Pederson (co-chairs)

Schedule and Table of Contents

User-Oriented Methods for Early Design and Evaluation of Intelligent Environments

Arjan Geven
CURE - Center for Usability Research & Engineering
Hauffgasse 3-5
1110 Vienna, Austria
geven@cure.at

Johann Schrammel
CURE - Center for Usability Research & Engineering
Hauffgasse 3-5
1110 Vienna, Austria
schrammel@cure.at

Manfred Tscheligi
CURE
&
University of Salzburg, ICT&S Center
tscheligi@cure.at

ABSTRACT

In this paper, we present our position on user experience methods for designing and evaluating intelligent environments in a user assistance context; a combination of narrative interviews, a scenario-based approach and finally Wizard-of-Oz prototyping. Users typically do not have any experience with intelligent environments, making it difficult to reliably assess user experience with such systems. An integrated scenario helps us focus on the user from the beginning on, narrative interviews help us understand the user in a very early phase of development and Wizard-of-Oz prototyping allow us to test early and often in the design of intelligent environments.

The application of these three methods in a specific project on cognitive vision is described, as well as their advantages and disadvantages as methods for assessing user experience in intelligent environments in general.

1.INTRODUCTION

Cognitive vision is the name of the combined technology that allows computer systems to 'see' and to make sense out of what they see. The computers are then able to acquire knowledge about objects and activities in the environment (the vision part of cognitive vision) and use this knowledge to improve the interaction and better serve users needs (the cognitive part of cognitive vision). Through its cameras, it can connect the real physical world with the virtual computational world and creates possibilities for proactive instead of reactive systems, systems that can detect, locate, recognize and understand objects and situations in the real world [8]. A cognitive vision system can show purposive goal-directed behavior, can adapt to unforeseen changes, and can anticipate the occurrence of objects and events [5]. This technology is still under heavy development and can be seen as the next step in computer development [3]. The introduction of systems that can see their environment and understand it requires a paradigm shift in the way we interact with a system. As computers acquire more human capabilities, human-machine-interactions can more and more approach human to human interaction instead of the more traditional way of interaction (GUI-style). For instance, a personal assistant in the office can be created using this technology 9. This assistant can take the form of an agent that can help the user locate his or her keys, books or notes, can automatically give background information to something the user touches, give deadline reminders, et cetera. In short, it can bring some advantages of the digital world into the physical world.

In this position paper, we present the three methods we have used successfully to analyze user interaction with an intelligent system designed to provide user assistance, in different phases of the development process. The first phase consisted of scenario-development to guide design and provide a general framework for the project development. This scenario helps us focus on the user already very early in the design process. Although this phase is focused on the user, users themselves are not directly involved in this phase. The second phase is the first phase in which users are included, where we investigate users' experiences and emotions evoked by current technology based on narrative interviews to create a better understanding of who the users are and how they feel about intelligent systems and how they experience and interact with technology. This phase includes no technology as such yet. In the third phase, then, still very early on in the development process, we experiment with user interaction with a 'real' cognitive vision system, by means of a wizard-of-oz prototype. We will discuss these three methods that allow us to gain insight in user experiences with intelligent technology in a very early phase of its application.

2.THE SCENARIO-BASED APPROACH

The first step we took in the cognitive vision project is based on providing an integrated scenario for guiding and evaluating the concept, the technology and the interactions associated with cognitive vision technology. Users themselves are not involved in this early phase in the design process. Instead, the scenario-approach can be seen as a meta-approach to the project, where it aids user-centered development, provides strong focus and makes challenges clear.

The development of (cognitive) vision technologies is making constant progress and first applications for the end-users can be expected soon. Typically the development of such applications is only driven by the availability of new technology and only minor effort is made to design the interaction between the user and the system. Here scenarios come in, as they have the power to visualize systems that do not exist yet vividly and can present possible ways to interact with them. Our starting point was that scenario-based process models could help significantly to deal with the characteristic challenges in this area provided the approach is adapted carefully to the specific requirements in this context, like privacy, usefulness, trust, but also metaphors, or interaction style.

There are three main reasons why our project approach for the intelligent environment is based on an integrating scenario.

Firstly, a *scenario-based approach helps to focus on interaction issues right from the beginning* of research and development and to identify the challenges for interaction. Issues related to human system interaction are of special relevance for intelligent environments and should be tackled systematically. These technologies have the potential to radically transform interaction paradigms, e.g. every object might become an input device with the use of advanced object and gesture recognition. This provides great opportunities for the interaction design but also great challenges - how is the user supposed to know which objects he can use; what makes sense; what is new, what old? Scenarios help us focus on these issues instead of only focusing on the technological side of development.

Secondly, a project in this context faces several challenges on the organizational level similar to industrial design. Multiple institutions are involved, they naturally have their own research agendas and therefore the overall goal of the project might be interpreted differently. Fast, easy and targeted communication is of major importance for the success of a complex project. One has to ensure that the involved stakeholders/parties can use a common language that bridges the different domains with specialized meanings. Work takes frequently place in parallel, geographically distributed and spread across different organizations. Also a project in the (applied) research context implies that the development process can not be planned in the same detail and with the same confidence as in an industrial development process. Scenario-based approaches are suited very well to address these challenges. *Scenarios have the ability to provide an overall guidance, to foster communication, to integrate evaluation by deriving test cases from the scenarios, and to support documentation* [9]. Our approach supported the stakeholders involved in this project by the following means: (1) it provided a well defined basis for the project evaluation, (2) it facilitated the documentation of the project's progress, (3) and finally it supported the iterative reshaping of the project plan.

Thirdly, in research, different techniques are typically developed using different assumptions about their context of use, so combining two techniques often involves modifying and adding new domain specific elements to the design [2]. Due to structural difficulties of integrating various techniques there is the risk to ignore challenging issues. For example, in the domain of intelligent systems, Brooks [1] expressed the concern that people are actually ignoring the true substance of intelligence if they do not focus on the interface between low- (e.g. eyebrow movement) and high-level issues (e.g. dialog turn-taking). By *making explicit the intended use contexts and by providing a guiding vision that has the capability to focus the work on all required issues*, scenario-based design helps to deal with these issues.

We based the design and development process of the cognitive vision project on a constitutional scenario. With constitutional scenario we mean a relatively general but in its internal logic cohesive task domain. To successfully serve as a constitutional scenario such a task domain must have certain characteristics: a) It must reflect the scientific key challenges from the targeted areas, b) it needs to be able to be applied to different domains, and c) it ideally also allows including aspects of human computer interaction. Additionally, we defined one multidimensional design space instead of sets of scenarios. To do so the constitutional scenario is carefully analyzed with regard to relevant variables influencing its complexity. The goal is to decompose it into several complexity dimensions that allow characterizing the task domain of the scenario in a multidimensional space. These dimensions then are studied with regard to their interrelation and finally systematically organized. For all dimensions characteristic values with increasing complexity are identified.

The long-term goal of our specific cognitive vision project envisions a scenario in that every person will interact in a natural way with artificial devices as an aid in daily life situations such as orientation, search and information retrieval. We refer to it as t*he Personal Assistance scenario*, where a combination of mobile devices and distributed ambient spaces unobtrusively support users by being aware of the present situation and by responding to user requests. Specifically, the technical goal is to devise Cognitive Vision methods to support scenarios that understand and support human daily-life activities. The Personal Assistance scenario requires cognitive abilities such as detection and recognition, spatial and temporal reasoning, embodiment of visual processes, and memory.

For us, the scenario has two additional advantages in this specific project. Firstly, that the *research and development challenges become very clear in an early stage of the project*. The common definition and structuring of the complexity dimensions by the research partners triggered targeted and fruitful discussions on conceptual key challenges and how to approach them. The definition of the dimensions also partly served as a starting point for conceptual work on how to approach these challenges. It allowed to easily identify key challenges not tackled by any partner very early and made this fact available for management. Division of labor between partners with related research interests could be done more rationally and with the common goal in mind.

The second advantage is that the integration of work could be done more easily as the different organizations already started working on their subparts with *a clear common goal* and application context in mind. Possible interfaces between the different techniques were also identified early and could be considered already during the development of the different techniques and therefore later efforts in adapting the methods could be minimized. For example the different developed technologies already assumed the need for the exchange and communication of certain tuning parameters and confidence values. However, making challenges clear doesn't mean somebody covers them. Even if the approach can be very helpful in identifying not tackled challenges special effort and management is needed to ensure this aspect is taken care of.

The definition of the complexity dimensions implies the risk to only tackle them on a one by one basis and to forget about interferences between them. Therefore one should take care to also analyze interaction effects and interdependencies between different dimensions.

3.NARRATIVE INTERVIEWS

Still in a very early phase, users can be made part of the project in a rather general way, unrestricted by technical preconditions, to provide valuable input on the new domain of intelligent environments and user interactions. This domain is rather tricky, as it is difficult to do user testing with users of intelligent environments. Interviews and focus groups and other similar methods are not easily applied as users are not familiar with the technology yet.

Narrative interviews can help us realize what users experience and what they feel when they interact with technology and can help us gain a better understanding of today's experiences that take place in a real context when interacting with technology, in order to get more feeling for how interaction with intelligent systems might be experienced by users. They also provide us with a soft measure of how intelligence in current technology is seen and judged by users. Furthermore, they tell us which interaction factors are mainly responsible for creating a positive user experience, which can be applied in the development process.

The narrative interview is an interview approach that is focused on starting narrations about real-life experiences, based on the work of [7]. We selected this method from a number of other methods for gathering user insight, as user experience research has no clear methods for assessing widespread and real-life experiences. Although many methods have been and are being devised, this particular focus was not found. The focus on eliciting narrations allows us to make use of the structural peculiarities story-telling follows, ie. that the emotional content of the story is re-enacted during the narration. Stories provide a more direct access to the experience than evaluative questions. With stories as base material the analysis can also consider structural elements of the narrations and characteristics of the used language.

The general goal of the study [6] was to better understand the experiences of the interaction with systems of all kinds, e.g. mobile devices, robots, personal computers, PDAs and consumer electronics. Each interview started with open questions about "emotional encounters with technology" which introduce the interviewee to the focus of the interview and creates the right mindset for follow-up questions. Users were asked to remember any situation with technology in which they experienced emotions. They were asked to recount these memories in detail and to induce stories as complete as possible. After these relatively unfocused questions, we asked participants for negative and positive experiences, as well as special emotional and user experience factors that were selected based on previous user experience work such as connectedness to other people and sharing experiences with others, feeling intimate with a system, trust in a system and flow. For each factor, participants were asked to narrate stories about situations in which they experienced it and elaborate on the precise circumstances under which the situation occurred.

Using the narrative method, we were able to identify interesting phenomena in everyday experiences evoked by today's technology, e.g. the overlap between emotion theory and technology practice as well as the differences between them, the dominance of negative experiences and the influence of usage on the user-system relationship. One main result of our analysis regarding the intelligence of current technology is that people didn't tend to characterize systems as intelligent at all. Attributions like "intelligent" or "clever" were not found anywhere in the interviews whereas characterizations like "stupid" or "dull" appear from time to time. On the other hand interviewees frequently mentioned negative and annoying experiences with systems that behaved "pseudo-intelligent". The typical dramaturgy in this cases consisted of the arousal of expectations regarding the system which then was disappointed. What is characterized as intelligent system by researches and developers doesn't mean to be filed the same way by users.

Applying this method in the context of technology teaches us some interesting things. Firstly, it is interesting to see how very similar the stories told by our participants were. We expected that participants would tell us many stories with very different kinds of experiences and emotions, but it was relatively easy to make sense out of the gathered data.

A lesson we learned was that it is not easy to bring people to really switch from "reporting" an event to telling a story. This is not only a theoretical difference, but it's practically also very important, as only story-telling allows participants to really re-live the experience including the related emotions that came up at the moment of the experience.

Participants are generally not familiar with an interview method in which they are asked questions very freely. This can create the uncomfortable situation where the participant wants to answer the request of the interviewer, but is afraid to give a wrong answer, and thus decides it'd be better not to say anything at all. These situations are not uncommon in free association interviews, but can be avoided by giving the participant some focus points for what kind of experiences you're looking. The participants then can go through their memory more easily, searching for experiences that fit specific leads instead of experiences in general and feel more confident to narrate about the experience. In our study, we used very broad focus points as 'positive experiences' or 'negative experiences', which give direction to what we are looking for, but still provide enough room for association on the participant's side to come up with 'free' past experiences.

As people get more comfortable with telling stories about their experiences, the questions also got more personal, and participants were able to find events in which they experienced the broadly described situations. Then, we found that it is useful to ask for extreme events; the 'best experience' and the 'worst experience' provide more powerful expression and allow for easier interpretation than just any positive or negative experience. Finally, we found it easier to reach closure in the interviews when we balanced the questions about experiences; about equal time was spent on both positive and negative experiences. This balanced the interviews and e.g. avoids it from turning into a rant against unusable technology. By putting a positive experience next to a negative experience, it made it easier for the participants to also see where the interview was going.

These narrative interviews help us to gain insight in user experiences with current technology, and provide meaningful insights on user experiences with intelligent systems .

4.WIZARD-OF-OZ PROTOTYPING

In a third step of the project we did focus some more on the technology side of the intelligent environment, with a simulation technique that allows us to experiment with user interactions with intelligent systems before the systems were in an advanced stage of development, with the help of a Wizard-of-Oz prototype.

Prototyping intelligent environments by means of simulation can provide useful information on user interaction with intelligent environments. Wizard of Oz prototyping does just that. [4] describe Wizard of Oz studies as a way to study user interaction with natural language systems.

A Wizard of Oz study is a study where subjects are interacting with a real intelligent system, but the intelligent system is simulated. Behind the interface is not an intelligent system, but instead the interaction is mediated by a member of the design team, the wizard, who performs the actions the intelligent system would take and provide the user with these actions through the interface. The large advantage of this approach is that the subject can be given much more freedom of expression.

The study we did was designed to find out how users interact with a personal (embodied) assistant that can see what the user is doing and react and help the user where necessary, in two different tasks.

The goal of the first task was to test user responses in a setting where the system had knowledge about the location of certain objects throughout the office (e.g. sticky tape). The participants were told that objects had been hidden by a previous participant under the careful eye of the cognitive vision system, which remembered the position of the objects. The system then told the user where to look for the specific item (e.g. “in the drawer on your right”). In each condition, the computer-voice, controlled by the Wizard, said where the target object could be found.

In the second task, participants were asked to assemble a 3d-structure puzzle. The puzzle consisted of six separate pieces, which, if fitted together in the right way, formed a cube, as depicted in Figure 1. There was only one way to solve the puzzle. The participants were asked to solve the puzzle as fast as possible. Participants were seated in front of the computer and everything was observed by the Wizard, who gave hints on solving the puzzle by means of a computer generated voice. The wizard typed the text, and a text to speech algorithm would then read the text out loud, to increase the realism simulated by the test setup. Each piece of the puzzle was numbered and color-coded for ease of recognition.

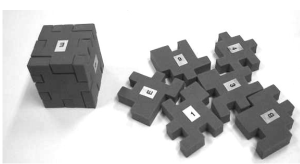

Figure 1: The 3d-Structure Puzzle.

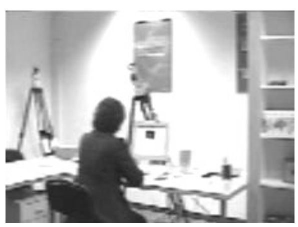

Figure 2: The user being observed by the Wizard of Oz.

The hints that were given were either reactive or proactive in nature. Reactive hints were given when the participant held their finger over a piece of paper on the table that said “Hint”. This was then observed by the Wizard, who gave a hint through the embodied agent. Proactive hints were given without the need for user action, but comprised the same kind of hints. A proactive hint would be given after the user had not initiated a reactive hint for more than 45 seconds. After this time, the system first asked whether it could be of any help to the user, and if nothing happened, it would give a hint from itself after a while. Hints that did not need a vision system were of the type “piece A and B border on each other” or “side A belongs on the inside of the cube”. Other hints that did require a vision system were of the type “piece A is not on the right position” or “piece A is in the right position, but is not turned in the right direction”. There was no difference in content between proactive and reactive hints.

The lab room where the test took place was equipped with four video-cameras and a microphone, which allowed the operator (the “Wizard of Oz”) to observe everything from a separate control room, and create a realistic atmosphere where the user had the idea that the system was fully operational and that he or she was being observed by a computer system (see also Figure 2).

In the setup described above, we tested with twelve participants who received the information and hints from the Wizard-controlled system. None of the users noticed that it was no real system they had been working with.

We also got some experience on the limitations of a Wizard of Oz approach.

The first limitation is that the illusion towards the user can easily be broken down by small mistakes from the side of the wizard. Small typing errors, slow response, no response, or the wrong response are examples of things that can go wrong when using a human operator to simulate a computer.

Secondly, the wizard has to know exactly what is going on. A good choice as a wizard would be someone from the design team itself, but it definitely has to be someone who knows every possible step the user might take next to be able to anticipate on these possible actions of the user to be able to simulate the system in such a way that users do not realize that they are in fact not interacting with an intelligent system.. If the wizard is not from the design team, extensive learning is required so that the wizard does not make mistakes in critical situations.

Thirdly, the amount of interaction that can be simulated by a wizard is fairly limited. The real-time work that has to be done “behind the scenes” can increase quickly at certain moments, overloading the wizard resulting in a loss of illusion for the user or a system that is responding very slowly or not at all to users’ actions. Especially when the user takes a few unexpected steps, this might ask too much from the wizard.

A final implication that comes from using a human to act like a computer is the fact that the human wizard will inherently react differently in similar situations with different test participants. These differences might be in the order of seconds or in the order of milliseconds, but it is a fact that it is physically not possible for a human to react at exactly the same speed at multiple occasions the way a computer would. This variance in responses might not only degrade the user experience of the participant, but also influences measurements that you want to perform yourself during the user testing. Every extra factor that is introduced into the test environment has an impact on the variance between participants and makes it more difficult to draw conclusions from the data.

5.Comparing Approaches

For an integrated study of user experience in intelligent environments for user assistance, a combination of approaches from various perspectives gives the best overview of possible interactions and ways in which users might be expected to use them once. Narrative interviews, scenario-based approaches and simulated technology combined can provide us with a common understanding of the user and user interactions with intelligent environments, in our case a cognitive vision environment.

Narrative interviews mostly teach us a lot about user experience with current technologies in a *real* setting filled with context information, which is insightful from a field in which we want to understand the user but only have just started developing the necessary tools to do so. This 'state of the art' view on user experience can also provide us with valuable information regarding future interactions with an intelligent system. However, such an extrapolation towards a new kind of system, towards a system that crosses the boundaries between the virtual and the physical worlds, towards a system that "invades user space" remains a step associated with uncertainty, for which other methods may be more applicable.

In parallel, a scenario-based development path and an iterative approach to this scenario, extending it as research progresses and using it to guide design, helps us focus on the ultimate goal of developing an intelligent environment that is made for users and not for technicians. An integrated scenario is not at all an exhaustive set of use cases, but it creates a common understanding of where the long-term research is heading. They keep us focused on interaction issues from the beginning of research and development and as such define which kind of user studies we actually want to perform.

Wizard of Oz is performed on a third level, (partly) filling the gap that cannot be filled by narrative interviews; allowing us to test the user assistance technology that does not actually exist yet. Simulating such technology at least gives us some possibilities for laboratory experiments and gives us some leads as to what shapes future user experiences. However, also this approach has its drawbacks. These drawbacks are mainly related to the use of a wizard to fulfill the role of a computer as described above; the wizard introduces more variance in the test setup and requires very extensive and precise knowledge of the system to perform as similar to system behavior as possible. Inherent to the wizard of Oz approach is the requirement to perform the test in a lab setting, where as many variables as possible are under the control of the researchers. This setting can be designed to remotely resemble the expected setting in which the intelligent environment is used, but it will always remain artificial and different. Compared to the context-rich narrative interviews lab experiments are relatively sterile and limited in their contextual setting.

Although we described three approaches to assess user experiences of an intelligent environment in an early development stage, each of them has its advantages and limitations and operates on a slightly different level. We believe that a combination of these three methods in a broader framework for user experience assessment would help us address relevant interaction topics, early, but not exhaustively. However, the last two methods mentioned, narrative interviews and wizard-of-Oz prototyping, can be used iteratively to constantly receive feedback on the way the development is heading straight from the user.

6.Conclusion

Cognitive vision shows to be a valuable improvement to the current interaction paradigm, but also a very new territory with lots of interesting research areas. By focusing research efforts in this domain (using scenarios) we can develop one "showcase" application for personal assistance on which we can base further development. Narrative interviews give us an integrated approach to user experience with general technology and uncovers relevant interaction issues. Wizard-of-Oz prototyping then gives us the possibility to iteratively test various would-be implementations in both exploratory and confirmative user testing.

Combined, these methods address those issues that are relevant for interaction design that help designers focus attention on the aspects that need work..

7.Future Work

Until recently, the cognitive vision prototypes had been restrained to fixed set ups in which one room could be overseen by the computer system. Recent developments have made this technology mobile, in the sense that a laptop can be equipped with the necessary tools to provide vision and recognition capabilities similar to the non moving cognitive vision system. In respect to this shift towards more mobile development, we want to continue to use Wizard of Oz style user evaluations of the mobile environment as it is being developed, in order to identify relevant user experience factors and to take into account the changes that occur when the system is transferred to a mobile medium and the implications that this has on user interactions.

8.ACKNOWLEDGMENTS

This work has been supported by the Austrian Science Foundation (FWF, project S9107-N04).

9.REFERENCES

[1] Brooks, R. (1991). Intelligence without representation. *Artificial Intelligence, 47*, 139-159.

[2] Cassimatis, N.; Trafton, G.; Schultz, A.; Bugajsk, M. & Adams, W (2002). A Task Domain for Combining and Evaluating Robotics and Cognitive Modeling Techniques. Measuring the Performance and Intelligence of Systems: *Proceedings of the 2002 PerMIS Workshop. NIST Special Publication 9902002*.

[3] Cole, R., Vuuren, S.v., Pellom, B., Hacioglu, K., Ma, J., Movellan, J., Schwartz, S., Wadestein, D., Ward, W. and Yan, J. (2003). Perceptive Animated Interfaces: First Steps Toward a New Paradigm for Human–Computer Interaction, *Proc. IEEE,* 91(9), 1391-1405.

[4] Dahlbäck, N., Jänsson, A. and Ahrenberg, L., 1993. Wizard of Oz Studies - Why and How, *Proc. ACM International Workshop on Intelligent User Interfaces '93, 193-200.*

[5] Geven, A., Schrammel, J. and Tscheligi, M. (2006). Personal assistance in the office: Interacting with embodied animated agents equipped with cognitive vision. *Nordichi 2006,* in press.

[6] ECVision, online at http://www.ecvision.org. Last accessed: January 2006.

[7] Schrammel, J.; Geven, A. and Tscheligi, M. (2006). Using Narration to Recall and Analyse User Experiences and Emotions Evoked by Today's Technology. *International Conference on Design and Emotion (In Press).*

[8] Schütze, F. (1976) „Zur soziologischen und linguistischen Analyse von Erzählungen" (Transl: On the sociological and linguistic analysis of narrations). *Internationales Jahrbuch für Wissens- und Religionssoziologie* 10, 7-41.

[9] Vernon, D., (2004). Cognitive Vision – The Development of a Discipline, *Proc. IST 2004 Event 'Participate in your future'*.

[10] Weidenhaupt, K.; Pohl, K.; Jarke, M. & Haumer, P. (1998). Scenarios in System Development: *Current Practice IEEE Softw.*, IEEE Computer Society Press, 15, 34-45.

A Sweet Dish for the Intelligent Coffee Table

Khairun N Fachry[1], Henk Eertink[2], Ingrid Mulder[2], Ruud Janssen[2]

Eindhoven University of Technology
P.O.Box 513
5600 MB The Netherlands
+31 40 247 5230
k.n.fachry@tm.tue.nl

Telematica Instituut
P.O. Box 589, 7500 AN The Netherlands
+31 53 485 0485
henk.eertink@telin.nl
ingrid.mulder@telin.nl
ruud.janssen@telin.nl

ABSTRACT
In this project, we focus on intelligent applications that can be deployed in a coffee corner in an office environment to make this environment more sociable, given the fact that the coffee corner is a place where professional and social interactions intermingle. We conducted a series of in-situ user studies in an office coffee corner that is equipped with context aware devices. Designing intelligent applications for the coffee corner requires a deep understanding of the office workers' needs and the situated interactions. First results of the early phase of the design process includes a list of user requirements and a conceptual design of a potential application are described.

Keywords
Design, Human Factors, Context-aware Systems.

1. INTRODUCTION
Intelligent devices such as wireless LAN, (indoor and outdoor) location sensing technology, and short range network functions are increasingly being used to support people's social life by means of context-aware systems. A system is context-aware if it uses the user's context to provide relevant information and/or services to the users, where relevancy depends on the user's task [4]. Many researches of context aware systems are from the technical point of view [7] e.g. on how to gather context from the users or how to increase the accuracy of the context aware technologies. This research, however, focuses on the user perspective on how to facilitate social interaction mediated by the context aware system which is in keeping with [11].

Previous researches show that many context aware technologies are designed for outside of the office environment, e.g. [3, 7, 8]. This research deals with how to support social interaction in a shared area of an office environment. When developing such a technology in an office environment, we are challenged with the question of how to combine professional and social interactions in a system. On one hand, the system should maintain the professional atmosphere in the office by taking into account situated interaction and privacy. On the other hand, the system should support sociality among colleagues.

In this paper we present a system that supports semi-formal activities of the office workers while they are in an intelligent shared office area in general and in an intelligent coffee corner in particular. The coffee corner is used by the office workers during breaks to meet each other. It is usually occupied by the office workers who share the same floor which consequently triggers social interaction among the office workers. Along with the emerging use of intelligent technologies in office environments, the coffee corner in focus is equipped with intelligent devices such as device-discovery, context-aware services, sensors, Bluetooth, RFID reader, camera, speaker and advanced displays (see Figure 1). Using these intelligent devices, the current coffee corner system allows the office workers to locate their colleagues within the office.

Inspired by the intelligent coffee corner available "in the house", we are particularly interested in the following research questions: how is the coffee corner being used now and how to enhance sociality among the office workers mediated by context aware systems.

2. METHOD
Given the large design space of intelligent applications, we opted for a system that is driven by, designed by, and is useful for the users. As a system driven by the users, we chose to collect user requirements through a series of in-situ user studies.

2.1 Direct Observation

2.1.1 Method
As an exploratory study, we observed in one of the coffee corners and investigated how the office workers behave in the intelligent coffee corner. The direct observation was open-ended and unstructured. This means that we observed the full situation of everything that is going on [9] in the coffee corner. In parallel, a literature study was conducted on the topic of intelligent shared areas.

2.1.2 Result
The observation and the literature study resulted in finding that a shared physical environment promotes informal social communication [12]. Although the shared area is located in a professional environment, the office workers tend to perform informal and semi-formal communication while they are in the coffee corner. Aside of the company project topics, topics such as upcoming events, weekends, sports, weather are several out of many topics that the office workers discussed about during their breaks.

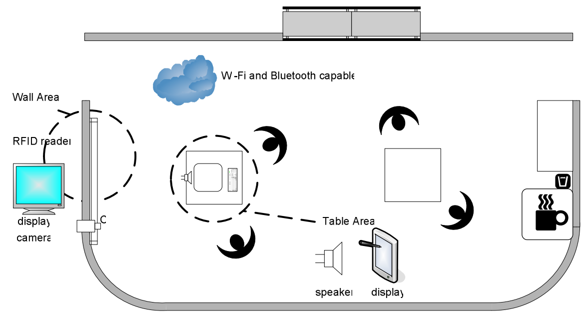

Figure 1 Observed Coffee Corner

Furthermore, we noticed that the behaviors of the office workers while they are in the coffee corner vary from person to person. Therefore, it is important to know the type and characteristics of the office workers. Based from our observation, we distinguished three groups of users[1]:

1. Frequent users.
   They have coffee regularly in the observed coffee corner and perform social interaction with each other. Frequently they invite their colleagues to drink coffee together. They often refer each other as “coffee buddy”.
2. “Looking familiar” users.
   They have coffee occasionally in the observed coffee corner. Their visits are incidental. They are employees of the office but they neither frequently work together with the frequent users nor share the same floor with the frequent users. They could be project partners from other companies or universities.
3. Guest users, visitors of the employees.
   They are visitors of the employees, thus they are seldom in the coffee corner. In most cases they are always accompanied by the visited employee due to their unfamiliarity with the area.

It was observed that in order to share information, the office workers brought along printed document, printed pictures or parts of newspaper to the coffee corner. Some of this hard-copy document were then collected back, left on the coffee table or hung on the wall. Many office workers brought along their personal devices such as hand phones or PDA's so that they can show their information i.e. pictures through their personal devices. Occasionally, they went to the coffee corner displays to look for some information on the internet and shared the information with their colleagues.

As for the interaction with the system, the behavior was different for each group of users. The frequent users were familiar with the area and interacted frequently with the intelligent devices. Their familiarity with the system caused them a tendency to occasionally try out the system if there is any new feature has been added to the system. The looking familiar users needed to be informed and encouraged to interact with the system. They tended to interact with the system after the frequent users has interacted with it. The guest users had to be invited, explained and offered to interact with the system.

[1] In this paper, the terms office workers and users are use interchangeably

The current input devices for interacting with the system were a keyboard and a mouse. These input devices force the office workers to take turn in interacting with the system and pass around the input devices. In many occasions, one frequent user was in charge of the input device and the other office workers watched or gave instructions on what interaction they should do. This fact hinders the social interaction in the social setting of the coffee corner.

Due to the nature of the coffee corner, the system is only used during breaks. Since the office workers have a limited time to spend during breaks, they appreciate a low effort, fun interaction and immediate feedbacks from the system. For example, a smart display on the table which detected where the user was standing (by placing sensors on the carpet) and changed its orientation based on the user’s standing location was highly appreciated and played by the office workers.

2.2 Collecting Ideas from the Users

2.2.1 Method

One finding from the observation is that the users discussed many ideas of the application that they would like to have while they are in the intelligent coffee corner. The intelligent technologies available around them might have stimulated this idea generation process. However, those ideas disappeared since nobody was there all the time to collect them. Therefore, we created an idea “submitter” and displayed it on the coffee corner display[2]. The idea submitter copies the function of an idea box (a form where the office workers can submit their ideas to us). Other considerations displayed via the idea submitter on the coffee corner display were to socialize the existence of the shared display and to entice the office workers to interact with the display. Furthermore, the idea submitter was available online and was running for two weeks. By collecting the idea online we gave the office workers the opportunity to submit their ideas from their personal PCs as well. Finally, in order to protect privacy and to promote creativity, the office workers could choose to submit their ideas anonymously.

2.2.2 Result

Many received ideas are examples of applications that the office workers would like to have and technologies that we can use in developing those applications. In total 23 different ideas of applications were received. We analyzed the proposed ideas by organizing and grouping them based on similarity of interests, resulting in five categories of potential applications.

1. **Company Awareness**
   The office workers would like to have an application that makes them aware of the events that are happening within the company. The events could be about upcoming presentations, new employees, etc.
2. **Personalized Information**
   The office workers would like to share their selected personal information with their colleagues. The selected information could be the users' interesting links, pictures, videos, etc.

[2] The online idea capturer is available on http://wwwhome.cs.utwente.nl/~meppelin/Nisa/start.html

3. **Games**
   The office workers would like to have short and interactive games during breaks to play with their colleagues. Competitions among office workers were also favored by providing information about the progress of the game or who wins the game. The games could be polling games, puzzles, quizzes, etc.
4. **Idea Publisher**
   The office workers would like to have a system that supports exchanging ideas among the office workers. The system should allow the office workers to submit, publish and comment ideas. The ideas could be about project, team building, social activities, etc.
5. **Who is where and what**
   The user would also like to know where their colleagues are and their current activities. They would like to know where their "coffee buddies" are and be able to contact them from the coffee corner.

One finding was that all of the people who submitted their ideas together with their names were the frequent users of the observed coffee corner. We argue that it was perhaps due to the familiarity with the environment, and were not feeling timid to submit their ideas onymously.

2.3 Coffee Break Brainstorming

2.3.1 Method

Since the office workers are in the coffee corner during their breaks and the intelligent devices are installed in the coffee corner, we chose the coffee corner as the place to brainstorm with the office workers. An invitation email was sent to all of the office workers to join the brainstorming sessions. A preset script of questions was used as a framework to keep the discussion running. This gave us the opportunity to get a deeper analysis about the potential applications, to know the concerns regarding the applications and to discuss possible solutions to overcome those concerns.

We visualized all five possible applications in short scenarios. We chose to visualize the scenario with animation because we think the animations can capture the office workers attention, thus trigger discussions among the office workers[3].

The visualization was uploaded on the coffee corner display. A moderator was present in the coffee corner and was in charge to lead the discussion. The comments of the office workers were noted on post-it notes by two note keepers. Since we have five categorizations of ideas, we presented one idea each morning. When there was a crowd of office workers in the coffee corner, the moderator went there and collected more users' opinions. Post-it notes, pens and a flip board were placed next to the display so the office workers can see each other comments.

2.3.2 Result

In general, every brainstorming session took 30-45 minutes and was attended by 10-14 office workers. A considerable amount of qualitative feedback from the office workers was collected. While

Figure 2 Brainstorming Sessions

the idea submitter merely answered the kind of interactions that can be occurred in the coffee corner, the brainstorming sessions allowed us to get more insight into the submitted ideas.

To clarify the approach, a part of the result of the brainstorming is presented below. The texts in brackets, italic and right aligned are the comments from the office workers. The selected comments were representative of the brainstorming session. Our interpretations of the comments are presented right after the comments. The complete list of comments and their interpretations for each scenario collected is described in [5].

"I would like to know what is going on today"

> The scenario was highly appreciated by the users. They would like to know what is currently happening in the office. They provide more examples of information that they would like to know, i.e. information of the colleagues who are being ill today, who are on business trips, etc. This finding concludes that the content of the information should be semi-formal information and the information should be updated timely (requirement no.1).

"No typing for keyboard, but for example scanning your badge for information"

> The office workers were not in favor of having keyboard or mouse as their interaction input devices. They claimed that keyboard and mouse only allow one person access. Moreover, keyboard and mouse give them the feeling of being behind their computer again during their break time (requirement 4 and 5).

We adopted affinity diagrams [6] to organize and analyze the data from the brainstorming sessions resulting in a list of user requirements for the design of applications for the intelligent office coffee corner. A list of requirements that is generic for all scenarios is presented.

1. The content should be semi-formal information e.g. upcoming presentation, new employees, etc.
2. The system should support personalization by providing a possibility to present the office workers' selected personal information.
3. The system should entice curiosity of the user by providing them daily updated information and an interesting interaction input device.
4. The system should support access from multiple users so that the user does not need to take turn in interacting with the system.

[3] The animated scenarios are available on http://wwwhome.cs.utwente.nl/~meppelin/Nisa/scenario.html

5. The interaction should be easy and fun to use and should not remind the office workers of their personal PCs.

2.4 Experts Design Session

An experts' design session was organized to collaboratively select and design scenarios that will be implemented for the intelligent coffee corner. Six experts coming from various expertise groups were invited. The experts are the office workers of the office in focus, thus they are also the potential users of the system.

The session started by reflecting the results from the brainstorming sessions. A moderator was present to give the overview of the result of the brainstorming sessions i.e. to present the user requirements and the pros and cons of each scenario. The experts were asked to discuss the best scenarios that will be implemented in the observed coffee corner by considering the result of the brainstorming sessions.

The session resulted in a decision to implement scenario 1 and scenario 2 because of several reasons.

a. They appeared to be the most favorite scenarios by the users. It is shown by the number of positive feedbacks (written on the post-its notes) that we collected and the frequency of the scenarios was raised back by the users during the brainstorming sessions.
b. They support social interaction among the office workers. They are a giving and receiving system, which provides exchanging of semi-formal information among the office workers.
c. Finally, they give the possibility to entice the office workers to interact with the system by giving them "unexpected" feedbacks. From on the study of Agamanolis [1], half the battle in designing an interactive situated or public display is designing how the display will invite that interaction. According to Churchill *et al.* [2] users needed constant encouragement and demonstration to interact with the interactive public display. The systems will shows daily information based on the received information from the office workers. Thus, by showing updated information, we aim to entice curiosity of the office workers so that they will interact with the system.

Based on this result, new scenarios refining scenarios 1 and 2 are designed.

About company awareness

As usual, as Dan arrives at his office, he likes to start his day with a cup of coffee. He goes to the coffee corner and he meets his colleagues there. Dan and his colleagues wonder what is going on in the company today. They approach the coffee comer display and ask the coffee corner system to present the events of the day. The coffee corner display presents them the events of the day including a birthday of Peter and a presentation of Karin.

About personalized information

Dan is sitting in front of his personal PC. He is browsing on the internet and he founds interesting information about a cycling trip. He thinks this information could be interesting to share with his colleagues. He published the information to his personal space in Coffee Corner System.

On Friday afternoon, Dan is in the coffee corner with his other colleagues. They were talking about their plan for the weekend. Dan remembers about the interesting link that he found. Dan gives his consent to display his personal space on the display. The display presents his personal space and Dan suggests his colleagues to join him for the cycling trip.

3. CONCEPTUAL DESIGN

The basic concept of the system is that it broadcasts general (requirement 1) and personalized (requirement 2) information. General information supporting company awareness can be published and accessed by every office worker. In this way, the system acts as an interactive broadcasting system. The personalized information acts a personal space where the office workers can publish their selected information to the coffee corner. It can only be displayed on the coffee corner display if the owner has given his consent for publishing. In this way, the system acts as a personal public display that provides an easy way to share personalized information to the office workers present in the coffee corner.

3.1 Technology Alternatives

As mentioned in the observation result, the current input devices force the user to perform turn taking interaction with the system. Therefore, on the technology choice we focus on how to make the interaction access in the coffee corner more "social" by allowing multiple people to access the system without needing to take turn (requirement 4).

We considered several technologies as input devices i.e. speech, RFID and Bluetooth technologies (see Figure 3). These technology alternatives were sketched on paper and being discussed with two experts who work in the office in focus and know very well the constraint of the coffee corner area. They are being compared with each other by discussing the benefit and limitation of each technology.

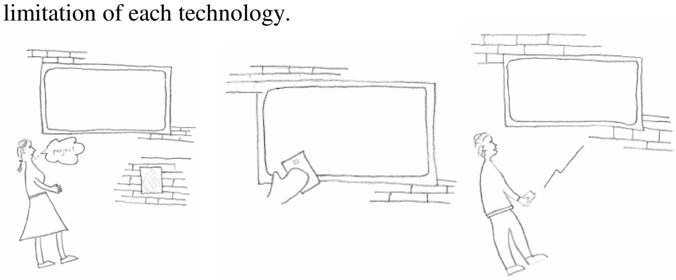

Figure 3 Technology Alternatives

The speech interface was not chosen because the speech interface is unreliable in a noisy environment (with multiple audio inputs). We considered using Bluetooth technology from the users' mobile phone to interact with the system. However, we think that this will make the coffee corner less social because every office worker has their own input device so that they can access the shared display without needing to negotiate with each other.

RFID tags, which are covered in interesting forms are chosen (named as interaction tags for the rest of the paper). The interaction tags are spread on the table so that the user can easily choose and pick up the tags, thus involving multiple users in deciding what information should be displayed on the coffee corner. It does not require the input devices (like keyboard and mouse) to be passed around. This input device makes all of the office workers to have the same right in accessing the system. The user can choose the interaction and even disrupt each other in interacting with the system. By this way, we copy the way people make decision in regular conversation by giving the office worker chances to negotiate and disrupt each other while interacting with the system. It facilitates a social way for interaction allowing a group of people interaction with the system simultaneously, especially in the given social setting.

Scenario company awareness does not require consent from the users. However, scenario personalized information requires the system (1) to recognize the user who attempts to use the system and (2) to get consent from the receiver that his or her personalized information will be displayed on the shared display.

As a next step, a technology that can recognize the user and facilitates an explicit consent from the user needs to be chosen. The office workers badges have been chosen for identification to get access to their personalized information. The main reasons are the office workers can use their own badge to initiate the interaction. Since every office worker is required to wear his or badge everywhere in the office, we do not need to force them to carry along a new device for personalization. More importantly, it is an unobtrusive and accurate way for recognizing the office worker and giving consent to the system to display the user's personalized information.

3.2 Interaction Tag Design (Cards vs. Sweets)

Other than functional aspects, we are also interested in aesthetic ways to entice curiosity of the users (requirement 3). We design the interaction tags so that they are easy and fun to use and do not remind the office workers of their personal PCs (requirement 5).

1. Form factor of colorful cards.

A collection of available cards is spread on the table (A) and the reader is embedded on the table (B). This form factor provides easy interaction from the users. It only requires the user to select the card that corresponds to the information that user wants to display (C) and drag the card without needing to lift up the user's hand.

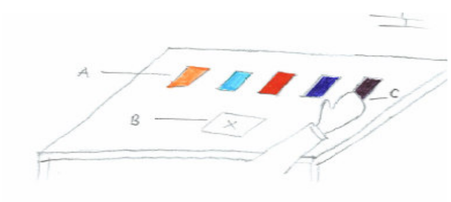

Figure 4 Form Factor Cards

B. Form factor of candies. A collection of available candies is placed in a box (A) and the reader is embedded on a plate (B). The metaphors of candies and the plate match the environment of the coffee corner. Instead of dragging (as in form factor cards), this form factor requires the user to take the candy out of the candy box and move it to the plate (C).

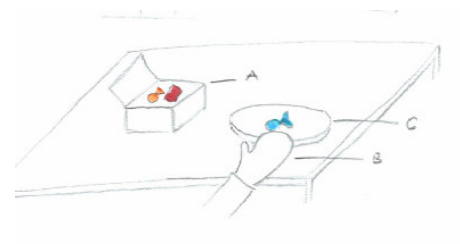

Figure 5 Form Factor Sweets

Both form factors are designed in these forms in order to make the input device playful and interesting for the office workers. More importantly, the forms will raise curiosity of the office worker, especially for the looking familiar and guest users.

Having considered the coffee corner environment, we opted for form factor of candies. The office workers bring along their badges while they are in the coffee corner and occasionally put their badge on the coffee table. By having the reader on the plate, the reader is obvious to the office workers, thus prevent the system to present personalized information by mistake.

3.3 Information Design

The home screen presents two categories i.e. general information supporting company awareness and personalized information. To help the use of the menu, navigational cues for possible presentation are presented on the display (see Figure 6). A red candy will present the general information and the badge presents personalized information. The users choose the category that they want to present by placing either the red candy or the user's badge on the plate. The display then presents the requested information, updates its screen and presents the next possible presentation. The same concept is valid for the next presentation. The user needs to see the display and then selects the interaction tag corresponds to the information that the user wants to display.

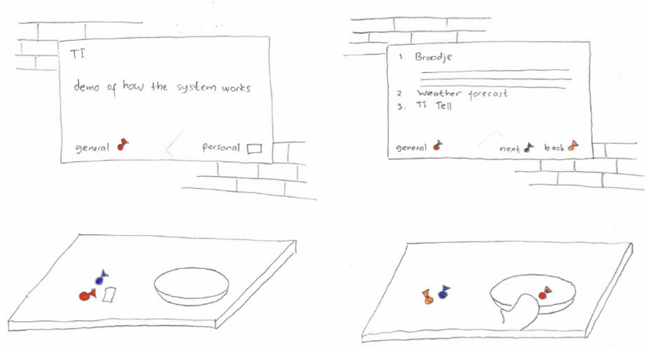

Figure 6 Navigational Cues on the Display

4. IMPLEMENTATION

The prototype has been implemented as a web-based application running on the coffee corner display. The information presented on the coffee corner display is facilitated by two databases. The first database is the information database that contains data that will be presented on the coffee corner display. For this purpose, a web interface has been implemented in order to push information to the information database. The second database is the context database that contains the history of RFID identification from the information retriever. The RFID identification is provided by the Context Management Framework (CMF) [10]. This is a distributed system that aggregates context-information about users. In this situation, it is used to control the interaction between the user and the RFID-readers on the coffee table, whereas the web-based application actually executes on the Coffee Corner Display. It provides the attributes location, time, and RFID tag number to the database.

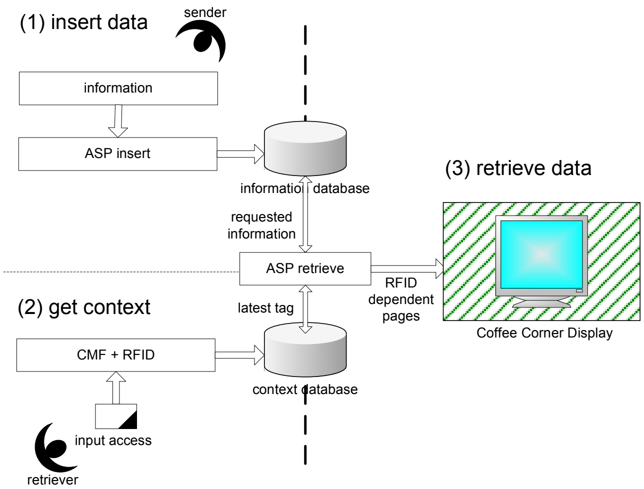

Figure 7 System Architecture

The dynamic retriever-side presentation on the coffee corner display (the 'ASP retrieve' function) queries the latest RFID tag from the context database. Using the interpretation of the latest tag, the system selects information from the information database and pushes the information to the coffee corner display.

5. FURTHER WORK

The functional prototype is scheduled for user evaluations at the coffee corner in focus in August 2006. The focus of the evaluation is the sociability of the system. The prototype will be refined on the basis of the finding from the evaluation.

6. ACKNOWLEDGMENTS

The authors thank to Hans Zandbelt and Martin Wibbels for their contribution in the prototype implementation and Lianne Meppelink for her contribution in setting up the user studies.

Biography of Authors

Khairun N. Fachry is a post-graduate student of User-System Interaction program at the Eindhoven University of Technology. Currently, she is conducting her final project in the Telematica Instituut in the scope of interactions in the intelligent coffee corner mediated by context aware systems. She is M.Sc in Telematics from Twente University (in The Netherlands).

Henk Eertink is a senior researcher at Telematica Instituut, and leads the Intelligent Communications group, that focuses on mobility management, context awareness and the related trust and privacy issues. Henk holds a PhD in Computer Science, from Twente University.

Ingrid Mulder holds a Ph.D. in Behaviour Sciences (2004). Since 1998 she is doing human-centric innovation research at the Telematica Instituut. She has been developing methods and techniques for capturing innovation, experiences, and latent user needs, as well as methods for the design and evaluation of innovative technologies.

Ruud Janssen studies knowledge sharing and information exchange between information workers in offices using a mixture of ethnographical research methods and user centered design. He holds a PhD degree (1999) in visual perception.

7. REFERENCES

[1] Agamanolis, S. Designing Displays for Human Connectedness. In O'Hara, K., Perry, M., Churchill, E. and Russell, D. (eds), Public and situated displays: Social and interactional aspects of shared display technologies, Kluwer, 2003.

[2] Churchill, E.F., Nelson, L., Denoue, L., Helfman, J. and Murphy, P. Interactive systems in public places: Sharing multimedia content with interactive public displays: a case study. In Proceedings of the 2004 Conference on Designing Interactive Systems (DIS2004) (Cambridge, MA, Aug. 2004), 7–16.

[3] CNN.com. Japan's Lonely Hearts Find Each Other with "lovegety". CNN.com, 7 June 1998, http://www.cnn.com/WORLD/asiapcf/9806/07/fringe/japan.lovegety/ (accessed 16 June 2006).

[4] Dey, A.K. Understanding and using context. Personal and Ubiquitous Computing 5: 20-20, 2001.

[5] Fachry, K.N., Post-graduate Final Report on "interactions in the intelligent coffee corner mediated by context aware systems"., User-System Interaction, Eindhoven University of Technology, The Netherlands, to appear in October 2006.

[6] Hackos, J.T., Redish, J.C. User and Task Analysis for Interface Design. John Wiley & Sons: USA, 1998.

[7] Kaasinen, E. User needs for location-aware mobile services. Personal Ubiquitous Computing 7: 70-70, 2003.

[8] Patrik Floreen, et al., "Towards a Context Management Framework for MobiLife," Proc. 14th IST Mobile & Wireless Summit 2005, Dresden, Germany, June 2005.

[9] Taylor-Powell, E., & Steele, S. Collecting evaluation data: Direct observation (G-3658-5). Madison: University of Wisconsin-Extension, Cooperative Extension, 1996, http://cecommerce.uwex.edu/pdfs/G3658_5.PDF (accessed 20 June 2006).

[10] van Kranenburg H. and Eertink H. "Processing Heterogeneous Context Sources", Proc. Next-Generation IP-based Service Platforms for Future Mobile Systems Workshop 2005, Trento, Italy, Jan. 2005, pp. 140-143.

[11] van Kranenburg, H., Koolwaaij, J., Mulder, I., & Eertink, H. Context Awareness Targeting User Needs. In Proceedings of Workshop on Context Awareness for Proactive Systems, June 12-13, 2006, Kassel, Germany, pp. 5-17.

[12] Whittaker, S., Frohlich, D., and Daly-Jones,O. (1994). Informal communication: What is it like and how might we support it? Proceedings of CHI 1994.

A Development Cycle for Ambient System Design

Tim Schwartz
Saarland University
66123 Saarbrücken
Saarbrücken, Germany
schwartz@cs.uni-sb.de

Ralf Jung
Saarland University
66123 Saarbrücken
Saarbrücken, Germany
rjung@cs.uni-sb.de

ABSTRACT
This paper describes a design development cycle for ambient systems. Other than the traditional software development concepts, we concentrate on the challenges of ambient designs. First we introduce the Ambient Development Cycle which consists of four designing stages. By trying to find a formula to rate the "level of ambience", we identify a few factors that seem to be important for ambient systems. We think that having such a formula would be of great help for developing, designing and comparing different concepts of systems.

1. INTRODUCTION
The increasing demand for ubiquitous computing and disappearing computational power embedded in our everyday environment created the idea of ambient intelligence. On one hand such systems should work in an intelligent way to support persons in their everyday environment by providing different services. On the other side such a system that consists of software and hardware must be integrated seamlessly in the environment. In this paper we will focus our attention on the *ambient* part of ambient intelligence, especially the requirements for artifacts that have to be integrated in a room to get an instrumented environment. That also includes the ambience of the uninstrumented room which represents the starting condition and at the same time the presettings for design choices.

In our notion we understand the paradigm of ambient intelligence, especially the ambient part, as the disappearance or seamless integration of smart objects in an environment to provide unobtrusive user support [3]. We will start with the assumption that a given environment has a natural level of ambience in terms of "atmosphere" as a result of existing furniture and objects. It is hard to evaluate that level because of the individual and subjective understanding of esthetics. This is the reason why we should try to hold up the original room style, adapt the hardware to it and compare the instrumented room with the original.

The question is how can we build an ambient system starting with the goal that describes what services the system should provide, followed by the requirement analysis, the design process, the modification step and finally the integration in the environment. The complex process of developing software products asks especially for efficient design models where we should also take economical constraints into consideration. In this paper we introduce an assistance-concept for developing ambient systems. We take a look at the economical interplay between technical hardware and the environment in which the hardware should be seamlessly integrated. The question why it could be interesting to have a description for the "level of ambience" will also be introduced in this paper and hopefully builds the starting point for a constructive discussion.

2. AMBIENT DEVELOPMENT CYCLE
The well-known *waterfall model* [9] introduced by Winston W. Royce in 1970 was one of the first practical models for software development. Despite of many justified criticism and modifications the core idea is still comprehensible and feasible for smaller developments . Improved design strategies came up, e.g. the *Prototyping Model* or the iterative and incremental approach of the *Unified Software Development Process (USDP)* [6]. Another famous and vivid model is Barry Boehm's *Spiral Model* [2] which takes changing requests into account (one of the main problems of the waterfall model). The main idea of the spiral model is the refining of the four development steps namely analysis, concept, coding and testing by repeating these steps until the product is satisfying. It provides the basis of our *Ambient Development Cycle* (Figure 1) that will be introduced in the following in which we turn our attention to an extension for ambient systems.

2.1 Phase I: Analysis
The goal is in the center of the development cycle which means we have to make sure that in all four phases we check whether the actual step fulfills the goal or is part of the goal. A typical goal could be the unobtrusive augmentation of a room with smart objects to provide intelligent human-supporting services.

In the beginning of the analysis we divide the the main goal into three analysis components in which we define the requirements in the fields of *human-factors*, *environment* and technology (*hardware and software*). The main focus in technical ambient systems lies on the unobtrusive and ubiquitous assistance of human beings with computational power integrated in their everyday environment. On the one

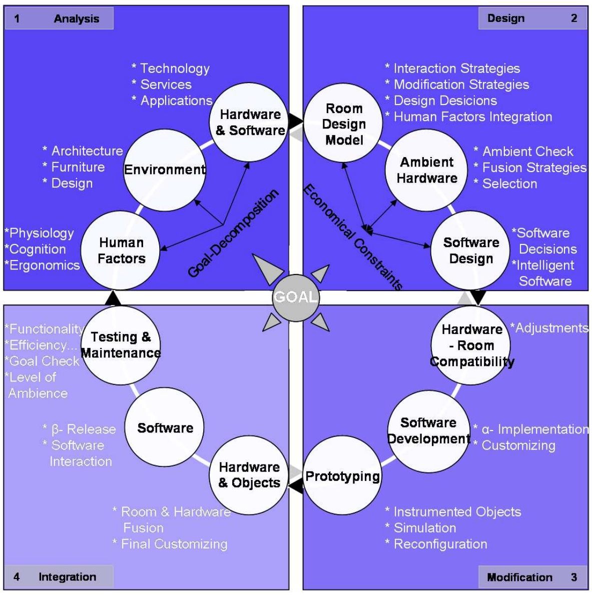

Figure 1: The Ambient Development Cycle

hand, the efficient integration of hardware in everyday objects requires the knowledge how persons interact with their environment and what services are useful to support[7]. On the other hand we must have a closer look on how to enrich everyday objects with hardware with respect to aesthetical aspects to increase the acceptance of such instrumented objects and to make a system ambient.
In the analysis phase we not only determine what is needed but also what is already available.

2.2 Phase II: Design

The design-step includes the creation of an adequate *Room Design Model* based on the analysis of the environment and the human-factors. An important question is how the environment (e.g. a room) can be modified without violating the original "design style". That is also the claim for the hardware that ought to be integrated in the room. Primarily, such an *ambient hardware* has to meet the functionality of the goal but it should also take room design constraints into account. It would be helpful if we could rate hardware by their "level of ambience", that should describe how easy hardware could be integrated in an environment with respect to ambient attributes e.g. invisibility and unobtrusiveness. We will have a closer look on this topic in Section 3
Another important aspect that also influences design decisions are *economical constraints*. Whatever the motivation for developing a system is, you are liable to financial conditions. For ambient systems the selection of appropriate hardware that is easy to integrate in the environment can decrease the costs for room modifications. On the other hand, such hardware may be more expensive than traditional hardware. As mentioned above the design style of the room should be changed as little as possible, hence the room modification possibilities are limited. In the *software design* step, it is important to select "intelligent algorithms" for supporting and adapting the user and his/her behavior. If there is an existing software that meets the requirements under certain aspects, you should ask yourself the question whether to modify the existing software or to build up a new system. The use of an user modeling service (e.g. UbisWorld [5]) should also be taken into consideration to make the services more personalized in the sense of the human-centered approach.

2.3 Phase III: Modification

The chosen design strategies for room model, hardware and software of the previous phase will be realized in the modification phase. First, there is a hardware-room compatibility test to find out whether the selected hardware and the room can be brought together to an instrumented ambient room. Slight modifications can be done in this step without going back to Phase II to redesign the room and the hardware. The *software development* process based on the software design decision of Phase II with a new implementation or the customizing- and modification-process of the existing software package. The fusion of room objects and the selected hardware takes place in the *prototyping* process where a mock-up will be realized. The prototype is also important for first interaction tests and simulation testing.

2.4 Phase IV: Integration

After prototyping, the system can be implemented in the real-world environment by integrating the instrumented objects in the room. The developed software should be free of programming bugs to make sure that no unwanted effects occur. The final step in the development circle for ambient system is the *testing and maintenance* part. Especially the testing phase should include important usability test to guarantee a human-centered system with its characteristic features. Furthermore the "level of ambience" after the modifications and integrations should be compared to the beginning level and the predicted level of the design phase. With these experiences we can learn for future developments how the ambience of an environment can be rated.

3. HOW AMBIENT IS MY SYSTEM?

In the introduction we stated that we want to put our attention on the ambience of a system. Imagine a room that is stuffed with electronic devices which offer a lot of nice services to the user but are bulky, ugly, connected by wires and have several lights that somehow show their activity coded in cryptic blinking which can only be understood by experts. You wouldn't want to have this as your living room (unless you are a complete geek), would you? Clearly this room has a low ambience or, as we call it in the following, it has a very low ambience level.

It would be nice to have a way to determine the ambience level of a system in order to compare two designs and to successfully complete the development cycle. A natural point for this would be in the integration part of the development cycle. Here, user studies can be used to test the β-release (or any other release) for its ambience level. Of course at this point of time lots of design decisions have already been made and thus modifications of the room or the hardware can be quite costly. A better point would be in the design part of the cycle. It would be nice to have some sort of function, formula or other technique that could be used to compare the ambience level of two concepts.

In this section we want to discuss various aspects (or variables) that such a formula should include and hope that this will lead to an interesting discussion.

3.1 A First Approach

It is clear that the ambience of a system is not a binary function (in other words it is either ambient or not) but is rather a gradual one (e.g. concept A is more ambient than concept B and less ambient than concept C). As stated in the introduction, the term ambience combines several aspects and some of them (e.g. aesthetics, atmosphere) are subjective or depend on design principles of a time era. One general goal of an ambient intelligent environment is to provide additional functionality (with the help of computing devices) where the devices themselves and their interfaces should be as unobtrusive as possible. To reduce the impact of the subjective aspects of ambience, one could compare the (electronically) uninstrumented room, that may have been carefully designed by some interior designer, with the instrumented room. Here the assumption is that the uninstrumented room is crafted in a fashion that takes the subjective aspects into account (in other words: we are shifting the task of dealing with these aspects to somebody who knows what he is doing). It is clear that an invisible modification of an environment does not change the ambience level of that environment. So at a first glance the change of ambience level seems to depend on the number of visibly added modifications. To extend this idea to two concepts C and C', that both realize the same system and are based on the same uninstrumented environment or where C' is a modification of C, one could just simply subtract the number of visible devices of each concept:

$$\Delta Ambience = \#VisbleDevices_C - \#VisbleDevices_{C'} \quad (1)$$

Here $\#VisibleDevices_x$ is the number of visible devices in a concept x and C and C' are two concepts that should be compared. If $\Delta Ambience$ gets negative, concept C is better than C'. Obviously $\Delta Ambience$ gets zero if both concepts have the same amount of visible devices or if all devices are hidden. Getting a positive $\Delta Ambience$ means that C' is better, in terms of ambience, than C. This seems to roughly fit the intuition but Equation 1 will change regardlessly of the form factor of the installed devices.

3.2 Evaluating the Ambience of Devices?

Equation 1 has at least one major backdraw: It drops with every visible device that is added to C' regardless of the device's form factor (or other visible attributes).One idea to overcome this would be to introduce a weighting factor for each device that expresses its level of ambience or better its level on non-ambience (e.q. some kind of ambience penalty). $\Delta Ambience$ can then be expressed as the difference between the sum of the Ambience Penalties AP of all devices in concept C and the sum of all APs of the devices in concept C':

$$\Delta Ambience = \sum_C AP_C - \sum_{C'} AP_{C'} \quad (2)$$

The idea behind this is, that non-visible devices get a penalty of zero ($AP(\textit{non-visible device}) = 0$). All other devices get a penalty greater than zero, depending on their level of non-ambience (the higher the penalty the worse the ambience).

3.2.1 Partial Hiding of Devices

From the section above one could conclude that hiding electronic devices is always a good idea but of course this is not always possible. In [4][10] we describe an indoor positioning system that works with active RFID tags and infrared beacons that are installed in the environment. Since RFID tags are based on radio technology they can easily be hidden behind the wall or in the ceiling and can still provide their functionality. Infrared beacons however work with infrared light, which means that they cannot be completely hidden or otherwise they will loose their functionality.

Another example for devices that cannot be completely hidden are displays that are used to inform or entertain the user, therefore they have to have some AP greater than zero. Since there is a vast variety of displays with different form-factors one idea could be to take the volume of a device as a rough measure for its AP. An old fashioned, bulky CRT monitor would get a high penalty (high AP) whereas some slim LCD screen with the same screen size gets a lower penalty, because it takes up less space in the environment. An ever lower AP can be achieved by mounting a projector under the ceiling. Using back projection should result in a zero AP for the projector itself because it is hidden behind the wall and therefore invisible, but intuition tells us that it should get some AP because of the space that is used up by the projection itself.

This shows that the ambience level should not only depend on the device itself but also on how or where it is installed. It is also possible to break a hole into the wall and install the CRT screen in that hole. This should at least be equivalent to the LCD screen that is mounted on the wall. As a conclusion we could say that the ambience penalty should correlate with the visible volume of the device after it is installed in the environment. Furthermore, the ceiling seems to be a better place to mount devices than the wall or some other place that is in the normal visual field of the user. The equation should therefore be extended with some kind of Mounting Factor MF:

$$\Delta Ambience = \sum_C MF_C \cdot AP_C - \sum_{C'} MF_{C'} \cdot AP_{C'} \quad (3)$$

Here $0 \leq MF \leq 1$, which means that devices that are completely hidden get a mounting factor of zero and devices that are completely in the visual field of the user get a mounting factor of one. Devices at the ceiling can get an MP between zero and one.

3.3 To be Ambient or not to be Ambient

Before we proceed with some more factors that should be considered when evaluating the ambience of a system, we want to investigate an example that shows that partial hiding of a device and maximizing the ambience level is not always favorable:

Consider an example where you have to instrument a museum that exhibits the workroom of Ludwig van Beethoven in such a way that some lifelike character guides the visitors through a room (see [8] for an example of such a system). Apart from a PDA (Personal Digital Device) this system uses public displays to show movies, animations and the lifelike character itself. Of course in the times of Beethoven

Figure 2: Loudspeaker camouflaged as a rock and as a picture[1].

Figure 3: A sphere shaped TV from the 60's.

no such technology as television existed and so there is no way to integrate the hardware into the workroom without violating historical facts. In that case it is better to expose the hardware as something that clearly does not fit in the environment. So the desired ambience level seems to also depend on the goal.

3.4 More Design Factors

3.4.1 Camouflage

Instead of complete or partial hiding of devices there is also the possibility of camouflaging the hardware. Figure 2 shows a loudspeaker in the design of a rock and one designed as a picture. Clearly, if the design of a device fits into the environment the MP factors need to be adjusted accordingly.

3.4.2 Historical Era

In Section 3.3 we described a very restrictive museum scenario which forced us to be historically correct. However, there are similar scenarios in which the restrictions are not that hard. A room, for example, that is designed in the style of the late 60's or early 70's just because the owner or original designer digged that style. That particular era's design style is very futuristic, which can be a great advantage for an easy integration of electronic hardware. Think about 60's hifi systems, that had lots of knobs, analog meters and blinking lights or think about sphere shaped television sets (see Figure 3) that should remind the observer of an astronaut helmet. These design elements seem to make it even easier to seamlessly integrate high tech devices in the environment and it may even be the case that they *add* to the ambience of the room (in terms of atmosphere).

Clearly the historical era influences the design (the look) of electronic hardware which also means that a room that has a specific ambience level today may have another ambience level at some point in the future. In situations where historical correctness is inevitable some design elements or devices simply cannot be used. These considerations need also be taken into account in the formula, maybe as another weighting factor similar to the mounting factor.

3.5 Economical Constraints

In Section 3.2.1 we took public displays as an example and we concluded that installing a CRT screen into the wall should have the same ambience rating than a flat LCD screen. Intuitively there should be some difference between both approaches: Breaking a hole into the wall is of course a financial factor and so is CRT or LCD screen. It may be that an LCD screen of a certain size is much more expensive than a similar CRT monitor, so that breaking a hole in the wall plus the expenses for the CRT device is cheaper than the flat screen. To take these things into consideration we would need a new measure that expresses the "economical ambience":

$$economical\ ambience = \frac{LevelOfAmbience}{Costs} \tag{4}$$

4. CONCLUSION

As we have seen in this paper the development process for ambient systems differs from the classical software design cycle especially in the design phase. The necessity of a careful hardware selection under consideration of not only goal-oriented aspects but also under observance of a given room style provides the basis for an effective enrichment with computational power in our everyday environment. The question about the ambience of a system brings up the need to identify more ambient parameters, their interaction, evaluation and manipulation. Perhaps the biggest challenge will be to find an appropriate graduation or scale for describing the ambience level of rooms and hardware. We have to lift up the subjective feel for style elements to a more abstract and scaleable level. Furthermore, if we are able to model the ambient factors that combine hardware and environment with respect to economical constraints, the development process will be more flexible and personalized.

Acknowledgments

Thanks to the reviewers for valuable input and suggestions.

Part of this research is being supported by the German Science Foundation DFG (Deutsche Forschungsgemeinschaft) in its Collaborative Research Center on Resource-Adaptive Cognitive Processes, SFB 378, Project EM 4, BAIR, and its Transfer Unit on Cognitive Technologies for Real-Life Applications, TFB 53, Project TB 2, RENA.

Short Author Biographies

Tim Schwartz studied computer sciences and computational linguistics at the Saarland University in Saarbrücken. Since 2004 he is a PhD student at the chair of Prof. Wahlster where he is working in a project called RENA, which is short for REsource-adaptive NAvigation. His main research interests lie in the field of indoor positioning and hybrid indoor-outdoor-navigation.

Ralf Jung also studied computer sciences at the Saarland

University. He wrote his diploma thesis in a junior researcher group called FLUIDUM, which is short for FLexible User Interfaces for Distributed Ubiquitous Machinery. He joined the chair of Prof. Wahlster in 2005 as a PhD student and is now part of the BAIR project, which deals with user adaptation in instrumented rooms. Currently he is working on an audio notification system.

5. REFERENCES

[1] Ela Data Ingenieursgesellschaft für Elektroakustik. http://www.ela-data.de/.

[2] B. Boehm. A Spiral Model of Software Development and Enhancement. In *SIGSOFT Softw. Eng. Notes*, pages 14–24, New York, NY, USA, 1986. ACM Press.

[3] J. Bohn, V. Coroama, M. Langheinrich, M. Mattern, and M. Rohs. Social, Economic, and Ethical Implications of Ambient Intelligence and Ubiquitous Computing. In *Ambient Intelligence*. Springer-Verlag, 2004.

[4] B. Brandherm and T. Schwartz. Geo Referenced Dynamic Bayesian Networks for User Positioning on Mobile Systems. In *Proceedings of the International Workshop on Location- and Context-Awareness (LoCA), LNCS 3479, ISSN 0302-9743*, volume 3479 / 2005 of *Lecture Notes in Computer Science*, pages 223–234, Munich, Germany, 2005. Springer-Verlag Berlin Heidelberg.

[5] D. Heckmann, T. Schwartz, B. Brandherm, and A. Kröner. *Decentralized User Modeling with UserML and GUMO*, pages 61–66. Decentralized, Agent Based and Social Approaches to User Modeling, Workshop DASUM-05 at 9th International Conference on User Modeling UM2005. Edinburgh, Scotland, 2005.

[6] I. Jacobson, G. Boock, and J. Rumbaugh. *The Unified Software Development Process*. Addison Wesley, 1999.

[7] S. Kurniawan. Interaction Design: Beyond Human Computer Interaction by Preece, Sharp and Rogers (2001), ISBN 0471492787. *Univers. Access Inf. Soc.*, 3:289–289, 2004.

[8] C. Rocchi, O. Stock, M. Zancanaro, M. Kruppa, and A. Krueger. The Museum Visit: Generating Seamless Personalized Presentations on Multiple Devices. In *IUI '04: Proceedings of the 9th International Conference on Intelligent User Interfaces*, pages 316–318, New York, NY, USA, 2004. ACM Press.

[9] W. W. Royce. Managing the Development of Large Software Systems: Concepts and Techniques. In *ICSE '87: Proceedings of the 9th International Conference on Software Engineering*, pages 328–338, Los Alamitos, CA, USA, 1987. IEEE Computer Society Press.

[10] T. Schwartz, B. Brandherm, and D. Heckmann. Calculation of the User-Direction in an Always Best Positioned Mobile Localization System. In *Proceedings of the International Workshop on Artificial Intelligence in Mobile Systems (AIMS)*, Salzburg, Austria, September 2005.

Customer Vendor Intervention Instruction Information

Nathan W. Everett
Microsoft Corporation
One Microsoft Way
Redmond, WA 98052
+1 425 246 8913
nathan.everett@microsoft.com

ABSTRACT
The connotations of the term "user assistance" depend on the context in which it is used. In this paper we propose a user assistance matrix for mobile devices in intelligent environments based on the kind of assistance being offered and the origin of the call for assistance. Creating this matrix makes it possible to understand what issues need to be addressed in order to provide effective user assistance in various scenarios.

Categories and Subject Descriptors
H.5.2 [**Information Interfaces and Presentation**]: User Interfaces – *Theory and Methods, User-centered design, Interaction Styles*; *User Assistance.*

General Terms
Documentation, Design, Theory.

Keywords
Models, Principles, Intelligent Environments, Ubiquitous Computing, Mobile Computing, Intervention, Instruction, Information, User Assistance

1. INTRODUCTION
Our fundamental premises are: a) Mobile devices are ubiquitous; b) Mobile devices are evolving an expanded feature-set; and c) Users and vendors see mobile devices as a desirable way to send and receive personalized information. This paper examines the type of assistance that might logically be offered to mobile device users on their devices.

As mobile devices become more powerful, they also become more complex, potentially leading to greater user confusion and requiring a larger base of learning in order to operate them effectively. This is especially true in the early stages of adoption or during a time of rapid increase in functionality. Usability testing suggests that more powerful devices also lead to higher user expectations for flawless and error-free operation. Users expect to find stored personal information (contacts and phone numbers) and on-demand information (directions and traffic conditions) on their mobile devices. Vendors of mobile content see this as a largely untapped means of promoting their goods and services, but because mobile devices present an equally large target for malicious attack and theft of personal data, individual privacy and security concerns must be considered carefully.

In the Customer Vendor Intervention Instruction Information (CVIII) Matrix, we propose a taxonomy for defining the different kinds of user assistance from the perspectives of the customer (device user) and the vendor (information source).

2. The CVIII Matrix
The active categories in the matrix are spelled out in the title. Columns are headed "Customer" and "Vendor." Rows are labeled "Intervention," "Instruction," and "Information." Table 1, below, describes each intersection of the matrix.

Table 1. CVIII Matrix.

	Customer	**Vendor**
Intervention (Help)	Originates with customer's statement of failure and requires an immediate response.	Originates with vendor's need to get urgent info to the customer; ften regarding a failure not yet recognized by the user.
Instruction (Training)	Originates with customer's question; does not require immediate assistance; customer has time and desire to learn.	Originates with vendor's suggestion offering a reason or motivation to learn more about a service.
Information (Advertising)	Originates with customer's search for information; Results may include paid advertising that gets preferred placement.	Originates with vendor's opportunistic suggestion based on user context, but without a specific user.

The matrix recognizes a difference between assistance in the form of intervention (help), instruction (training), and information (advertising). It distinguishes among these according to who initiates the UA request. The "Vendor" column closely parallels [2]'s "incidental interactions," defined as "where actions performed for some other purpose, or unconscious signs, are interpreted in order to influence/improve/facilitate the actor's future interaction or day-to-day life."[1] Where [2] explores "designed interactions," the CVIII matrix attempts to

[1] "Beyond Intention—Pushing Boundaries with Incidental Interaction," by Alan Dix, Lancaster University, Lancaster, UK.

extend that concept to include broadcast information as well as secondary use of actor (customer) stimuli.

Intervention is always offered when a failure has occurred or is imminent. It is immediate and must be available when the user is in a "non-connected" state. Instruction, on the other hand, is offered at the user's leisure when there is no imminent failure, but a desire to learn. Information is offered opportunistically in response to user requests or location and context awareness.

It is worth mentioning that the "information" row typically does not relate to the device itself. That is, customers seldom initiate requests for information about their mobile devices; they initiate requests for information about the world around them. Similarly, vendors want to push world-related information to the user.

The matrix also differentiates between assistance initiated by the customer and that initiated by the vendor. While help and assistance that address customer-initiated requests are available on nearly every kind of device, there are currently few environments that enable vendor-initiated assistance.

2.1 Customer/Intervention

This scenario typically begins with a failure in expected functionality or in the user experience, which generates a statement of condition. These failures may be widely disparate, but typically the customer knows something is wrong but may not have adequate information to ask a specific question. Some examples include:

- "Help! I'm lost."
- "Help! It didn't work."
- "Help! I can't connect."

Note that these scenarios imply a request for intervention. An immediate task or function has failed and the device is non-operative. The user does not know what kind of question to ask, only that a problematic situation exists.

2.2 Customer/Instruction

This scenario typically begins with an inquiry about a task for which there is inadequate knowledge. Any aspect of the device may be a subject for training, but the key element here is that the customer has a specific scenario in mind when asking. Examples include:

- "How do I use my GPS tracker?"
- "How do I manage music on my device?"
- "How do I send an international text message?"

Note that, in contrast to help requests, customer training requests are typically "how-to" questions. They relate to a scenario, not to a feature; in fact, the user may not know if the scenario is supported at all. In each instance, the user opens a door for exposing functionality.

2.3 Customer/Information

This scenario typically begins with a search query. Through the widespread use of internet search engines, customers have become accustomed to receiving search results that include paid placement, or advertising. This began in telephone directories, where all businesses with telephone service were listed, but a business could pay for various levels of enhanced placement or expanded content. Advertising in this context responds to a customer's request for information. Examples include:

- "What department stores are in this shopping mall?"
- "Where can I buy fresh fish?"
- "Is there a good restaurant near here?"

Most information requests can be phrased as a question, though they may simply be based on a keyword search, as in "restaurants, Bellevue, downtown." This topical selection helps customers get results that are more relevant to their actual request.

2.4 Vendor/Intervention

This scenario typically begins with the vendor identifying a potentially disruptive experience, resulting in vendor-initiated assistance. The customer may not be able to see, or even recognize, the system failure. Nonetheless, the vendor/intervention scenario is defined by the timeliness of the intervention. Some examples of vendor/intervention scenarios include:

- "Help! My customer has left my service area."
- "Help! My customer's device is not responding."
- "Help! My customer turned into a blind alley."

Note that these are statements of a situation the vendor feels merits intervention. From the vendor's perspective, the system has failed. This is the vendor's opportunity to intervene on the customer's behalf, and to provide critical information to the customer *before* it becomes a customer/intervention scenario.

2.5 Vendor/Instruction

This scenario typically begins when the vendor identifries an opportunity to optimize device usage. Customers use a fraction of the features and capabilities of their mobile devices. This is increasingly true as customers are presented with a wider range of usage scenarios. This is an opportunity for vendors to push training to the user rather than waiting for the user to ask for it. Some examples of vendor instruction scenarios include:

- "You have used that feature three times recently. Would you like to automate the process? Here's how."
- "Did you know you can send the photos you have taken to other people? Here's how."
- "There are seven others playing this game within 100 meters of you. Would you like to contact them for joint play? Here's how."

We see the pattern here identifying a scenario and offering "how-to" instruction. The vendor instruction scenarios are opportunistic in that they expand or improve the customer's use of the device and/or services. We have learned from experience, however, that users show little patience for unsolicited tips that are imposed on them unless they have "opted in" for such information.

2.6 Vendor/Information

This scenario typically begins with a vendor's unsolicited broadcast of information. This is advertising. It may take advantage of user context to give actionable suggestions. Information messaging may be highly targeted and individualized. In this instance, messages may be received because of the customer's location. For example:

- Flash: Mariners beat Yankees 6-1. To subscribe to local sports press #618-Send.
- Flash: Did you forget the cat food? Your brand is in aisle seven! (Customer is checking out with kitty litter in basket.)

It is easy to see that such pinpoint targeted advertising might be highly successful, if the customer has elected to receive such offers. If those offers come to the customer by surprise, they may be resented or even feared as an invasion of privacy.

3. USING THE CVIII MATRIX

A clear taxonomy of the problem space will help to define the ways user assistance can be offered. Each cell in the matrix has unique issues that need to be resolved in order to suggest a comprehensive solution. As we compare these issues, we can also see where there are common solutions that can be used across assistance models. We will observe, in fact, that mobile user assistance itself crosses over into user assistance for ubiquitous computing in all its forms. The following table begins the development of the issue matrix but is not meant to be a definitive list.

Table 2. Issues of CVIII user assistance

	Customer	Vendor
Intervention (Help)	- On-device footprint - Context sensitivity - Interrupted connectivity - Screen presentation - Updatability	- Privacy/trust - Context sensitivity - Interrupted connectivity - Screen presentation
Instruction (Training)	- Download - Stream - Discoverability - Consistency - Synchronization - Screen presentation	- Push - Media - Discoverability - Consistency - Screen presentation
Information (Advertising)	- Context - Bandwidth/cost - Search - Subscription - Non-interruption of device use - Screen presentation	- Ubiquity - Context - Bandwidth/Cost - Discoverability - Privacy - Annoyance - Security - Screen presentation - Media - Cost of entry

Identified issues like privacy, trust, interruption, and annoyance must be investigated concurrently with the design of user assistance. [4] defines "implicit HCI" as "an action, performed by the user that is not primarily aimed to interact with a computerized system but which such a system understands as an input. It is further identified that perception and interpretation of the user, the environment, and the circumstances are key concepts for implicit HCI."[2] In many cases, however, the actions that set context for interpretation do not need sensing apparatus beyond that which is already present on advanced mobile devices. Both GPRS and WiFi networks can identify location (where the customer is) and presence (whether the customer is available for contact on the device). Technologically, we have the ability to provide user assistance in many intelligent environments; but emotionally, we may find resistance to any "help" that is offered without a request.

It would be a mistake to emulate failed designs for such assistance. A noted example of this is Microsoft Office's "Clippy" assistant which has been removed from their products. We will find that the amount of intervention that customers will tolerate likely will vary by context and by the customer's understanding of what is offered, at what cost (monetary, privacy, annoyance, etc.), and with what benefit (convenience, efficiency, etc.).

Our proposed next steps in this investigation are to 1) identify more completely the matrix of issues, 2) determine user tolerance for vendor-initiated UA in different contexts, 3) recommend contextual designs for UA that satisfy vendor desires and customer concerns.

4. ACKNOWLEDGMENTS

Thanks to the Microsoft Industry Innovations Group for video example of Retail Vision. Thanks to reviewers of this paper for excellent suggestions on references. Thanks to Ethan Casey and Michele Palmer for editing.

5. ABOUT THE AUTHOR

Nathan Everett is the User Assistance Strategist for Microsoft Mobile and Embedded Devices. He is not an academic researcher. He has been the UA Strategist for Microsoft MED since 2004. In that capacity he researches trends in mobile device usage and determines how Microsoft can best provide user assistance for the Windows Mobile platform.

6. REFERENCES

[1] Chen, Eugene, and Marcus, Aaron. Designing the PDA of the Future. *Interactions*. 9, 1 (January 2002), 34-44.

[2] Dix, A. "Beyond Intention - pushing boundaries with incidental interaction." Proceedings of Building Bridges: Interdisciplinary Context-Sensitive Computing, Glasgow University, 9 Sept 2002

[3] Sands, I. "Retail Vision: Transforming Business with Technology." Microsoft Industry Innovations Group, April 2005.

[4] Schmidt, A. "Implicit Human Computer Interaction Through Context." Personal Technologies, 4(2&3), Springer-Verlag, pp. 191-199, June 2000.

[2] "Implicit Human Computer Interaction Through Context" by Albrecht Schmidt, University of Karlsruhe, Karlsruhe, Germany.

CHAT – Towards a general-purpose infrastructure for multimodal situation-adaptive user assistance

Carmelo Ardito[1] Thomas Pederson[1,2] Maria Francesca Costabile[1]

[1]Dipartimento di Informatica
Università di Bari
70125 Bari, Italy
+39 080 544 3300
{ardito, costabile, pederson}@di.uniba.it

[2]Department of Computing Science
Umeå University
SE-90187 Umeå, Sweden
+46 90 786 6548
top@cs.umu.se

ABSTRACT

In this position paper, we describe CHAT, an upcoming project aimed at providing multimodal context-sensitive services to mobile users. We specifically describe two conceptual corner stones of the project of relevance to the topic of user assistance in intelligent environments: a) a multimodal interaction framework targeted at providing service access through different modalities for different real-world situations and for improving interaction with mobile devices in general, b) an "egocentric interaction" model for framing interaction with objects in the vicinity of a mobile user, including also other real-world and/or computational entities than the mobile device itself, ranging from computationally "stupid" everyday objects to more advanced interactive devices such as desktop PCs. The final section of the paper is devoted to open issues in the design of the CHAT infrastructure related to the topic of user assistance in intelligent environments.

Categories and Subject Descriptors

D.3.2 [**Design Tools and Techniques**]: User interfaces; D.2.11 [**Software architectures**]: Domain-specific architectures; H.1.2 [**User/Machine Systems**]: Human factors; I.3.6 [**Methodology and Techniques]**: Interaction techniques; K.8 [**Personal Computing**]: Miscellaneous.

General Terms

Design, Human Factors.

Keywords

Multimodal interfaces, mobile human-computer interaction.

1. INTRODUCTION & PROBLEM DESCRIPTION

Mobile computing implies computing in more varied physical contexts than desktop computing. Different situations enforce different constraints as to what kind of device interaction and assistance that is needed; technically possible to offer; socially acceptable to perform; etc. To some extent, interaction style is also a matter of preference (e.g. some people prefer headsets when using cellular phones, others prefer using only the phone device itself).

In the CHAT project, "Cultural Heritage fruition & e-learning applications of new Advanced (multimodal) Technologies", we intend to develop a software infrastructure that allows services accessed through thin clients such as cellular phones or PDAs to be a) adaptable to personal preferences of the user, with focus on the choice of interaction modalities, and b) adaptive to the physical-virtual context of the human actor carrying the device. In both cases, the proposed architecture should be open both for channeling interaction between services and user through the mobile device itself, as well as through available input and output facilities in the vicinity. Furthermore, real-world phenomena sensed by the device itself or indirectly through external sensor pools will be made available through the CHAT infrastructure as a resource for service developers to effectively design "intelligent" environments.

The multimodality of interest for us permits users to interact with the system using several input channel simultaneously, classified by W3C as simultaneous co-ordinated multimodality [8]. In the research context there are already similar systems (e.g. [2]) and empirical studies (e.g. [1]) targeting this kind of multimodality, but also commercial tools with these interaction features are currently developed.

The architecture for supporting these kind of multimodal systems is more complex than traditional interactive systems, because we have to consider:

- parallel recognition modules for each input channel: every module produces fragments of the overall input that must be combined to become meaningful;
- a general methodology to interpret the meaning of the input fragments;
- a time-sensitive analysis process to determine which fragments must be combined to become meaningful;
- a module to manage the overall user/system dialogue;
- criteria to adapt the input/output modalities to the users' needs and the environment in which they actually are.

We deliberately try to make the infrastructure as general as possible because we believe that multimodal adaptability and adaptive features are beneficial independent from the area of application. For the purpose of evaluation however, the CHAT project will develop mobile multimodal prototype applications related to two particular activities: e-learning and exploration of cultural heritage.

1.1 E-learning

E-learning is an area investigating the possibilities of improving the learning process by using information technology. It enables new forms of learning including distance learning and just-in-time learning. Advantages include new possibilities in distributing learning material through both computer and communications technology. Such devices can include personal computers, CDROMs, Television, PDAs, and Mobile Phones. Communications technology enables the use of the Internet, email, discussion forums, collaborative software, classroom management software. Courses can be tailored to specific needs and asynchronous learning is possible. The "any time, any place" nature of e-learning could be a winning strategy for particular needs, such as decongestion of overcrowded education facilities, support for learners or lecturers who live far from schools and universities, life-long education. Moreover, it could be a valuable opportunity for specific groups of learners, such as disabled learners, if the learning material is actually accessible to them.

1.2 Cultural heritage exploration

Cultural heritage is the legacy of physical artefacts and intangible attributes of a group or society that are inherited from our ancestors, maintained in the present and bestowed for the benefit of future generations. The term "cultural heritage" has not always meant the same thing. Having at one time referred exclusively to the monumental remains of cultures, heritage as a concept has gradually come to include new categories such as the intangible, ethnographic or industrial heritage. A noteworthy effort was subsequently made to extend the conceptualization and description of the intangible heritage. This is due to the fact that closer attention is now being paid to humankind, the dramatic arts, languages and traditional music, as well as to the informational, spiritual and philosophical systems upon which creations are based. Physical or "tangible cultural heritage" includes buildings and historic places, monuments, artefacts, etc., that are considered worthy of preservation for the future. These include objects significant to the archaeology, architecture, science or technology of a specific culture. "Natural heritage" is also an important part of a culture, encompassing the countryside and natural environment, including flora and fauna. These kind of heritage sites often serve as an important component in a country's tourist industry, attracting many visitors from abroad as well as locally.

2. FRAMEWORKS

The system development is guided by a multimodal design framework for ensuring state-of-the art support for multimodal interaction, and an egocentric interaction design framework for guiding the analysis and design of context-aware services.

2.1 Multimodal design framework

We propose a framework inspired by W3C's Multimodal Interaction Framework that identifies major components for any multimodal system (see Figure 1).

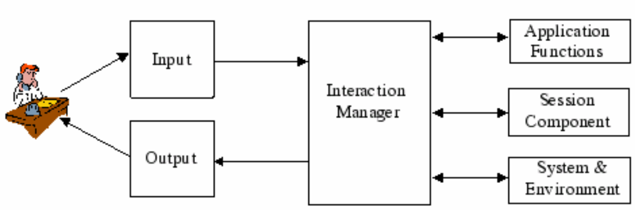

Figure 1. The W3C Multimodal Interaction Framework [7].

The W3C multimodal interaction framework is not an architecture, but a modeling framework one step above in abstraction. It describes neither how components are allocated to hardware devices, nor how the communication system enables the hardware devices to communicate. [7]

2.1.1 *Components of the W3C multimodal interaction framework*

Human user - A user who enters input into the system and observes and hears information presented by the system.

Input - An interactive multimodal implementation will use multiple input modes such as audio, speech, handwriting, and keyboarding, and other input modes.

Output - An interactive multimodal implementation will use one or more modes of output, such as speech, text, graphics, audio files, and animation.

Interaction manager - The logical component that coordinates data and manages execution flow from various input and output modality component interface objects. The interaction manager maintains the interaction state and context of the application and responds to inputs from component interface objects and changes in the system and environment. The interaction manager then manages these changes and coordinates input and output across component interface objects.

Application Functions – The services that should be offered to the users.

Session component - An interface to the interaction manager to support state management, and temporary and persistent sessions for multimodal applications.

System and Environment component - A component that enables the interaction manager to find out about and respond to changes in device capabilities, user preferences and environmental conditions. For example, which of the available modes, the user wishes to use — has the user muted audio input? The interaction manager may be interested in the width and height of the display, whether it supports colour, and other capability and configuration information.

The framework designed for the CHAT project (see Figure 2) is compatible with the Multimodal Interaction Framework previously discussed.

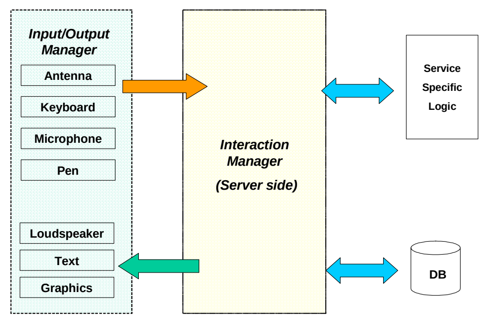

Figure 2. CHAT Multimodal Interaction Framework.

The *Input/Output Manager* is a "lightweight" software running on the user device with the responsibility of managing the input and output channels. I/O Manager captures the input fragments coming from several channels and transmits them to the Interaction Manager running on a server using suitable standard protocols.

The *Antenna* in the framework symbolizes high-level position information received through for instance GPS, GSM, or RFID technologies.

The *Interaction Manager (IM)* receives the multimodal input fragments from the user's device and processes them to obtain a meaningful input.

The *Service Specific Logic* acts on the input received and potentially produces an output transmitted to the IM.

Finally, the IM generates a multimodal representation of the required service that will be presented to the user by output channels suited to the user's preferences and needs as well as current environmental context and device type.

2.2 Egocentric interaction framework

The user interface design in the CHAT project will be guided by the *egocentric interaction* framework [3], inspired by the currently popular view within cognitive science that human individual actions are to a large degree influenced by what the specific individual can perceive of the surrounding environment. Based on a an integrated view on physical and virtual space[1] [4] objects in the proximity of a particular human actor can be categorized as being situated in one out of four spaces, at any given point in time (see Figure 3).

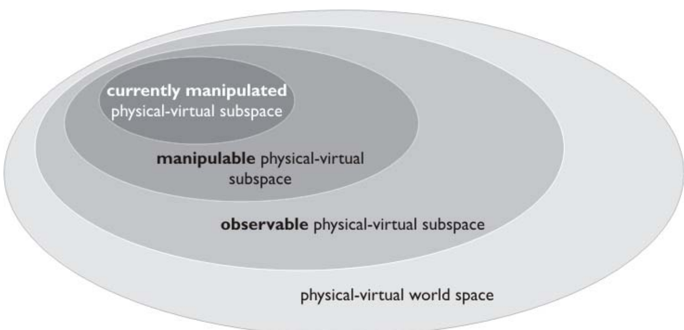

Figure 3. A situative model of physical-virtual space from the perspective of a specific human actor at a specific point in time. Adapted from [4].

The borders between the different subspaces are based on assumptions of the specific human actor's perceptive and cognitive experience of the current physical-virtual environment, e.g. real-world objects and/or virtual objects (presented by, and accessed through computing devices) in the immediate vicinity of the specific human actor. The egocentric interaction framework is based on the belief that computer system models should be closely tied to the cognitive and perceptive models that human actors construct and maintain as part of everyday life.

The term 'egocentric' has been chosen to signal that it is the human body and mind of a specific human individual that (sometimes literally) acts as a centre of reference to which all interaction modeling and activity assistance is anchored.

2.2.1 Components of an egocentric interaction system

The conceptual system architecture (illustrated in Figure 4) is based on a wearable computing/sensing hardware configuration consisting of a *private black box* offering computing power and storage space for data generated by an *egocentric interaction sensor pool* monitoring object-centric phenomena within the observable physical-virtual subspace of a specific human actor. Furthermore, the private black box runs a *physical-virtual operating system* hosting both advanced physical-virtual applications developed by software developers as well as simpler programs designed by the user her/himself. Such applications can incorporate the manipulation of both physical objects (e.g. a sculpture at a museum) and virtual objects (e.g. a web page describing the same sculpture). Explicit interaction with the physical-virtual operating system is performed through a *private black box user interface*, either fitted onto the private black box itself, or running on a general-purpose device like a PC. Implicit interaction [5] with the physical-virtual operating system emerges whenever the user interacts with a physical or virtual object inside the manipulable physical-virtual subspace (see Figure 3) monitored by the private black box. The local computing of the private black box can optionally be enhanced by communication with publicly and ubiquitously accessible *shared object knowledge repositories*, distributing anonymous data about objects and their everyday use.

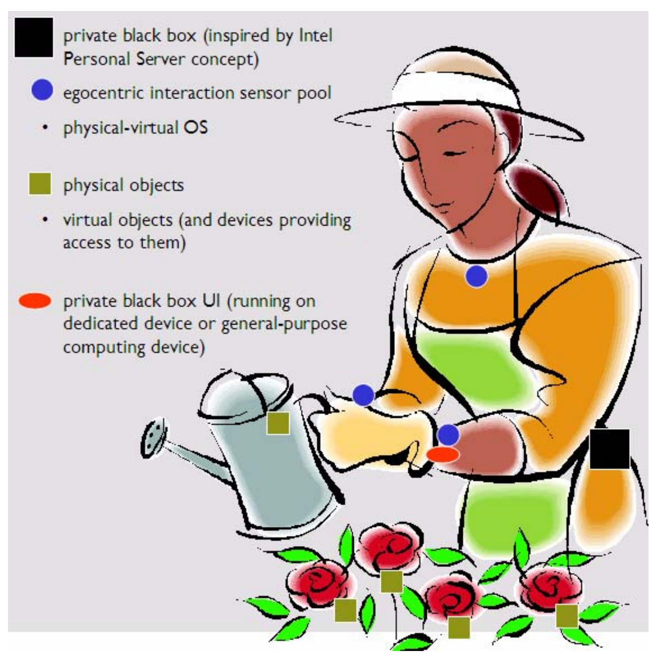

Figure 4. General components of an egocentric interaction system. (Virtual objects and shared object knowledge repositories not pictured.)

[1] In short, "physical space" denotes the euclidean three-dimensional space of the real world. "Virtual space" is the multi-dimensional hyperspace (locally often euclidean two-dimensional) accessed through interactive computing devices such as cellular phones, PDAs and desktop computers.

2.2.2 Potential application areas for egocentric interaction systems

- (physical and virtual) object logistics support — make sure you have everything you need, make sure unnecessary distracting objects are out of the way
- activity sequence support — make sure you do it in the order necessary
- physical-virtual gap bridging support — allow for smooth transitions between performing actions in the physical world and in the virtual world
- provide access to virtual environments and objects, including to place virtual objects in physical space, to place physical objects in virtual space
- provide manual search and recall functionalities within the log database of previously performed activities
- provide simple support for end-user development / programming of physical-virtual applications
- extension of the user's senses: providing info on when the bus will arrive, how far away the person you should meet is located

3. FUTURE WORK

The first step of the project will be to identify a set of concrete use cases in the two application areas of interest. We hope to acquire knowledge and hints from this workshop that can help us make important design decisions, e.g. when in the studied activities user assistance would be useful and in what form (e.g. which kind of assistance, through what modality, and when).

4. OPEN ISSUES AND WHY WE WOULD LIKE TO ATTEND THE WORKSHOP

We would come to this workshop with more questions than solutions although we do believe that the multimodal framework and the egocentric interaction model could contribute and help framing the discussion on user assistance in intelligent environments.

Of the questions listed in the workshop call, we find the following very relevant for our aim in CHAT:

- **Which user activities and tasks require assistance?** What method or heuristic should we use to identify the most useful and practically achievable assistance for e-learning and exploration of cultural heritage?
- **How should the designer choose the best sensing and interaction technologies for a scenario?** Having a small mobile device as computation and communication hub between user, ("intelligent") environment, and server (see Figure 4): how should the service logic (see Figure 2) be developed in order to seamlessly cope with ad-hoc appearance and disappearance of sensors and actuators external to the device itself?
- **How should multiple users with concurrent activities be supported?** E-learning and exploration of cultural heritage are activities which often dynamically change from being performed in isolation to being performed in groups. When is synchronous communication in a community better then asynchronous? When experiencing a common phenomena or facing a common problem, in what way can a group of users be encouraged to assist each other in order to increase the knowledge level and experience quality as a group?
- **How should the current state of the user assistance system be represented, especially when dealing with multiple tasks?** We believe that the egocentric interaction perspective, although still very abstract, constitutes an interesting step towards modeling complex and concurrent tasks because a) it attempts to capture what the user views as important at any given time, b) it does so by making no distinction between objects of interest residing in the physical world and objects in the virtual world (i.e. inside computing devices).

Of the key topics mentioned in the workshop call we have the following starting points for discussion:

- **How to unify the complementary concepts of public and personal devices in IEs**. Being inspired by the Intel Personal Server concept [6], We propose to regard interaction *data* as private, and interaction *devices* (including mobile and wearable ones) as potentially public. Distribution of interaction devices among individual users is transparent and can follow established social rules, while distribution of digital data is by nature invisible and thus has to be restricted. From another point of view, the sought-for seamless interplay between (personal) mobile devices and more stationary computational "intelligence" is an important requirement, because without smooth such mechanisms, knowledge of individual preferences will have to be inferred on the fly rather than retrieved (at least in an "occasional user" scenario), making the environment look significantly less "intelligent".
- **How to model user activity (terminology, structure, notation) for the design of IEs.** We believe to offer some ideas towards framing activity in intelligent environments by proposing the egocentric interaction framework which centres the attention of the modeler to the users body and mind, and which regard physical and virtual objects as residing in the same physical-virtual space. This stance should be interesting because intelligent environments are almost per definition relying on a combination of real-world and computational phenomena. There is of course more to human activity than mere object manipulation and we look forward to getting feedback and ideas for complementing the model.

5. SHORT AUTHOR BIOGRAPHIES

Carmelo Ardito is, since November 2004, Ph.D. Student of the School of Informatics at the Dipartimento di Informatica, University of Bari. He got the laurea degree in Computer Science at the University of Bari in June 2002 and the laurea specialistica degree in Computer Science at the University of Bari in October 2004, discussing a thesis on "an experimental evaluation of the use of audio in virtual environment navigation". Since 2001 he has been research collaborator of the Department of Computer Science in various projects, sponsored by European Union and Italian organizations. His current research interests are in the Human-Computer Interaction field, particularly Information Visualization Usability Techniques, Mobile Systems, Web-based systems.

Thomas Pederson received his PhD from the Dept. of Computing Science at Umeå university, Sweden, in December 2003. Faculty opponent at the dissertation event was professor William Buxton, University of Toronto, Canada. His PhD thesis proposes a uniform and integrated view on physical and virtual objects, environments, and activities. Thomas did his MSc project at Ericsson Media Lab in Stockholm, and has spent six months as visiting research assistant at Fraunhofer IPSI in Germany prior to his PhD studies as well as two months in the ACES research group in Rennes, France, after his PhD dissertation. He is currently project manager of easyADL, a two-year project intended to investigate the use of Ubiquitous Computing technology to support activities of daily living for individuals suffering dementia disease. In addition to the managing of the easyADL project in Sweden, Thomas holds a post doc fellowship at Bari university in Italy for 2006 and 2007.

Maria Francesca Costabile is full professor at the Computer Science Department of the University of Bari, Italy, where she teaches HCI for the computer science curriculum. Her current research interests are in both theoretical and application oriented aspects of visual formalisms for information representation and querying, visual system design, visual data mining, adaptive interfaces, user models, multimodal and multimedia interaction, web interfaces, system usability. She is regularly in the program committees of international conferences and workshops and has been Program Co-Chair of INTERACT 2005. She has been General Chair of the International Working Conference on Advanced Visual Interfaces (AVI) in 2004 and Program co-Chair of AVI '96 and AVI'98; she is in the AVI steering committee (AVI is organized in cooperation of ACM SIGCHI). She is senior member of IEEE and member of ACM and ACM SIGCHI. She was founding member of the Italian Chapter of ACM SIGCHI (SIGCHI Italy) and served as chair from 1996 to 2000.

6. REFERENCES

[1] Oviatt, S., De Angeli, A. and Kuhn, K. (1997). Integration and synchronization of input modes during multimodal human-computer interaction. In Proceedings of CHI'97, Atlanta, Georgia, USA, 18-23 Apr, 1997, pp. 415-422.

[2] Paternò, F., & Giammarino, F. (2006). Authoring interfaces with combined use of graphics and voice for both stationary and mobile devices. Proceedings of the International Conference on Advanced Visual Interface 2006 (AVI 2006), Venice, Italy, May 23-26, 2006, pp. 329-335.

[3] Pederson, T. (2006). Egocentric Interaction. Workshop on What is the Next Generation of Human-Computer Interaction?, CHI2006, April 22-23, Montréal, Canada.

[4] Pederson, T. (2003). *From Conceptual Links to Causal Relations — Physical-Virtual Artefacts in Mixed-Reality Space*. PhD thesis, Dept. of Computing Science, Umeå university, report UMINF-03.14, ISSN 0348-0542, ISBN 91-7305-556-5. Permanent URL: http://urn.kb.se/resolve?urn=urn:nbn:se:umu:diva-137

[5] Schmidt, A. (2002). Ubiquitous Computing – Computing in Context. PhD thesis, Computing Department, Lancaster university, U.K.

[6] Want, R., Pering, T., Danneels, G., Kumar, M., Sundar, M., Light, J. (2002). The Personal Server: Changing the Way We Think about Ubiquitous Computing. Proceedings of UBICOMP 2002, Göteborg, Sweden, Springer Verlag, September 2002, pp. 194-209.

[7] W3C Multimodal Interaction Framework. http://www.w3.org/TR/mmi-framework/

[8] W3C Multimodal Interaction Activity. http://www.w3.org/2002/mmi/

A Study of Situation Understanding Method Utilizing Environments for Multimodal Observation of Infant Behavior

Shogo Ishikawa
Shizuoka University
3-5-1 Johoku Hmamatsu
Shizuoka Japan
cs1007@s.inf.shizuoka.ac.jp

Naofumi Otani
Shizuoka University
3-5-1 Johoku Hamamatsu
Shizuoka Japan
nao@takebay.net

Shinya Kiriyama
Shizuoka University
3-5-1 Johoku Hamamatsu
Shizuoka Japan
kiriyama@inf.shizuoka.ac.jp

Hiroaki Horiuchi
Shizuoka University
3-5-1 Johoku Hmamatsu
Shizuoka Japan
horiuchi@inf.shizuoka.ac.jp

Yoichi Takebayashi
Shizuoka University
3-5-1 Johoku Hmamatsu
Shizuoka Japan
takebay@inf.shizuoka.ac.jp

Shigeyoshi Kitazawa
Shizuoka University
3-5-1 Johoku Hmamatsu
Shizuoka Japan
kitazawa@inf.shizuoka.ac.jp

ABSTRACT
We have developed ubiquitous and wearable environments which allow us to observe infant behavior from various perspectives. Our goal is to represent human commonsense knowledge as the computational models, which are applied to the spoken dialogue systems that realize smart and clever man-machine interactions by understanding speakers' intentions and emotions appropriately. We have established infant learning environments using natural materials of the cedar, which enables infants to be absorbed in the activities comfortably and safely. We have constructed environments for infant behavior recording; we have set up multiple cameras and microphones in the Cedar yurt. We have also developed a wearable speech recording device to capture infant utterances with high quality. Developing a comment-collecting system which allows everyone to make comments easily from the multi-viewpoints, we have realized a framework for multimodal observation of infant behavior. Utilizing the multimodal environments, we propose a situation description model based on observation of demonstratives uttered by infants, since demonstratives appear frequently in their conversations and become a precious clue to understand situations. The proposed model, which represents the mental distances of speakers and listeners to objects on a general and simple model, enables us to predict speakers' next behavior. The consideration results enable us to conclude that the constructed environments lead to development and realization of human interaction models applicable to user assistance systems for elder people supporting.

Categories and Subject Descriptors
H.1[Models and Principles]: User/Machine Systems; -*Human factors, Human information processing*

General Terms
Human Factors.

Keywords
user assistance systems, situation understanding, multimodal observation environments.

1. INTRODUCTION
As more and more progress has been made in the field of ubiquitous and wearable technologies, abilities and areas of human activities have been more widespread. As for “how to support human activities”, the focal method has moved from “what device should be useful for a target activity”[1] to “how the total environments for the activity should be designed”[2, 3]. Information systems supporting human activities require a function to supply appropriate information in good timing. In order to realize such a function, it is indispensable to elucidate fundamental mechanisms of intelligence controlling human activities, such as “how humans learn things”, or “how they decide their behavior”. In the existing circumstances, however, such mechanisms are little understood.

In order to solve the above problems, we aim at revealing the mechanisms of learning or behavior decision of humans by observing particularly infants whose intelligence is under development.

From this viewpoint, we have designed and developed environments for recording and observing infant behavior in the following methods;

1. Establish an infant school to observe infant behavior.
2. Develop ubiquitous and wearable environments enabling us to record infant utterances and actions in the classes of the school comprehensively.
3. Construct workflows and frameworks enabling many people of our research group to observe infant behavior from their own viewpoints.

Many researches on the observations of infant behavior have been conducted; the language acquisition process has been investigated by analyzing the sounds of infant utterances[4, 5]. Using video data recording the scenes of playing of infants with parents, the acquisition processes of spoken language and gesture have been examined[6–8]. Most of such researches, however, are researches that test a hypothesis postulated exclusively for their own purpose, and the data collected by the observations are limited to those of single modality. On the contrary, we utilize multi-stream data recorded by cameras, microphones and sensors, and observe infant behavior from multimodal viewpoints, focusing on multiple modalities such as utterances, gestures, and sights.

Utilizing these environments, we have started constructing a 'Multimodal Infant Behavior Corpus,' which makes a valuable contribution to the elucidation of human commonsense knowledge and its acquisition mechanism [9]. This paper describes our efforts at extracting commonsense knowledge which can apply to dialog systems in a wide range of domains by means of utilizing this corpus. First, we have observed the utterances of demonstratives which appear most frequently in the actions of one- or two-year-old infants. Then, we have proposed a model which can describe mental states for each utterance using uttered demonstratives as clues. The proposed model has enabled us to study a framework of situation understanding which can be incorporated into user assistance systems of various domains.

In the next section, we describe our ubiquitous and wearable environments developed for infant learning. Section 3 consists of the practice of our infant school using the developed environments, and the framework of infant behavior observations. A situation description model is proposed as the results of the observations of demonstrative utterances in section 4. We consider the changes of mental states using the instances of behavior description by the proposed model and study how to apply the results to predict the actions of speakers in section 5. Section 6 concludes the paper.

2. ENVIRONMENTS FOR INFANT LEARNING AND BEHAVIOR RECORDING

We have designed infant learning environments to manage our infant school smoothly and to proceed infant behavior observation efficiently. For the purpose of realizing learning environments which allow infants to learn and play at ease, the school is constituted of the cedar PAO yurt as shown in Figure 1. The environments provide a calm and comfortable atmosphere of the cedar to the participants, infants and their parents, which enables infants to behave naturally.

In the following subsections, we expound recording environments for capturing thoroughly the features of infant behavior from various perspectives.

2.1 Infant Behavior Recording at the School

In order to capture infant behavior in detail, many pieces of equipment such as cameras and microphones are required. On the other hand, the sight of the equipment may cause infants to behave unnaturally. The yurt avoids this problem. Utilizing the beams of the yurt, we can arrange displays, lights, speakers, cameras, and microphones freely and inconspicuously, instead of sticking the equipment on the wall or the ceiling of a room. Furthermore, we can change the arrangement of the equipment flexibly corresponding to the variation of activities in the class and of observation methods. Figure 2 shows the recording equipment in our school.

Fig.1 An infant learning school constituted of a cedar yurt.

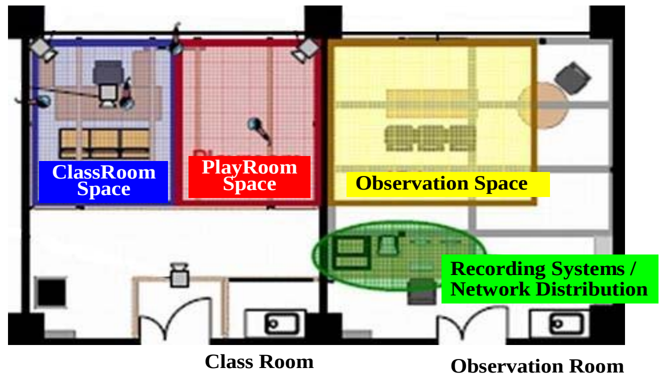

Fig.2 Multimodal infants behavior observation environments

Each class of our school consists of two parts; one is a classroom time on the desk, and the other is a playroom time on the floor. The recording in the classroom time is easily conducted by installing cameras and microphones in their suitable places. On the contrary, the playroom time is hard to record, because infants move around freely. In the case of video recording in the playroom time, the following problems have occurred;

1. Since many cameras are required, it is difficult to find target scenes from the huge amount of recorded data.

2. Infant behavior needs to be recorded not only in an angle just around infants, but also in many angles capturing hand motions or facial expressions.

To cope with the above problems, we place four cameras which allow us to operate them by the remote control shown in Figure 3. Through the Internet, we can follow movements of each individual infant using the cameras.

As mentioned above, we have succeeded to develop the ubiquitous environments enabling us to record the behavior of infants wherever they move around.

2.2 Recording of Infant Utterances

Utterances include important information which explains mental behavior such as intentions or emotions. To analyze properly such information, which is conveyed by prosodic features of speech, it is essential to record utterances with less noise and high quality. Speech data recorded by microphones embedded in the beams of the yurt had the following problems:

1. Large ambient noise was included.

Fig.3 A network camera which enables remote control.

Fig.4 A wearable speech recording device.

Fig.5A Web-based comment collecting system.

2. Sound volume was not stable.

3. Speaker identification was difficult.

To cope with these problems, we have developed a wearable speech recording device shown in Figure 4. Two condenser microphones are arranged near both shoulders. Recorded speech is stocked in a voice recorder stored inside of the rucksack.

We have investigated the quality of speech data recorded by the developed device. The result proved that the new device overcame all of the above problems.

3. PRACTICE OF AN INFANT SCHOOL AND BEHAVIOR OBSERVATION

3.1 Management of the Infant School

We have held an infant school with two sixty-minute classes every week since June, 2005. One of the classes is for the one-year-old and the other is for the two year-old. Each class consists of a teacher and three pairs of infant and mother.

Two teachers manage each of the classes under the supervision of an infant education expert who has held his own infant school for more than thirteen years. The former thirty minutes of each class is a classroom time; the teachers show infants various activities using clay, paper, scissors and pencils. The latter half hour of each class is a playroom time; while mothers get advices about parenting from the teacher, infants play together using toys such as blocks and balls. More than sixty classes have been held for ten months. The length of recorded video and speech data far exceeds 120 hours.

3.2 A Tool for Infant Behavior Observation

In order to grasp the features of rapidly growing infants, it is necessary to continuously record multimodal data including video and speech streams. However, it is far more important to add meta-information about not only general records of classes such as dates, names of participants, and contents of activities, but also the results of infant observation.

In observing infants, it is indispensable to collect comments from not only researchers and teachers, but also parents who know best what infants have learned and they have liked to play recently.

In order to achieve this, we have developed a Web-based comment collecting system that enables us to access anytime via the Web and to make comments in natural language at any position of the video stream as shown in Figure 5. The system allows various observers such as teachers, researchers and parents to access the data and to add or modify comments freely.

According to the practice of the infant school, the developed system has been used for ten months. Fourteen people made comments for infant behavior in the classes. The comments for 1,832 scenes were collected.

4. THE CONSTRUCTION OF SITUATION UNDERSTANDING MODEL BY THE OBSERVATION OF DEMONSTRATIVES

4.1 The Observation of Infant Utterances Focused on Demonstratives

We have observed demonstratives uttered by infants by using recorded video and speech data. First, the items of annotation required to describe the features of demonstrative utterances were enumerated. As the result of preliminary experiments using a part of recorded data, we have selected out the following eleven items:

(1): utterance time

(2): utterance content

(3): behavior with an utterance

(4): object

(5): listener

(6): physical distance between a speaker and an object

Table.1 A situation description model based on demonstrative utterances.

(9)(10) psychological distance between a speaker or a listener and an object	(2) utterance content	(6) physical distance between a speaker and an object	(7) physical distance between a speaker and a listener	(8) physical distance between a listener and an object
I S L ★ ×	kore(This)	0	0 or close	close
		distant	0 or close	distant
II S L ★ ×	atchi(There) are(That)	distant	close or distant	distant
III S L ★	kore(This) kotchi(This way) koko(Here)	0	0 or close or distant	close
		close or distant	0 or close or distant	close or distant
IV S L ★	kore(This) kou(Like this)	0	0 or close	close
V S L ★	are(That)	close	0 or close	close
		distant	0 or close	close or distant

※S: Speaker L: Listener ★: object ×: don't care

(7): physical distance between a speaker and a listener

(8): physical distance between a listener and an object

(9): mental distance between a speaker and an object

(10): mental distance between a listener and an object

(11): notes

The items (1) to (3) are essential to the observations based on video and speech data. The next three items (4) to (6) are added by referring to a relevant study [10] which has observed the practical use of demonstratives focused on the physical distances between speakers and objects. The items (7) to (10) are incorporated for the analyses based on the theory of "territory of information."[11]. Supplementary description is made possible by the note (11).

4.2 Concrete Examples of Situation Description

Since the two-year-old infant has a stable skill to distinguish between *"kore"* (this) and *"are"* (that), his demonstrative utterances have been analyzed in detail. We have examined the situation description of demonstrative utterances through the analysis. Mental states of the infant as a speaker have been considered by using the one example scenes as follows:

Scene A (Figure 6): the infant says "kore" (this) referring to a close object, and after a while, he says "are" (that) to another distant object with pointing finger at it. In these cases, there is no discrepancy between Physical and mental distances.

A-1: The infant says "kore" (this) touching a close chair. This shows that the chair is in territory of the infant. Also, his mother as a listener is aware of that and responds with a nod.

A-2: The infant says "are" (that) referring to another object far from the chair with finger pointing. Since this object is in territory of his mother as a listener and distant physically from him, the object is also distant mentally.

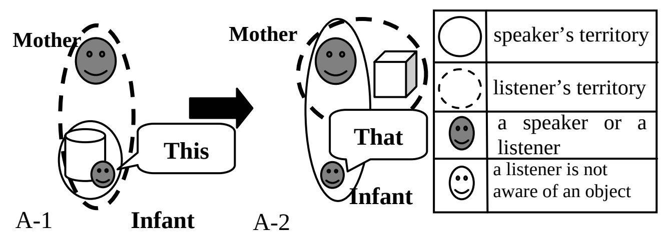

Fig.6 A example of physical and mental distances have no discrepancy between them.

4.3 The Proposition of a Situation Description Model

Focused on the two kind of mental distance (hereafter MD), "(9) MD between a speaker and an object" and "(10) MD between a listener and an object" in Section 4.1, we have classified the observed situations into five categories as shown in Table 1. We propose this categorization as a situation description model. The third to sixth columns of Table 1 indicate the examples of the items (6) to (8) about the physical distance for each pattern. The estimation of mental states is realized by utilizing the following information jointly: the physical distances calculated using video and sensory data and the kinds of uttered demonstratives.

Category I: the MD of a speaker is <in territory>, and the MD of a listener is <don't care>. This is the most popular pattern in the observation. In the most cases of this category, a speaker utters to take a notice of a listener when the speaker wants to do something and the listener is not aware of it.

Category II: the MD of a speaker is <out of territory>, and the MD of a listener is <don't care>. For instance, the following case belongs to this category: a speaker is interested in a distant object, and the speaker make a listener be aware of it. The difference from **Category I** is the physical distance between a speaker and an object.

Category III: the MD of a speaker is <in territory>, and the MD of a listener is also <in territory>. In many cases, a listener recognizes clearly an object, or an object belongs to a listener. Moreover, the case that a speaker is about to be robbed of an object by a listener belongs to this category. **A-1** in section 4.2 are examples of this category.

Category IV: the MD of a speaker is <in territory>, and the MD of a listener is <out of territory>. The typical case of this category is that a speaker explains an object which the speaker has to a listener, or that a listener mentions an object which a speaker has to the speaker.

Category V: the MD of a speaker is <out of territory>, and the MD of a listener is <in territory>. The following cases are examples: a speaker recognizes that an object belongs to a listener, and it is impossible for a speaker to get an object by oneself. **A-2** in section 4.2 belongs to this category.

5. MENTAL SITUATION UNDERSTANDING BY THE PROPOSED MODEL

5.1 The Observation of Mental States Transition in Scenes

Using the situation description model proposed in section 4.3, it is possible to represent the transition of mental states as the change of the proposed five **categories I** to **V**. From this viewpoint, we have proceeded to observe demonstratives in infant utterances further closely. As the examples of state transition patterns, the following two scenes were extracted.

Scene B(Figure 7): An example of mental state transition **"I => III => V"**.

B-1: The two-year-old infant points a finger at a toy on a distant wall, and says *"kore"* (this). On the other hand, the teacher as a listener next to the infant is not aware of the toy. This utterance is to get the notice of the teacher, and has the function of drawing the listener into the territory of the speaker. Although the MD of the speaker for the object is <in territory>, the listener is <don't care> for it; this belongs to **Category I**.

B-2: Though the infant says *"kore"* (this) with pointing the object again and again, the teacher only says "Later!" and tries to make the infant do another thing. So the infant comes close to the object on the distant wall with pointing it, and says *"kore"* (this) with looking at the teacher. Since the speaker and the listener are close to the object mentally, this is classified into **Category III**.

B-3: The teacher recognizes the object, repeats the same words "After the class finished." The infant hears it, goes back to the teacher's knee, and says *"are"* (that) for the object. After then, the infant never mentions the object. The change of demonstratives indicates that the infant's desire has been suppressed for some reasons. The infant insists on his desire repeatedly, which is not accepted by the teacher as a listener. At last, he gives up his desire. It is highly possible that the reason of the suppression is the determinate attitude of the teacher. Since the infant recognizes that it is impossible to obtain the object, he utters *"are"* (that). This situation belongs to **Category V**.

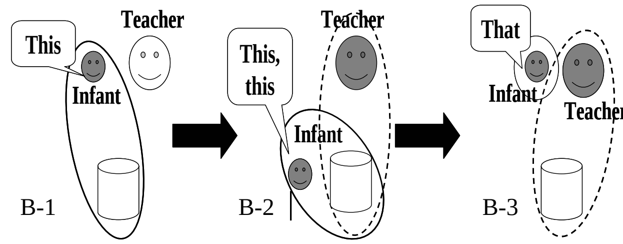

Fig. 7 An example of mental state transition "I => III => V".

As shown in the above example, when the physical distance between a speaker and an object is <distant> but the mental distance between the speaker and the object is <close>, the speaker make an action from a desire to get the object. It is highly possible that the action is created to bridge the gap between the physical distance and the mental distance. It is also considered that the action includes the speaker's intention of making the listener aware of the object to satisfy the speaker's desire. The action transfers the mental states from **Category I** to **Category III**. The grasp of gaps between the physical distance and the mental distance realizes appropriate behavior supports by proposing actions which operate well to bridge the gaps.

On the other hand, it is possible to suppress the desire by the intervention of another person. In Scene **B-3**, for instance, the teacher's insistence suppresses the infant's desire. Such situations in which mental suppressions appear are represented as the transitions from **Category III** to **Category V**. In those situations, the mental distance between a speaker and an object changes from <close> to <distance> by the operation of suppression factors. This action is regarded as an adaptation to present circumstances like an acceptance of a listener's insistence. From an opposite viewpoint, a situation in which a transition from **Category III** to **Category V** is observed has a possibility of including behavior suppression by a certain factor. The consideration of the factor provides a clue to solve the speaker's worries or stresses.

As mentioned above, we have proved that the proposed situation description model enables us to consider the essential mechanism of human behavior.

5.2 The Behavior Prediction Based on the Analysis of Mental States

As considered in Section 5.1, the proposed situation description model has the faculty for explaining the changes of the mental states of speakers and listeners according to the shift of scenes.

It is possible to utilize the described changes of the mental states for the behavior prediction of speakers. For example, when a scene which belongs to **Category I** is detected, a speaker of the scene is predicted to make an action which is classified into **Category III** to bridge a gap between the physical and the mental distances. Additionally, the detection of a **Category III** scene predicts that it is possible for a speaker of the scene to shift the mental state to **Category V** by the suppression of a desire.

Therefore, to accumulate behavior observation results as the change patterns of the mental states enables us to predict a speaker's behavior by the following steps:

(1) For new scenes, describe mental states of actors (speakers) by the proposed model using the behavior history up to that time.

(2) After representing the transitions of mental states as the change patterns among the five categories of the proposed model, extract patterns matched by the prefix search from the accumulated patterns.

(3) For each extracted pattern, a category located next to the end of a matched part is a candidate for the next mental state at the time.

By implementing the above steps, it is possible to predict the next behavior of a speaker from the present circumstances. The incorporation of this framework into the strategy managers of user assistance systems realizes smart and accurate supports according to the situations of users in the domains of not only the elder people supporting, but also the personal secretary system and the conference support system.

6. CONCLUSION

Aiming to represent commonsense knowledge on computing models, we have developed ubiquitous and wearable environments which allow us to observe infant behavior from various perspectives. Infant learning environments have been established using natural materials of the cedar, which enables infants to be absorbed in the activities comfortably and safely.

Environments for infant behavior recording have been constructed; multiple cameras and microphones were set up in the Cedar yurt. A wearable speech recording device was also developed to capture infant utterances with high quality. Developing a comment-collecting system which allows everyone to make comments easily from the multi-viewpoints, a framework for multimodal observation of infant behavior has been realized.

Using the recorded data, we have observed the occurrences of appear expressing in infant utterances. As the result of the observations, we have proposed a situation description model which enables us to classify mental states by means of the mental distance between a speaker or a listener and an object.

Additionally, in order to develop user assistance systems which appropriately support human life according to each situation, we have considered a mechanism which realizes the behavior prediction of speakers by using the proposed model, and accumulating the description of scenes.

7. REFERENCES

[1] Clarkson, B., Mase, K., Pentland, A. *Recognizing user's context from wearable sensors: Baseline system.* International Symposium on Wearable Computers(ISWC) (2000) 69–76.

[2] Yoshitaki, S., Sakane, Y., Takebayashi, Y. *New cycling environment using multimodal knowledge and ad-hoc network.* IEICE Transactions on Inf. and Syst.(Special Issue on Human Communication) E-87-D(6) (2004) 1377–1385.

[3] Sakane, Y., Takahata, M., Otani, N., Kuroki, Y., Takebayashi, Y. *A design for the ubiquitous karate school for learning mind and body interaction.* Pervasive 2004 Advances in Pervasive Computing (2004) 345–350

[4] DK, O. *Metaphonology and infant vocalizations.* In precursors of Early Speech, (1986) 21–35.

[5] Ejiri, K. *Relationship between rhythmic behavior and canonical babbling in infant development.* Phonetica 54 (1998) 226–237

[6] Hsu, H., Fogel, A., Cooper, R.B. *Infant vocal development during the first 6 months: Speech quality and melodic complexity.* Infant and Child Development 9(1) (2000) 1–16.

[7] Pizer, G. *Baby signing as language socialization: The use of visual-gestural signs with hearing infants.* Proc. of the 11th Annual Symposium about Language and Society 47 (2004) 165–171.

[8] Petters, D. *Building agents to understand infant attachment behaviour.* International Joint Conference on Artificial Intelligence 2005 (2005) 158–165.

[9] Otake. K, *A Consideration on Infant Behavior Corpus,* FIT2005, Vol.2, 2005, 129-130 (in Japanese)

[10] Kamio. A, *Territory of Information: Pragmatics and Beyond New Series* (John Benjamins Pub Co, 1997)

[11] Suga. T, *Solving the Riddle of Language Acquisition: How to Design a Virtual Infant (MLAS):*

http://www10.ocn.ne.jp/~mlas/JSLS_Paper.htm

Simulating and Evaluating Public Situated Displays in Virtual Environment Models

Christoph Stahl

Saarland University
P.O. Box 15 11 50
D-66041 Saarbrücken, Germany
+49 681 302 3191

stahl@cs.uni-sb.de

Jens Haupert

Saarland University
P.O. Box 15 11 50
D-66041 Saarbrücken, Germany
+49 681 302 64047

haupert@xantippe.cs.uni-sb.de

ABSTRACT

In this position paper, we address the problems that arise during the design phase of intelligent environments that make use of public situated displays for user assistance. We propose a new design process, which is based on an architectural model of the environment. It employs a toolkit that allows the designer of an intelligent environment to verify the visibility of displays from various virtual viewpoints before they are actually acquired and deployed. Our approach furthermore allows the early evaluation of application prototypes. We mirror the screens of real machines to the virtual displays and use an avatar as a positioning loopback device in order to test the behavior of context-aware presentations.

Categories and Subject Descriptions

H.5.2 [**Information Interfaces and Presentation**]: User Interfaces – *Theory and Methods, User-centered design, Interaction Styles*

D.2.1 [**Software Engineering**]: Requirements/Specifications – *Methodologies*

General Terms

Design, Human Factors

Keywords

Models, Principles, Intelligent Environments, Ubiquitous Computing

1. MOTIVATION

Public situated displays have great potential for user assistance in intelligent environments, and several applications have been published in [1]. They can be used to complement personal mobile devices such as mobile phones and PDAs. These small devices are always available to the user, but their display resources are strictly limited due to the constraints of their form factor. Whereas public displays offer large screen spaces, their visibility is restricted to certain locations. However, this restriction can be beneficial for presentations that are designed for a specific spatial context.

Traditional signage applies their limited visibility to convey context-related information, such as sales offers, opening times, room occupancy, but also local rules and prohibitions. A sign that reads “no smoking” for example is not intended to have a global effect, but only to stop people from smoking inside the room where they can read it. In addition, signs have a fixed orientation with respect to their surrounding environment, so that they are especially useful to assist us in our wayfinding tasks.

Electronic displays, even when they are non-interactive, offer more flexibility and can adapt their content to the situation. In intelligent environments, they can even sense the users and adapt their presentations according to the users’ profiles and cognitive resources. For example, an arrow might dynamically switch its direction according to the goals of each individual user.

However, the design of signs and displays poses difficulties since the designer has to consider the visibility of the displays from various viewpoints, and the interpretation of directions heavily depends on their viewing angle. We all might have experienced traffic situations, where it is unclear whether the sign tells us to turn left or to go ahead. Personalized information is yet more difficult to design, since the presentations are automatically generated and have to be evaluated for a variety of events and situations, such as different user profiles and goals.

Furthermore, testing and evaluating electronic displays comes with high cost for the actual hardware and their installation. It is also clearly unfeasible to try out different positions for a large wall-mounted display, since it is an undesirable and irreversible process to drill holes into concrete and steel.

Hence we propose a new design process, which is based on a virtual model of the environment and employs a toolkit that allows the designer of an intelligent environment to simulate and evaluate public displays from various viewpoints before they are acquired and deployed.

2. THE DESIGN PROCESS

In this position paper, we propose a new approach for the design of user assistance in intelligent environments that is centred on 3D models of the environment. These models represent the architectural features of a place. We make the assumption here that the environment, which is going to be instrumented with an information system for its visitors or inhabitants, already exists (or is at least at the end of the planning phase) and architectural drawings are available to create such a model.

By using our method, the interaction designer will be aware of the spatial context during the whole design process, from the first scenarios to the evaluation of a prototype in a virtual environment. The environment model helps the designer to decide which kind of assistance is required by the user in certain places, and to choose the appropriate interaction devices, such as displays and sensors, based on their visibility and range.

To support this, we have extended the *YAMAMOTO* map modelling toolkit with a new feature that allows the designer to model displays, and we have implemented an egocentric perspective, so that the designer can virtually explore the instrumented environment in order to validate the placement and content of the displays.

The proposed approach is especially beneficial for the design of user assistance applications that include a wayfinding component, which provides navigational aid to the user on public displays. Our method comprises the following steps:

- Modelling architectural features (walls, doors, stairs) and furniture
- Virtually instrumenting the environment and visualizing the placement of displays from an ego perspective
- Prototyping the user assistance application
- Evaluating the assistance application prototype within the virtual environment model
- Deploying the system
- Monitoring the display content in ‘mixed reality’

In the following sections, we describe each step in more detail, before we close the paper with a summary and outlook.

2.1 Modelling and Browsing the Environment

For the modelling of the architectural features of the environment, we use the *YAMAMOTO* map modelling toolkit, which has been published in [3]. The editor allows one to visually mark-up walls and doors based on architectural ground plans through polygon meshes. In order to keep the modelling process as simple as possible, each floor is represented in two-dimensional space only. However, through the use of layers, multiple floors and interconnections such as staircases can also be represented. Complex environments may consist of multiple buildings and outdoor areas, so that a single scale large model may not be appropriate for the design process. Hence the toolkit allows one to recursively refine buildings and rooms through submodels of a smaller scale.

The resulting model can be shown either in a top-down projection, birds-eye view, or egocentric perspective. The necessary 3D geometry for a realistic rendering of the building is automatically created on the fly from the 2D outlines by extending the z-axis through an extrusion operation. Some additional object attributes such as ceiling height and object type (wall, door, stairs) are considered.

In addition to the plain floor layout, the modelling of furniture arrangements or queuing barriers would be clearly desirable for a realistic simulation of environments. Unfortunately our toolkit does not yet provide the means to do this.

2.2 Virtually Instrumenting the Environment

Once the environment has been modelled, the *YAMAMOTO* editor provides an interactive view, which helps the interaction designer to identify the places where users are likely to require and request information from the environment in order to pursue their activities and goals.

The designer can now begin to add ubiquitous and pervasive computing elements into the model, such as sensor devices or positional beacons. These devices are modelled through geometric primitives (such as points, circles and sections), that represent their position, range and orientation. Additionally, a symbolic name and ID can be specified for hardware devices. In the *RENA* project for example, we have used the toolkit to model infrared beacons and RFID-tags, which we use for our indoor positioning system, which estimates the position of a mobile device based on the signals received.

In order to model displays within the environment, we have enhanced the *YAMAMOTO* editor, so that the designer of an instrumented environment can choose the display size, aspect ratio (4:3 or 16:9, portrait or landscape) and orientation (tabletop or wall mounted). Since the crucial point in the setup of public displays is their visibility from different viewpoints, the egocentric perspective provided by *YAMAMOTO* helps the designer to verify that the displays are visible to the user and allows to interactively adjust their position, if necessary. The screenshot in Figure 1 shows two wall-mounted public displays, which have been placed next to a door in a corridor.

The designer can assign static screenshots as mock-up presentations to the modelled displays, which helps to judge their comprehensibility with respect to the spatial context. The semantics of a sign showing an up-arrow for example is often ambiguous – does it point forward or upward to the next level? Nearby an elevator, the sign would have a different interpretation than at the beginning of a long corridor.

The most interesting feature however is the ability to mirror the content of networked displays to the virtual environment in real time, which will be explained in the following section.

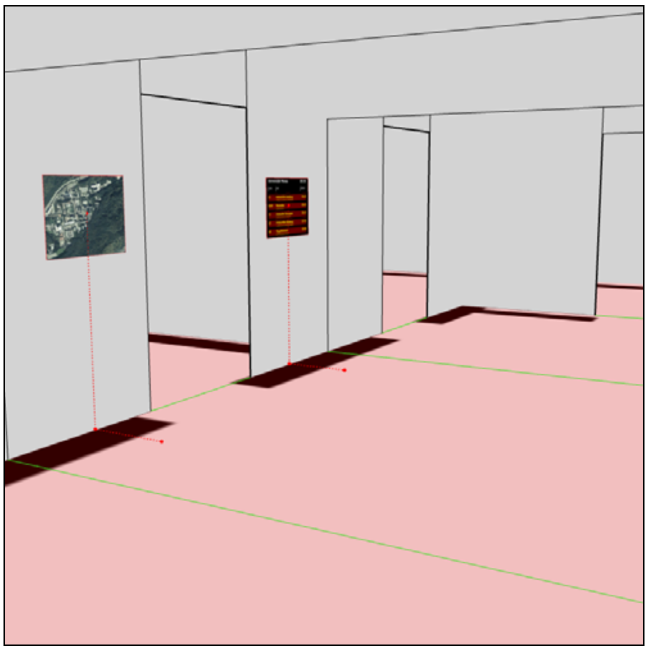

Fig. 1: Modelling public displays with a static mock-up presentation (left) and a VNC real-time capture (right).

2.3 Simulation and Evaluation

In the previous design phases, the designer has created a model of the environment and has virtually instrumented it with ubiquitous computing elements, such as beacons for indoor positioning and public displays. According to the third phase of our proposed design approach, the designer would now develop a first prototype of a user assistance system that provides location-based presentations. However, setting up the indoor positioning infrastructure and public displays in the real environment takes a lot of cost and effort, and is sometimes not even possible at all. In public areas, such as shops or airports, no experimental setup would be tolerated by the management, and the developer has to simulate the whole system somehow before it can actually be installed.

Here the modelling effort of the first two phases begins to pay off, since the developer gets the necessary virtual testing environment for free. Figure 2 shows the evaluation setup, where the real instrumented environment (IE) is shown on the left hand side, and the virtual environment (VR) is depicted on the right. The instrumented environment encloses the user and comprises display devices, distributed computing power for the application logic, and navigational beacons. According to the setup in our RENA project, the user holds a PDA which receives the beacon signals and estimates its position. The position of the user is sent to the application, which sends corresponding presentations to the public displays.

In the virtual environment, the positioning system is simulated through the avatar, which acts as a kind of loopback device. As the designer moves the avatar, location events are sent to the application, to which the whole simulation process is completely transparent. The presentations are rendered on real machines and mirrored to the virtual displays. We have therefore integrated a *VNC* (Virtual Network Computing) [2] client into the *YAMAMOTO* rendering process, which achieves a display update rate of approximately one frame per second, and the virtual environment is rendered with approximately 20 frames per second.

The benefit of this testing arrangement is that both the designer and the system developer can actually see the real outcome of the application logic, which runs in a fully distributed computing environment as is typical for intelligent environments. This approach should ensure that the system will perform well according to the designer's intention once it is deployed in the real environment. Additionally, a first study with real users could be easily conducted in this lab setting.

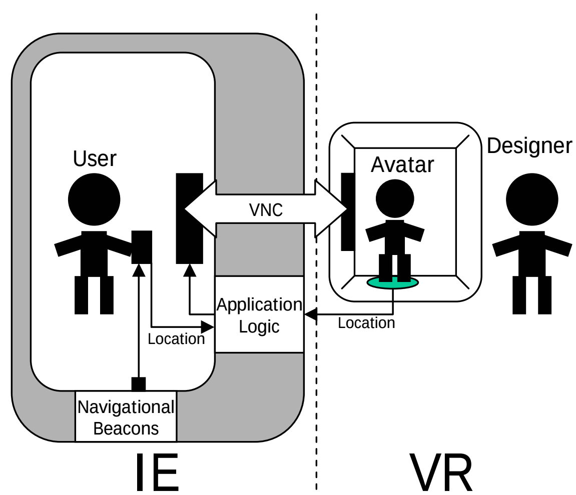

Fig. 2: The VR-Environment provides a positioning loop-back device to evaluate location-aware assistance services.

2.4 Monitoring Real Display Environments

The simulation setting described above can be useful even beyond the evaluation phase. For the administration staff, which is in charge of feeding the installed application with location-based data, such as room occupancy or route knowledge, it might be cumbersome to check the content of the actual displays, which are distributed all around the environment. The virtual environment would provide an intuitive interface for them to monitor and validate the presentations in real time, conveniently from their office. Viewing the display content from an egocentric perspective would help to identify wrong presentations (especially misleading wayfinding directions) much easier than in a simple collection of plain windows without their spatial context.

3. SUMMARY AND OUTLOOK

We have outlined an approach for the design of intelligent environments, which is based on a geometric 3D model of the environment. The model allows the designer to virtually instrument the environment with sensor devices and public displays and can be used to evaluate application prototypes. We mirror the screens of real machines to the virtual displays and use an avatar as a positioning loopback device for location-aware user assistance applications.

The necessary enhancements of the YAMAMOTO toolkit are currently under development, so we could not yet apply the proposed design process to a real-life problem. However, we are planning to do so in the near future. We will also extend the capabilities of the editor so that furniture arrangements can be modeled and the virtual testing environments become more realistic.

4. SHORT AUTHOR BIOGRAPHIES

Jens Haupert is a student of Computer Sciences at Saarland University and has implemented the first version of the *YAMAMOTO* editor in a practical course on software engineering in 2003. Since then he has been working in the Artificial Intelligence Group of Prof. Dr. Wolfgang Wahlster as a research assistant and received his Bachelor degree in October 2005 for his thesis on the hierarchical geo-referenced modeling of environments with the *YAMAMOTO* toolkit. He is currently writing his Master thesis, which includes implementing the features that are described in this paper.

Christoph Stahl has received his diploma in computer science in 2001, and has been working since then as a Ph.D. student in the projects REAL and RENA. There he has implemented several mobile and kiosk-based pedestrian navigation systems and set up the architecture of an instrumented environment. It currently assists its users in their wayfinding and shopping tasks. Currently he is working on a methodology to design and develop user assistance systems that integrate support for multiple activities and make intelligent use of Ubicomp technologies and environments.

5. REFERENCES

[1] O'Hara, K., Perry, M., Churchill E. and Russel, D. (2003). Public and Situated Displays: Social and interactional aspects of situated display technologies, ISBN 1-4020-1677-8, Kluwer, The Netherlands, 2003.

[2] Richardson, T. (2005). The RFB Protocol, Version 3.8, RealVNC Ltd., Website: http://www.realvnc.com/docs/rfbproto.pdf

[3] Stahl, C. and Haupert, J. (2006). Taking Location Modelling to new Levels: A Map Modelling Toolkit for Intelligent Environments. 2nd International Workshop on Location- and Context-Awareness. In: M. Hazas, J. Krumm, and T. Strang (Eds.): LoCA 2006, LNCS 3987, pp. 74 – 85, 2006. Springer-Verlag Berlin Heidelberg.

Rapid prototyping and evaluation of intelligent environments using immersive video

Pushpendra Singh, Hai Nam Ha, Patrick Olivier, Christian Kray, Zhiwen Kuang
Amy Weihong Guo, Phil Blythe, Phil James
Newcastle University
Newcastle upon Tyne, United Kingdom
p.l.olivier@ncl.ac.uk

ABSTRACT
A major problem in developing ubiquitous computing applications is the simulation of the required infrastructure and the environment in which they are to be deployed. Developers have to rely on either low-fidelity techniques (such as paper prototypes and mental walk-throughs) or simply wait for a full scale deployment. Similarly, even where an existing infrastructure exists preliminary evaluations of such applications is problematic due both to the logistics involved in coordinating subjects and experimenters on site, and the uniqueness of each interaction (it is impossible to reliably repeat specific visual, spatial and sensory contexts in an actual deployment). We propose an approach using immersive video with surround sound and a simulated infrastructure to create a realistic simulation of a ubiquitous environment in the software design and development office. We describe the motivations and design of our immersive video system and discuss our initial experiences of how it can be used to aid developers in the requirements capture, design, development and deployment of ubiquitous computing applications and intelligent environments.

General Terms
Human Factors, Experimentation

Keywords
Immersive video, rapid prototyping, evaluation, modeling.

1. BACKGROUND AND MOTIVATION
Development and implementation of ubiquitous computing applications require rapid prototyping and early evaluation. However, the infrastructure to support such applications is often not available at the time of application development and developers have to wait for a full-scale deployment of sensors and other components before their prototypes can be deployed and evaluated. In addition, interaction between a ubiquitous computing system and its users takes place in a specific context, which may not only change frequently but also impact how users perceive the system and how they interact with it. Oftentimes, it is very difficult to assess in advance which contextual factors are relevant and how they influence interaction. This frequently requires full deployment and/or a redesign of a ubiquitous system in order to account for specific contextual configurations or situations.

Research has shown that prototyping systems for studies is important to understand how a system is fit for purpose and will in practice be utilized. The degree to which an ubiquitous computing application can be useful depends on the ability of designer to understand and anticipate a user's needs. Prototyping helps in understanding these needs. The take-up of ubiquitous computing applications in everyday life is therefore dependent on the development of prototyping techniques that can provide useful insights into likely usage patterns.

Many ubiquitous computing applications depend upon the nature of a spatial interaction between a user and a physical environment, both in terms of a user's interest in the elements of the physical environment (e.g. landmarks in a mobile navigation application) and interactions between the user and sensors deployed either on the user or on objects in the environment (e.g. components of a positional tracking systems). To capture such spatial issues virtual reality models of a scenario have been proposed. Such systems utilize existing game engines in which a replica of a physical environment is constructed and augment this environment with virtual sensors.

The few reported deployments of virtual reality for prototyping ubiquitous applications have used game-like desktop simulations as high fidelity immersive visualisation of such models require large stereoscopic displays, such as in a CAVE, which are not widely available[7, 1]. However, although such virtual reality models are capable of capturing the spatial characteristics of a physical environment they inevitably lack the visual quality of real world. Furthermore, the process of creating a 3D model of a real scenario is well known to be a very time-consuming and costly process, and even the most sophisticated 3D designer is incapable of recreating real world experiences and unexpected events of environments such a shopping centers and airports.

Even where there is an underlying sensors and wireless network infrastructure many problems arise in relation to the evaluation of ubiquitous computing applications at the site of an actual deployment. Such processes are time-consuming and particular events and activities might only be intermittently available (e.g. such as the arrival of a particular plane or bus in intelligent transport systems). Evaluations of actual deployments typically also limit the experimenter's ability to elicit continuous feedback from users during the evaluation and preclude the possibility of recreating the same situation repeatedly to test problems that arise as a results of particular transient configurations of the user, world and infrastructure.

Thus, there is a pressing need for a tool which can facilitate both rapid prototyping and evaluation of computing ubiquitous applications within the software development office. Existing techniques, such as paper prototyping [5], suffer from their inherently low fidelity (in relation to both visual, physical and sensory phenomena) when applied to mobile and ubiquitous systems. These contextual factors (such as crowded environments, visual and aural impressions of a location) are hard to recreate or convey using standard methods. One possible solution is to use a Wizard of Oz methodology whereby a user's personal device (typically a mobile phone or other handheld computer) is controlled externally by an experimenter (the wizard) to simulate its planned behavior [4]. A number of problems remain for such WOz-based approaches: they still requires participants/experimenters to physically visit the place where the system is to be used; subsequent trials will not take place under exactly the same circumstances (e.g. it might be raining on one occasion, while it might be sunny another time, crowd densities will vary, etc.); significant effort may be required to build the control mechanism used by the Wizard; some aspects of the system (such as sensors/actuators embedded in the environment) may be very difficult to simulate.

2. THE IMMERSIVE VIDEO APPROACH

We propose the use of *immersive video*, to overcome limitations of applying traditional prototyping and early evaluation approaches to mobile and ubiquitous systems and to capture the sensory experiences that we expect users to be exposed to at locations of deployment of location-based services. By capturing video (imagery and sound) at the site of the intended deployment of a location-based service and simulating the sensor infrastructure, system developers can have ready access (i.e. in their development office) to a high fidelity recreation of a user's experience of using a prototype mobile system on location.

2.1 Simulating sensory experience

Immersive video systems have been proposed as an alternative form of entertainment multimedia (e. g. [6]), with the goal of extending the traditional constraints of film to provide the user with a sense that they are in some way truly immersed in an environment. Our system harnesses these techniques using one or more video cameras to capture synchronized footage at key locations within an intended usage scenario. Depending on the display facilities available to the designers, a wide field-of- view of the scene is captured (typically between 140-360°) and then replayed on multiple screens in front of (or within) which the user stands.

Footage captured at each location should be edited to create shots that can be looped to create a sense of being at a point in space and time, without the sensory disturbance that can result when observing a jump from the end of a shot back to the beginning (both in terms of the image and the sound). A *base shot* at each location is first created and a number of *event shots*, relevant to the location, may be added depending on the nature of the application under development. Depending on the field-of-view that has been captured, multiple base shots for a location may be created corresponding to a user standing in different orientations. Event shots correspond to activity that is typically salient to the application under development, for example, the arrival or departure of a bus in an intelligent transport system, or a change in the state of a pedestrian crossing signal in a mobile guide application.

2.2 Simulating Events

Base shot and event shots are organized into a simple state-based model of the environment. Two-way links are created between the base shots at a particular location, as transitions between these states corresponds to users reorienting themselves. One way links are formed between one or more of the base shots corresponding to spatially adjacent locations (according to the direction in which a user is allowed to move in the scenario that has been recreated). Finally, two way links are added between the base shots at a location and their associated event shots. A user's location and orientation in space, and their location in time (i.e. in relation to the events that may occur) thus correspond to their location in the state space of shots.

Physical movement (either reorientation or translation) is simulated by transitions between base states, and the passage of time at a location (either real or otherwise) by transitions from base shots to event shots (and back). These transitions, which correspond to the passage of a user in space or time, are not performed by the user but by a separate Wizard of Oz-style controller who drives that immersive video system in response to the user's verbal requests. Figure 2 shows state diagram for our application.

In our current system we have deliberately chosen this configuration in preference to augmenting the mobile device carried by the user with additional controls (which would do more to undermine the integrity of the users interaction with the device), or any hands free interaction technique, which would most likely be hard to deploy in a typical environment where software development is undertaken. Transitions between shots are realized using standard techniques from cinematography, such as cross-fading the image and sound of the source and destination shots, a visual style that is instantly recognizable, and readable, for users.

2.3 Simulating the technology infrastructure

Simulating the user's sensory experience of an environment is one component of our mobile services prototyping and evaluation system. The remaining component aims to simulate the sensor data corresponding to the base and event shots. During the creation of the state space of shots, XML-based sensor files are created for each shot specifying the values of different sensor readings that will occur at the corresponding time and location. This sensor abstraction layer allows the specification and simulation of GPS, Bluetooth, RFID and infrared inputs for each shot. The abstraction layer renders the origin of sensor data completely transparent to an application thereby allowing users to stand within the immersive video system and experience the behavior of an application exactly as if they were in such a context in a real sensor infrastructure. Figure 1 shows a fraction of a sample sensor file, with readings for GPS and related states that can be reached.

```
<?xml version="1.0" standalone="no" ?>
<ivideo>
<sensor stype="gps" value="" irid=""
   longitude="54.97750833333333"
   latitude="1.61345555555556"
   startframe="0" endframe="3000" />
<state name="station2_4" astartframe="0"
   aendframe="3000"
   centersrcfile="../AVI/c24z1.avi"
   leftsrcfile="../AVI/l24z1.avi"
   rightsrcfile="../AVI/r24z1.avi"
   xmlfile="../XML/s24.xml"
   istartframe="" iendframe="" />
<state name="station1_1" astartframe="0"
   ...
   ...
</ivideo>
```

Figure 1: XML file for one state showing sensor values and other possible states reachable from this state.

Crucially, from the user perspective, an application, and its user interface, will behave exactly as if they were in the actual location in a real-world deployment. The video and audio provide a rich impression of the actual space, and the large field of view heightens their sense of immersion as they interact with a device. As discussed above, control of a user's location in the state space is maintained by a wizard, a task that is facilitated by only having the allowable states available. We are currently using this approach to ease the evaluation and prototyping of mobile and ambient applications while enabling the discovery of context-related issues. Our experience has been that much of the fidelity of a user's real-world experience can be captured using the immersive video system and interface and application developers can very quickly generate prototypes and evaluate them using the immersive video.

2.4 Immersive video & requirements capture

An immersive video simulation of a mobile application scenario contributes to the development process in a number of ways [6]. Our primary goal was to create a mechanism by which the evaluation of such a system can be conducted both as early as possible in the development life-cycle and within a sensory context which the user will typically encounter when using the application. The use of immersive video addresses a number of additional problems in the development of mobile and ambient systems. Firstly, requirements capture for such systems is known to be significantly more difficult than for traditional static applications. By definition mobile systems will be used in a wide range of environments and the users of such systems will often be involved in multiple tasks simultaneously.

Capturing this richness of deployment is problematic using standard text or graphical formalisms and notations for characterizing use cases. Even seemingly rich approaches such as scenarios [2] and personas [3] fail to capture much of the richness of a real world deployment. In the light of this shortfall in the description of the requirements of mobile systems designers and developers struggle to understand the full implications of their design choices, and it is almost inevitable that major (and costly) interface design revisions will occur late in the process.

The process of developing an immersive video simulation goes some way to mediating this shortfall in the expressiveness of traditional approaches to requirements capture. On one hand the selection of the usage scenario, and the specific locations at which the filming is undertaken comprises an informal statement as to the anticipated context in which the system will be used, the range of relevant and incidental events that occur at different locations, and the data (i.e. sensor data) available to the mobile device. On the other hand the footage itself explicitly captures features of the environment that are impossible to include in a traditional scenario or use case description, such as the level and nature of ambient sound, the degree of clutter and motion in the user's visual field, and the precise spatial characteristics of important features in an environment.

3. INITIAL EXPERIENCES

We are currently using the immersive video approach as a prototyping and evaluation tool in the context of three projects, each of which have specific requirements and properties to be taken into account. In this section we will report on initial experiences we gained with one system: a prototype built within ASK-IT, a European Union Framework 6 project that aims to enhance inclusion through the application of ambient intelligence. One goal of ASK-IT is to provide mobility impaired users with seamless location sensitive guidance and information as they arrive at a metro station, exit and navigate to a cultural location such as a museum, and enter the museum and view its exhibits. The user interface is primarily map-based but also provides textual information on exhibits in a museum. Having an evaluation tool with which views can be elicited without exposing users to a real environment is particularly useful for such a vulnerable user group.

For ASK-IT we have prototyped a GPS-enabled PDA, Bluetooth indoor positioning, RFID e-ticketing and infrared beacons for pointing gestures. The actual application works in following manner: when a user arrives at metro station the PDA detects the platform using Bluetooth positional sensing and displays a map of the platform with directions to the escalator and lift. As the user exits the metro station a GPS signal is detected and a map of surrounding area is displayed with direction information towards the museum. A panoramic photograph (covering a 360° view) at the top of the display is used to help the user identify the direction he or she needs to walk.

When user enters into the museum, the PDA detects the new

Figure 4: Immersive video in action: a user interacting with a mobile device while being surrounded by a video replay of the location, where it is intended to be used

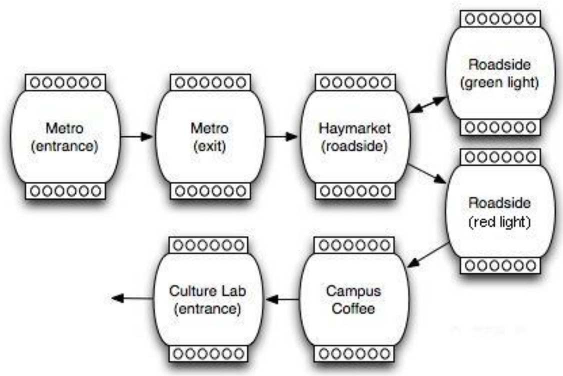

Figure 2: Fragment of the state space for the immersive video prototype of the museum application.

location (using Bluetooth positional sensing at the entrance to the museum) and displays a map of museum. The user is also issued with an e-ticket for the museum (using RFID tag attached to PDA) and information pertaining to the e-ticker is displayed on PDA. Within the museum, when user points the PDA at an exhibit, a supplementary multimedia presentation relating to that exhibit is displayed on the device. The detection of which exhibit the user is pointing the device at is implemented using infrared beacons. Figure 3 shows an example of the actual user interface displayed on the PDA. Immersive video of the environment was used in both the development and initial evaluation of the applications. Figure 2 show a fragment of the shot state space used.

During development, it reduced considerably the need to visit the test site which would have been both time-consuming and also required the installation of infrastructure elements, such internal Bluetooth devices at the platform of the metro station and in the museum. This gave developers considerable freedom to develop their application without committing either to the class of infrastructure devices to be used or their detailed configuration on site. Possible users were also able to evaluate the system at an early stage and highlighted some useful interface design issues (such as interaction with the panoramic images). Finally, with the aid of the demo, participants were able to gain first hand experience of navigating the simulation of the space and the behavior of the interface. Figure 4 shows our local set-up using a three walled CAVE-based display. However, a similar impact can be achieved by projecting video on walls of a corner of a room.

4. FUTURE WORK

In future, we plan to use immersive video as a standard prototyping tool other ubiquitous computing based projects. Two major projects, on which we are current working are (i) real-time 3D visualization of urban data for emergency response and (ii) the assessment of future traveler information systems. The emergency response scenarios project involves the development of a mobile device that provides real-time 3D visualizations of urban data for emergency response sce-

Figure 3: User interface on mobile device: interactive panoramic picture (top), map with route (middle), and menu-bar (bottom)

narios. Requirements capture for the application include working closely with police, fire and other emergency services. This means that application development needs to include input from different user groups. Also, access to information related to emergency facilities such as hospitals is required and then evaluation of such an application will be difficult due to its nature of emergency response scenarios.

Here the immersive video offers significant benefits in overcoming these problems. Limited access to important buildings such as hospitals, police or fire headquarters can be overcome by videoing them once and annotating it with all the information that will be available. Immersive video will also allow easy evaluation of application by different user groups because they can evaluate it in their office-space and can provide input. Video of real environment will provide easy means of assessing a user's ability to visually register the 3D graphics with the environment without having to be on-site.

Our future traveler information systems project is a response to the observation that increased car use in many urban areas has led to growing traffic congestion, which not only threatens economic growth but also results in poor air quality, noise and global warming. A range of marketing and management measures have been piloted in various areas in the UK including travel awareness campaigns, workplace and school travel plans, personalized travel planning and public transport information. These measures usually seek to offer better information and opportunities to influence the choice of transport mode made by individuals, and reduce the level of car use. These measures can have a significant effect on individual travel choices when individuals or particular groups are targeted and tailored information is provided based on individual journey characteristics.

The key problem relating to the provision of better information is that the locations and contexts that travelers find themselves in are constantly changing, thus their information needs are both immediate and highly context sensitive. However, a large-scale implementation that collects a full set of travel-related information in real-time, and disseminates such information in a context aware manner, is not possible within the existing information systems infrastructure. A future traveler information system (FTIS) will have relevant environmental information, live transport and traveler information, and will compile them proactively with stored historical information. Such a system will therefore always offer customized and timely travel-related information based on traveler context awareness. Such a scenario raises two important questions:

- How will people perceive such proposals for future traveler information systems?
- Does the proposed future traveler information system have the potential to bring about a significant modal shift?

The ideal way to explore the behavioral change that results from FTIS use is to allow users to interact with such systems as part of their everyday travel routine, such as commuting, shopping and recreation. However, FTIS supported by pervasive computing is still very much an aspiration as large scale deployments of pervasive computing environments do not exist. Building a physical mock-up of FTIS is expensive and time-consuming, it will also cover a large geographical area (usually a city) which will make evaluation of such a system cumbersome.

Immersive video will be very helpful in creating a realistic sense of a user's experience of an application and allow participants to interact with a simulated system. It will help in development and evaluations of applications related with FTIS. It will also be able to stimulate the experience of interacting with an FTIS by providing participants with a sense that they are in some way truly immersed in an environment where the FTIS has been implemented.

The utility of immersive video can be clearly visualized with respect to the projects mentioned above. For the development of immersive video as a standard prototyping tool, we aim to develop a range of applications to make it easier to build an immersive video environment after taking the videos. Two major components of immersive video system are:

1. Wizard: the wizard controls the sequence of video clips based on the state diagram as explained in section 2.2. Arranging clips into right sequence order and then feeding them to wizard is quite time-consuming.

2. Sensor information: we are using XML files to store the sensor information about every clip. Though this approach is general it is cumbersome to create individual XML files and make changes in them.

Furthermore a number of interaction modes within the immersive video system need to be explored:

1. User interaction:
   - How to move backward or forward in a scene (without intervention from the Wizard)?
   - How to select objects available in a scene (for enhanced interactivity)?
   - Physical interaction techniques, such as opening a door?
2. System interaction: presently only the wizard injects events according to its observation of the user's behaviour. But events can be generated based on the level and nature of interaction that a user is engaged in. Events are also not only the result of a user triggering sensors but a result of external events too, such as plane arriving, receiving a call.

5. CONCLUSION

A key issue in mobile and ambient computing is the effort required to rapidly prototype and evaluate user interfaces and applications. Existing technologies for these tasks suffer either from low fidelity (e. g. paper prototypes, mental walkthroughs) or effectively require a near full-scale deployment. In this paper, we proposed our approach using immersive video and a simulated infrastructure to create a realistic environment in the office or lab. Immersive video provides a low-cost and rapid means to prototype user interfaces and applications, and to evaluate them in a realistic simulation of the context, in which they are intended to be used. Our approach also allows for user studies in the lab, where participants can be easily exposed to specific environment conditions in a realistic and repeatable manner. We described our inital experiences and future plans of using immersive video for a range of problems and at the same time developing tools to make immersive video an easy-to-use rapid prototyping tool for ubiquitous computing based applications.

6. REFERENCES

[1] J. J. Barton and V. Vijayaraghavan. Ubiwise, a ubiquitous wireless infrastructure simulation environment. In *HPL-2002-303*. HP Laboratories Palo Alto, 2002.

[2] J. Carroll. *Making Use: Scenario-Based Design of Human-Computer Interactions*. MIT Press, 2000.

[3] A. Cooper. *The Inmates are Running the Asylum*. SAMS/Macmillan, 1999.

[4] M. T. J. H. Kaj Makela, Esa-Pekka Salonen and R. Raisamo. Conducting a wizard of oz experiment on a ubiquitous computing system doorman. In *Proceedings of the International Workshop on Information Presentation and Natural Multimodal Dialogue*, pages 115–119, 2001.

[5] L. Liu and P. Khooshabeh. Paper or interactive?: A study of prototyping techniques for ubiquitous computing environments. In *CHI '03: CHI '03 extended abstracts on Human factors in computing systems*, pages 1030–1031, New York, NY, USA, 2003. ACM Press.

[6] S. Moezzi, A. Katkere, D. Y. Kuramura, and R. Jain. Immersive video. *vrais*, 00:17, 1996.

[7] M. E. Papka and R. Stevens. Ubiworld: An environment integrating virtual reality, supercomputing and design. In *HPDC '96: Proceedings of the High Performance Distributed Computing*, page 306, Washington, DC, USA, 1996. IEEE Computer Society.

Pushpendra Singh

I completed Master of Technology in Computer Science at the Indian Institute of Technology, Delhi, India and Ph.D. at INRIA Research Lab, Rennes, France. My Ph.D. was focused on developing a fault tolerance framework for applications running on mobile devices. I then developed a mobile device based adaptive speech remediation software for Japanese English speakers as a postdoc in the Department of Psychology at the University of Portsmouth, U.K. After that I spent some time at Nokia, Japan working on multimedia and video quality evaluation.

I joined Newcastle University in December 2005 as a research associate. My current research interests include ubiquitous computing and 3D graphics and visualization. Presently I am working on real-time visualization of 3D Data on mobile devices and developing prototyping and evaluation tools for ubiquitous computing applications.

List of Publications

(# ::= out of print)

electronic copies can be found at:

http://w5.cs.uni-sb.de/Publist/ or ftp://ftp.cs.uni-sb.de/pub/papers/SFB378/

copies can be ordered from:

Doris Borchers, Universität des Saarlandes, FR 6.2: Department of Computer Science, Postfach 151 150, Im Stadtwald 15, D-66041 Saarbrücken, Fed. Rep. of Germany, e-mail: borchers@cs.uni-sb.de

Reports

B1 Schirra, J., Brach, U., Wahlster, W., Woll, W.: *WILIE — Ein wissensbasiertes Literaturerfassungssystem.* FB Informatik, KI-Labor, Bericht Nr. 1, April 1985. In: Endres-Niggemeyer, B., Krause, J. (eds.): Sprachverarbeitung in Information und Dokumentation. Heidelberg: Springer, 1985, 101–112. #

B2 Arz, J.: *TRICON — Ein System für geometrische Konstruktionen mit natürlichsprachlicher Eingabe.* FB Informatik, KI-Labor, Bericht Nr. 2, Mai 1985. In: Endres-Niggemeyer, B., Krause, J. (eds.): Sprachverarbeitung in Information und Dokumentation. Heidelberg: Springer, 1985, 113–123.

B3 Wahlster, W., Kobsa, A.: *Dialog-Based User Models.* SFB 314 (XTRA), Bericht Nr. 3, Februar 1986. In: Ferrari, G. (ed.): Proceedings of the IEEE 74 (7). July 1986 (Special Issue On Language Processing), 984–960. #

B4 Fendler, M., Wichlacz, R.: *SYCON — Ein Rahmensystem zur Constraint-Propagierung auf Netzwerken von beliebigen symbolischen Constraints.* FB Informatik, KI-Labor, Bericht Nr. 4, November 1985. In: Stoyan, H. (ed.): GWAI-85. 9th German Workshop on Artificial Intelligence. Proceedings. Heidelberg: Springer 1985, 36–45.

B5 Kemke, C.: *Entwurf eines aktiven, wissensbasierten Hilfesystems für SINIX.* FB Informatik, KI-Labor (SC-Projekt), Bericht Nr. 5, Dezember 1985. Erweiterte Fassung von: SC — Ein intelligentes Hilfesystem für SINIX. In: LDV-Forum. Nr. 2, Dezember 1985, 43–60. #

B6 Wahlster, W.: *The Role of Natural Language in Advanced Knowledge-Based Systems.* SFB 314 (XTRA), Bericht Nr. 6, Januar 1986. In: Winter, H. (ed.): Artificial Intelligence and Man-Machine Systems. Heidelberg: Springer, 1986, 62–83.

B7 Kobsa, A., Allgayer, J., Reddig, C., Reithinger, N., Schmauks, D., Harbusch, K., Wahlster, W.: *Combining Deictic Gestures and Natural Language for Referent Identification.* SFB 314 (XTRA), Bericht Nr. 7, April 1986. In: COLING '86. 11th International Conference on Computational Linguistics. Proceedings. Bonn 1986, 356–361.

B8 Allgayer, J.: *Eine Graphikkomponente zur Integration von Zeigehandlungen in natürlichsprachliche KI-Systeme.* SFB 314 (XTRA), Bericht Nr. 8, Mai 1986. In: GI – 16. Jahrestagung. Proceedings, Bd. 1. Berlin 1986, 284–298.

B9 André, E., Bosch, G., Herzog, G., Rist, T.: *Coping with the Intrinsic and Deictic Uses of Spatial Prepositions.* SFB 314 (VITRA), Bericht Nr. 9, Juli 1986. In: Jorrand, Ph., Sgurev, V. (eds.): Artificial Intelligence II. Proceedings of AIMSA-86. Amsterdam: North-Holland, 1987, 375–382. #

B10 Schmauks, D.: *Form und Funktion von Zeigegesten, Ein interdisziplinärer Überblick.* SFB 314 (XTRA), Bericht Nr. 10, Oktober 1986.

B11 Kemke, C.: *The SINIX Consultant — Requirements, Design, and Implementation of an Intelligent Help System for a UNIX Derivative.* FB Informatik, KI-Labor (SC-Projekt), Bericht Nr. 11, Oktober 1986. In: User Interfaces. Proceedings of the International Conference of the Gottlieb Duttweiler Institut. Rüschlikon/Zürich 1986.

B12 Allgayer, J., Reddig, C.: *Processing Descriptions containing Words and Gestures.* SFB 314 (XTRA), Bericht Nr. 12, Oktober 1986. In: Rollinger, C.-R., Horn, W. (eds.): GWAI-86 und 2. Österreichische Artificial-Intelligence-Tagung. Proceedings. Berlin/Heidelberg: Springer, 1986, 119–130. #

B13 Reithinger, N.: *Generating Referring Expressions and Pointing Gestures.* SFB 314 (XTRA), Bericht Nr. 13, November 1986. In: Kempen, G. (ed.): Natural Language Generation. Dordrecht: Nijhoff, 1987, 71–81. #

B14 Jansen-Winkeln, R. M.: *LEGAS — Inductive Learning of Grammatical Structures.* FB Informatik, KI-Labor, Bericht Nr. 14, November 1986. In: Hallam, J., Mellish, C. (eds.): Advances in Artificial Intelligence. Proceedings of the AISB Conference. Chichester: Wiley, 1987, 169–181. #

B15 Werner, M.: *RMSAI — Ein Reason Maintenance System für approximative Inferenzen.* FB Informatik, KI-Labor, Bericht Nr. 15, Dezember 1986. In: Stoyan, H. (ed.): Begründungsverwaltung. Proceedings. Berlin etc.: Springer, 1988, 86–110. #

B16 Schmauks, D.: *Natural and Simulated Pointing — An Interdisciplinary Survey.* SFB 314 (XTRA), Bericht Nr. 16, Dezember 1986. In: 3rd European ACL Conference, Kopenhagen, Danmark 1987. Proceedings. 179–185.

B17 Zimmermann, G., Sung, C. K., Bosch, G., Schirra, J.R.J.: *From Image Sequences to Natural Language: Description of Moving Objects.* SFB 314 (VITRA), Bericht Nr.17, Januar 1987.#

B18 André, E., Rist, T., Herzog,: G. *Generierung natürlichsprachlicher Äußerungen zur simultanen Beschreibung von zeitveränderlichen Szenen.* SFB 314 (VITRA), Bericht Nr. 18, April 1987. In: Morik, K. (ed.): GWAI-87. 11th German Workshop on Artificial Intelligence. Proceedings. Berlin/Heidelberg: Springer, 1987, 330–338. #

B19 Rist, T., Herzog, G., André, E.: *Ereignismodellierung zur inkrementellen High-level Bildfolgenanalyse.* SFB 314 (VITRA), Bericht Nr. 19, April 1987. In: Buchberger, E., Retti, J. (eds.): 3. Österreichische Artificial-Intelligence-Tagung. Proceedings. Berlin/Heidelberg: Springer, 1987, 1–11. #

B20 Beiche, H.-P.: *LST-1. Ein wissensbasiertes System zur Durchführung und Berechnung des Lohnsteuerjahresausgleichs.* SFB 314 (XTRA), Bericht Nr. 20, April 1987. In: Buchberger, E., Retti, J. (eds.): 3. Österreichische Artificial-Intelligence-Tagung. Proceedings. Berlin/Heidelberg: Springer, 1987, 92–103. #

B21 Hecking, M.: *How to Use Plan Recognition to Improve the Abilities of the Intelligent Help System SINIX Consultant.* FB Informatik, KI-Labor (SC-Projekt), Bericht Nr. 21, September 1987. In: Bullinger, H.-J., B. Shackel: Human–Computer Interaction — INTERACT 87. Proceedings of the 2nd IFIP Conference on Human–Computer Interaction. Amsterdam: North Holland, 1987, 657–662.

B22 Kemke, C.: *Representation of Domain Knowledge in an Intelligent Help System.* FB Informatik, KI-Labor (SC-Projekt), Bericht Nr. 22, September 1987. In: Bullinger, H.-J., B. Shackel: Human–Computer Interaction — INTERACT 87. Proceedings of the 2nd IFIP Conference on Human–Computer Interaction. Amsterdam: North Holland, 1987, 215–220. #

B23 Reithinger, N.: *Ein erster Blick auf POPEL. Wie wird was gesagt?* SFB 314 (XTRA), Bericht Nr. 23, Oktober 1987. In: Morik, K. (ed.): GWAI-87. 11th German Workshop on Artificial Intelligence. Proceedings. Berlin/Heidelberg: Springer 1987, 315–319.

B24 Kemke, C.: *Modelling Neural Networks by Means of Networks of Finite Automata.* FB Informatik, KI-Labor, Bericht Nr. 24, Oktober 1987. In: IEEE, First International Conference on Neural Networks, San Diego, USA 1987. Proceedings.

B25 Wahlster, W.: *Ein Wort sagt mehr als 1000 Bilder. Zur automatischen Verbalisierung der Ergebnisse von Bildfolgenanalysesystemen.* SFB 314 (VITRA), Bericht Nr. 25, Dezember 1987. In: Annales. Forschungsmagazin der Universität des Saarlandes, 1.1, 1987, S.82–93.
Wahlster, W.: *One Word Says More Than a Thousand Pictures. On the Automatic Verbalization of the Results of Image Sequence Analysis Systems.* SFB 314 (VITRA), Bericht Nr. 25, Februar 1988. In: T.A. Informations, Special Issue: Linguistique et Informatique en République Féderale Allemande, September 1988.

B26 Schirra, J.R.J., Bosch, G., Sung, C.K., Zimmermann, G.: *From Image Sequences to Natural Language: A First Step towards Automatic Perception and Description of Motions.* SFB 314 (VITRA), Bericht Nr. 26, Dezember 1987. In: Applied Artificial Intelligence, 1, 1987, 287–305. #

B27 Kobsa, A., Wahlster, W. (eds.): *The Relationship between User Models and Discourse Models: Two Position Papers.* SFB 314 (XTRA), Bericht Nr. 27, Dezember 1987. Both papers of this report appear in: Computational Linguistics 14(3), Special Issue on User Modeling (Kobsa, A., Wahlster, W. (eds.)), 1988, 91–94 (Kobsa) and 101–103 (Wahlster).

B28 Kobsa, A.: *A Taxonomy of Beliefs and Goals for User Models in Dialog Systems.* SFB 314 (XTRA), Bericht Nr. 28, Dezember 1987. In: Kobsa, A., Wahlster, W. (eds.): User models in dialog systems. Berlin etc.: Springer, 1988, 52–68.

B29 Schmauks, D., Reithinger, N.: *Generating Multimodal Output — Conditions, Advantages and Problems.* SFB 314 (XTRA), Bericht Nr. 29, Januar 1988. In: COLING-88, Budapest 1988. Proceedings. 584–588.

B30 Wahlster, W., Kobsa, A.: *User Models in Dialog Systems.* SFB 314 (XTRA), Bericht Nr. 30, Januar 1988. In: Kobsa, A., Wahlster, W. (eds.): User Models in Dialog Systems. Berlin: Springer, 1988, 4–34. (Extended and revised version of Report No. 3.)

B31 André, E., Herzog, G., Rist, T.: *On the Simultaneous Interpretation and Natural Language Description of Real World Image Sequences: The System SOCCER.* SFB 314 (VITRA), Bericht Nr. 31, April 1988. In: ECAI-88. Proceedings. London: Pitman, 1988, 449–454. #

B32 Kemke, C.: *Der Neuere Konnektionismus. Ein Überblick.* KI-Labor, Bericht Nr. 32, Mai 1988. In: Informatik-Spektrum, 11.3, Juni 1988, 143–162. #

B33 Retz-Schmidt, G.: *Various Views on Spatial Prepositions.* SFB 314 (VITRA), Bericht Nr. 33, Mai 1988. In: AI Magazine, 9.2, 1988, 95–105.

B34 Ripplinger, B., Kobsa, A.: *PLUG: Benutzerführung auf Basis einer dynamisch veränderlichen Zielhierarchie.* SFB 314 (XTRA), Bericht Nr. 34, Mai 1988. In: Hoeppner, W. (ed.): Künstliche Intelligenz. GWAI-88. 12th German Workshop on Artificial Intelligence. Proceedings. Berlin: Springer, 1988, 236–245.

B35 Schirra, J.R.J.: *Deklarative Programme in einem Aktor-System: MEGA-ACT.* FB Informatik, KI-Labor, Bericht Nr. 35, Mai 1988. In: KI, 2.3 / 2.4, 1988, 4–9/4–12.

B36 Retz-Schmidt, G.: *A REPLAI of SOCCER: Recognizing Intentions in the Domain of Soccer Games.* SFB 314 (VITRA), Bericht Nr. 36, Juni 1988. In: ECAI-88. Proceedings. London: Pitman, 1988, 455-457. #

B37 Wahlster, W.: *User and Discourse Models for Multimodal Communication.* SFB 314 (XTRA), Bericht Nr. 37, Juni 1988. In: Sullivan, J.W., Tyler, S.W. (eds.): Architectures for Intelligent Interfaces: Elements and Prototypes. Reading: Addison-Wesley, 1988.

B38 Harbusch, K.: *Effiziente Analyse natürlicher Sprache mit TAGs.* FB Informatik, KI-Labor, Bericht Nr. 38, Juni 1988. In: Batori, I.S., Hahn, U., Pinkal, M., Wahlster, W. (eds.): Computerlinguistik und ihre theoretischen Grundlagen. Symposium, Saarbrücken, März 1988. Proceedings. Berlin etc.: Springer, 1988, 79–103. #

B39 Schifferer, K.: *TAGDevEnv. Eine Werkbank für TAGs.* FB Informatik, KI-Labor, Bericht Nr. 39, Juni 1988. In: Batori, I.S., Hahn, U., Pinkal, M., Wahlster, W. (eds.): Computerlinguistik und ihre theoretischen Grundlagen. Symposium, Saarbrücken, März 1988. Proceedings. Berlin etc.: Springer, 1988, 152–171. #

B40 Finkler, W., Neumann, G.: *MORPHIX. A Fast Realization of a Classification-Based Approach to Morphology.* SFB 314 (XTRA), Bericht Nr. 40, Juni 1988. In: Trost, H. (ed.): 4. Österreichische Artificial-Intelligence-Tagung. Wiener Workshop - Wissensbasierte Sprachverarbeitung. Proceedings. Berlin etc.: Springer, 1988, 11-19. #

B41 Harbusch, K.: *Tree Adjoining Grammars mit Unifikation.* FB Informatik, KI-Labor, Bericht Nr. 41, Juni 1988. In: Trost, H. (ed.): 4. Österreichische Artificial-Intelligence-Tagung. Wiener Workshop - Wissensbasierte Sprachverarbeitung. Proceedings. Berlin etc.: Springer, 1988, 188-194. #

B42 Wahlster, W., Hecking, M., Kemke, C.: *SC: Ein intelligentes Hilfesystem für SINIX.* FB Informatik, KI-Labor, Bericht Nr. 42, August 1988. In: Gollan, W., Paul, W., Schmitt, A. (eds.), Innovative Informationsinfrastrukturen, Informatik-Fachberichte Nr. 184, Berlin: Springer, 1988.

B43 Wahlster, W.: *Natural Language Systems: Some Research Trends.* FB Informatik, KI-Labor, Bericht Nr. 43, August 1988. In: Schnelle, H., Bernsen, N.O. (eds.): Logic and Linguistics. Research Directions in Cognitive Science: European Perspectives, Vol. 2, Hove: Lawrence Erlbaum, 1989, 171–183.

B44 Kemke, C.: *Representing Neural Network Models by Finite Automata.* FB Informatik, KI-Labor, Bericht Nr. 44, August 1988. In: Proceedings of the 1st European Conference on Neural Networks "nEuro'88", Paris 1988. #

B45 Reddig, C.: *"3D" in NLP: Determiners, Descriptions, and the Dialog Memory in the XTRA Project.* SFB 314 (XTRA), Bericht Nr. 45, August 1988. In: Hoeppner, W. (ed.): Künstliche Intelligenz. GWAI-88. 12th German Workshop on Artificial Intelligence. Proceedings. Berlin: Springer, 1988, 159–168. #

B46 Scheller, A.: *PARTIKO. Kontextsensitive, wissensbasierte Schreibfehleranalyse und -korrektur.* FB Informatik, KI-Labor, Bericht Nr. 46, August 1988. In: Batori, I.S., Hahn, U., Pinkal, M., Wahlster, W. (eds.): Computerlinguistik und ihre theoretischen Grundlagen. Symposium, Saarbrücken, März 1988. Proceedings. Berlin: Springer, 1988, 136–151.

B47 Kemke, C.: *Darstellung von Aktionen in Vererbungshierarchien.* FB Informatik, KI-Labor, Bericht Nr. 47, September 1988. In: Hoeppner, W. (ed.): Künstliche Intelligenz. GWAI-88. 12th German Workshop on Artificial Intelligence. Proceedings. Berlin: Springer, 1988, 306–307.

B48 Jansen-Winkeln, R.M.: *WASTL: An Approach to Knowledge Acquisition in the Natural Language Domain.* SFB 314 (XTRA), Bericht Nr. 48, September 1988. In: Boose, J., et al. (eds.): Proceedings of the European Knowledge Acquisition Workshop (EKAW '88), Bonn 1988, 22-1–22-15.

B49 Kemke, C.: *What Do You Know About Mail? Representation of Commands in the SINIX Consultant.* FB Informatik, KI-Labor, Bericht Nr. 49, Dezember 1988. In: Norwig/Wahlster/Wilensky (eds.): Intelligent Help Systems for UNIX. Berlin: Springer, 1989. #

B50 Hecking, M.: *Towards a Belief-Oriented Theory of Plan Recognition* FB Informatik, KI-Labor, Bericht Nr. 50, Dezember 1988. In: Proceedings of the AAAI-88 Workshop on Plan Recognition. #

B51 Hecking, M.: *The SINIX Consultant — Towards a Theoretical Treatment of Plan Recognition —.* FB Informatik, KI-Labor, Bericht Nr. 51, Januar 1989. In: Norwig/Wahlster/Wilensky (eds.): Intelligent Help Systems for UNIX. Berlin: Springer, 1989. #

B52 Schmauks, D.: *Die Ambiguität von 'Multimedialität' oder: Was bedeutet 'multimediale Interaktion'?* SFB 314 (XTRA), Bericht Nr. 52, Februar 1989. In: Endres-Niggemeyer/Hermann/Kobsa/Rösner (eds.): Interaktion und Kommunikation mit dem Computer. Berlin: Springer, 1989, 94–103. #

B53 Finkler, W., Neumann, G.: *POPEL-HOW. A Distributed Parallel Model for Incremental Natural Language Production with Feedback.* SFB 314 (XTRA), Bericht Nr. 53, Mai 1989. In: IJCAI-89. Proceedings. 1518–1523. #

B54 Jung, J., Kresse, A., Reithinger, N., Schäfer, R.: *Das System ZORA. Wissensbasierte Generierung von Zeigegesten.* SFB 314 (XTRA), Bericht Nr. 54, Juni 1989. In: Metzing, D. (ed.): GWAI-89. Proceedings. Berlin: Springer, 1989, 190–194.

B55 Schirra, J.R.J.: *Ein erster Blick auf ANTLIMA: Visualisierung statischer räumlicher Relationen.* SFB 314 (VITRA), Bericht Nr. 55, Juni 1989. In: Metzing, D. (ed.): GWAI-89. Proceedings. Berlin: Springer, 1989, 301–311. #

B56 Hays, E.: *Two views of motion: On representing move events in a language-vision system.* SFB 314 (VITRA), Bericht Nr. 56, Juli 1989. In: Metzing, D. (ed.): GWAI-89. Proceedings. Berlin: Springer, 1989, 312–317. #

B57 Allgayer, J., Harbusch, K., Kobsa, A., Reddig, C., Reithinger, N., Schmauks, D.: *XTRA: A natural language access system to expert systems.* SFB 314 (XTRA), Bericht Nr. 57, Juli 1989. In: International Journal of Man-Machine Studies, 161–195.

B58 Herzog, G., Sung, C.-K., André, E., Enkelmann, W., Nagel, H.-H., Rist, T., Wahlster, W., Zimmermann, G.: *Incremental Natural Language Description of Dynamic Imagery.* SFB 314 (VITRA), Bericht Nr. 58, August 1989. In: Brauer, W., Freksa, C. (eds.): Wissensbasierte Systeme. Proceedings. Berlin: Springer, 1989, 153–162.

B59 Schirra, J.R.J.: *Einige Überlegungen zu Bildvorstellungen in kognitiven Systemen.* SFB 314 (VITRA), Bericht Nr. 59, August 1989. In: Freksa/Habel (eds.): Repräsentation und Verarbeitung räumlichen Wissens. Proceedings. Berlin: Springer, 1989, 68–82. #

B60 Herzog, G., Rist, T., André, E.: *Sprache und Raum: natürlichsprachlicher Zugang zu visuellen Daten.* SFB 314 (VITRA), Bericht Nr. 60, August 1989. In: Freksa/Habel (eds.): Repräsentation und Verarbeitung räumlichen Wissens. Proceedings. Berlin: Springer, 1989, 207–220. #

B61 Hays, E.M.: *On Defining Motion Verbs and Spatial Prepositions.* SFB 314 (VITRA), Bericht Nr. 61, Oktober 1989. In: Freksa/Habel (eds.): Repräsentation und Verarbeitung räumlichen Wissens. Proceedings. Berlin: Springer, 1989, 192–206. #

B62 Herzog, G., Retz-Schmidt, G.: *Das System SOCCER: Simultane Interpretation und natürlichsprachliche Beschreibung zeitveränderlicher Szenen.* SFB 314 (VITRA), Bericht Nr. 62, Oktober 1989. In: Perl, J. (ed.): Sport und Informatik. Schorndorf: Hofmann, 1989. #

B63 André, E., Herzog, G., Rist,: T. *Natural Language Access to Visual Data: Dealing with Space and Movement.* SFB 314 (VITRA), Bericht Nr. 63, November 1989. In: Nef, F., Borillo, M. (eds.): 1st Workshop on Logical Semantics of Time, Space and Movement in Natural Language. Proceedings. Edition Hermes. #

B64 Kobsa, A.: *User Modeling in Dialog Systems: Potentials and Hazards.* SFB 314 (XTRA), Bericht Nr. 64, Januar 1990. In: AI and Society. The Journal of Human and Machine Intelligence 4. 214–231. #

B65 Reithinger, N.: *POPEL — A Parallel and Incremental Natural Language Generation System.* SFB 314 (XTRA), Bericht Nr. 65, Februar 1990. In: Paris, C., Swartout, W., Mann, W. (eds.): Natural Language Generation in Artificial Intelligence and Computational Linguistics. Kluwer, 1990. 179–199.#

B66 Allgayer, J., Reddig, C.: *What KL-ONE Lookalikes Need to Cope with Natural Language — Scope and Aspect of Plural Noun Phrases.* SFB 314 (XTRA), Bericht Nr. 66, Februar 1990. In: Bläsius, K., Hedstück, U., Rollinger, C.-R. (eds.): Sorts and Types in Artificial Intelligence. Springer, 1990, 240-286.

B67 Allgayer, J., Jansen-Winkeln, R., Reddig, C., Reithinger, N.: *Bidirectional use of knowledge in the multi-modal NL access system XTRA* SFB 314 (XTRA), Bericht Nr. 67, November 1989. Proceedings of IJCAI-89. 1492–1497.

B68 Allgayer, J., Reddig, C.: *What's in a 'DET'? Steps towards Determiner-Dependent Inferencing.* SFB 314 (XTRA), Bericht Nr. 68, April 1990. In: Jorrand, P., Sendov, B. (eds.): Artificial Intelligence IV: methodology, systems, applications. Amsterdam: North Holland, 1990, 319-329.

B69 Allgayer, J.: *SB-ONE$^+$ — dealing with sets efficiently.* SFB 314 (XTRA), Bericht Nr. 69, Mai 1990. In: Proceedings of ECAI-90, 13-18.

B70 Allgayer, J., A. Kobsa, C. Reddig, N. Reithinger: *PRACMA: PRocessing Arguments between Controversially-Minded Agents.* SFB 314 (PRACMA), Bericht Nr. 70, Juni 1990. In: Proceedings of the Fifth Rocky Mountain Conference on AI: Pragmatics in Artificial Intelligence, Las Cruces, NM, 63–68.

B71 Kobsa, A.: *Modeling the User's Conceptual Knowledge in BGP-MS, a User Modeling Shell System.* SFB 314 (XTRA), Bericht Nr. 71, September 1990 (revised version). In: Computational Intelligence 6(4), 1990, 193–208.

B72 Schirra, J.R.J.: *Expansion von Ereignis-Propositionen zur Visualisierung. Die Grundlagen der begrifflichen Analyse von ANTLIMA.* SFB 314 (VITRA), Bericht Nr. 72, Juli 1990. In: Proceedings of GWAI-90, 246–256. #

B73 Schäfer, R.: *SPREADIAC. Intelligente Pfadsuche und -bewertung auf Vererbungsnetzen zur Verarbeitung impliziter Referenzen.* SFB 314 (XTRA), Bericht Nr. 73, August 1990. In: Proceedings of GWAI-90, 231-235.

B74 Schmauks, D., Wille, M.: *Integration of communicative hand movements into human-computer-interaction.* SFB 314 (XTRA), Bericht Nr. 74, November 1990. In: Computers and the Humanities 25, 1991, 129-140.

B75 Schirra, J.R.J.: *A Contribution to Reference Semantics of Spatial Prepositions: The Visualization Problem and its Solution in VITRA.* SFB 314 (VITRA), Bericht Nr. 75, Dezember 1990. In: Zelinsky-Wibbelt, Cornelia (ed.): The Semantics of Prepositions – From Mental Processing to Natural Language Processing. Berlin: Mouton de Gruyter, 1993, 471–515. #

B76 Kobsa, A.: *Utilizing Knowledge: The Components of the SB-ONE Knowledge Representation Workbench.* SFB 314 (XTRA), Bericht Nr. 76, Dezember 1990. In: Sowa, John (ed.): Principles of Semantic Networks: Explorations in the Representation of Knowledge. San Mateo, CA: Morgan Kaufmann, 1990, 457–486.

B77 Retz-Schmidt, G.: *Recognizing Intentions, Interactions, and Causes of Plan Failures.* SFB 314 (VITRA), Bericht Nr. 77, Januar 1991. In: User Modeling and User-Adapted Interaction, 1, 1991, 173–202.

B78 Kobsa, A.: *First Experiences with the SB-ONE Knowledge Representation Workbench in Natural-Language Applications.* SFB 314 (XTRA), Bericht Nr. 78, Juni 1991. In: AAAI Spring Symposium on Implemented Knowledge Representation and Reasoning Systems, Summer 1991, Stanford, CA, 125–139.

B79 Schmauks, D.: *Referent identification by pointing – classification of complex phenomena.* SFB 314 (XTRA), Bericht Nr. 79, Juli 1991. In: Geiger, Richard A. (ed.), Reference in Multidisciplinary Perspective: Philosophical Objact, Cognitive Subjet, Intersubjective Process. Hildesheim, Georg Olms Verlag, 1994.

B80 Tetzlaff, M., Retz-Schmidt, G.: *Methods for the Intentional Description of Image Sequences.* SFB 314 (VITRA), Bericht Nr. 80, August 1991. In: Brauer, W., Hernandez, D. (eds.): Verteilte KI und kooperatives Arbeiten. 4. Internationaler GI-Kongreß, 1991, Springer, 433–442.

B81 Schmauks, D.: *Verbale und nonverbale Zeichen in der Mensch-Maschine-Interaktion. Die Klassifikation von Pilzen.* SFB 314 (XTRA), Bericht Nr. 81, November 1991. In: Zeitschrift für Semiotik 16, 1994, 75–87.

B82 Reithinger, N.: *The Performance of an Incremental Generation Component for Multi-Modal Dialog Contributions.* SFB 314 (PRACMA), Bericht Nr. 82, Januar 1992. In: Proceedings of the 6. International Workshop on Natural Language Generation, 1992, Springer.

B83 Jansen-Winkeln, R.M., Ndiaye, A., Reithinger, N.: *FSS-WASTL: Interactive Knowledge Acquisition for a Semantic Lexicon.* SFB 314 (XTRA), Bericht Nr. 83, Februar 1992. In: Ardizonne E., Gaglio, S., Sorbello, F. (eds.): Trends in Artificial Intelligence, Proceedings of the second AI$*$IA 1991, Lecture Notes on Artificial Intelligence 529, 1991, Springer, 108–116.

B84 Kipper, B.: *MODALYS: A System for the Semantic-Pragmatic Analysis of Modal Verbs.* SFB 314 (PRACMA), Bericht Nr. 84, Mai 1992. In: Proceedings of the 5th International Conference on Artificial Intelligence - Methodology, Systems, Applications (AIMSA 92), September 1992, Sofia, Bulgaria, 171–180.

B85 Schmauks, D.: *Untersuchung nicht-kooperativer Interaktionen.* SFB 314 (PRACMA), Bericht Nr. 85, Juni 1992. In: Dialogisches Handeln. Festschrift zum 60. Geburtstag von Kuno Lorenz.

B86 Allgayer, J., Franconi, E.: *A Semantic Account of Plural Entities within a Hybrid Representation System.* SFB 314 (PRACMA), Bericht Nr. 86, Juni 1992. In: Proceedings of the 5th International Symposium on Knowledge Engineering, Sevilla, 1992.

B87 Allgayer, J., Ohlbach, H. J., Reddig, C.: *Modelling Agents with Logic. Extended Abstract.* SFB 314 (PRACMA), Bericht Nr. 87, Juni 1992. In: Proceedings of the 3rd International Workshop on User Modelling, Dagstuhl, 1992.

B88 Kipper, B.: *Eine Disambiguierungskomponente für Modalverben.* SFB 314 (PRACMA), Bericht Nr. 88, Juli 1992. In: Tagungsband der ersten Konferenz zur Verarbeitung natürlicher Sprache (KONVENS 92), Nürnberg, 1992, 258–267.

B89 Schmauks, D.: *Was heißt "taktil" im Bereich der Mensch-Maschine-Interaktion ?* SFB 314 (PRACMA), Bericht Nr. 89, August 1992. In: Proceedings des 3. Internationalen Symposiums für Informationswissenschaft ISI'92, September 1992, Saarbrücken, 13–25.

B90 Schirra, J.R.J.: *Connecting Visual and Verbal Space: Preliminary Considerations Concerning the Concept 'Mental Image'.* SFB 314 (VITRA), Bericht Nr. 90, Novenber 1992. In: Proceedings of the 4th European Workshop "Semantics of Time, Space and Movement and Spatio-Temporal Reasoning", September 4-8, 1992, Château de Bonas, France, 105–121.

B91 Herzog, G.: *Utilizing Interval-Based Event Representations for Incremental High-Level Scene Analysis.* SFB 314 (VITRA), Bericht Nr. 91, November 1992. In: Proceedings of the 4th European

Workshop "Semantics of Time, Space and Movement and Spatio-Temporal Reasoning", September 4-8, 1992, Château de Bonas, France, 425–435.

B92 Herzog, G., Maaß, W., Wazinski, P.: *VITRA GUIDE: Utilisation du Langage Naturel et de Représentations Graphiques pour la Description d'Itinéraires.* SFB 314 (VITRA), Bericht Nr. 92, Januar 1993. In: Colloque Interdisciplinaire du Comitée National "Images et Langages: Multimodalitée et Modélisation Cognitive", Paris, 1993, 243–251.

B93 Maaß, W., Wazinski, P., Herzog, G.: *VITRA GUIDE: Multimodal Route Descriptions for Computer Assisted Vehicle Navigation.* SFB 314 (VITRA), Bericht Nr. 93, Februar 1993. In: Proceedings of the Sixth International Conference on Industrial & Engineering Applications on Artificial Intelligence & Expert Systems. Edinburgh, U.K., June 1-4, 1993, 144–147.

B94 Schirra, J.R.J., Stopp, E.: *ANTLIMA – A Listener Model with Mental Images.* SFB 314 (VITRA), Bericht Nr. 94, März 1993. In: Proceedings of IJCAI-93. Chambéry, France, August 29 - September 3, 1993, 175–180.

B95 Maaß, W.: *A Cognitive Model for the Process of Multimodal, Incremental Route Descriptions.* SFB 314 (VITRA), Bericht Nr. 95, Mai 1993. In: Proceedings of the European Conference on Spatial Information Theory. Lecture Notes in Compute Science, Springer. Marciana Marina, Elba, Italy, September 19-22, 1993, 1–24.

B96 Jameson, A.: *Müssen Dialogsysteme immer objektiv sein? Fragen wir sie selbst!.* SFB 314 (PRACMA), Bericht Nr. 96, Mai 1993. In: Künstliche Intelligenz 7(2), 1993, 75–81.

B97 Schneider, A.: *Connectionist Simulation of Adaptive Processes in the Flight Control System of Migratory Locusts.* Fachbereich Informatik, KI-Labor, Bericht Nr. 97, September 1993. In: Proceedings of Artificial Neural Networks in Engineering 1993, St. Louis, Missouri, USA: Intelligent Engineerung Systems Through Artificial Neural Networks, Vol. 3, November 1993, ASME Press, New York, USA, 599–604.

B98 Kipper, B.: *A Blackboard Architecture for Natural Language Analysis.* SFB 314 (PRACMA), Bericht Nr. 98, Februar 1994. In: Proceedings of the Seventh Florida Artificial Intelligence Research Symposium (FLAIRS 94), May 5-7, 1994, Pensacola Beach, USA, 231–235.

B99 Gapp, K.-P.: *Einsatz von Visualisierungstechniken bei der Analyse von Realweltbildfolgen.* SFB 314 (VITRA), Bericht Nr. 99, April 1994. In: Tagungsband des 1. Workshops zum Thema *visual computing*, Darmstadt, 1994.

B100 Herzog, G., Wazinski, P.: *VIsual TRAnslator: Linking Perceptions and Natural Language Descriptions.* SFB 314 (VITRA), Bericht Nr. 100, April 1994. In: Artificial Intelligence Review Journal, 8(2), Special Volume on the Integration of Natural Language and Vision Processing, edited by P. Mc Kevitt, 1994, 175-187.

B101 Gapp, K.-P.: *Basic Meanings of Spatial Relations: Computation and Evaluation in 3D Space.* SFB 314 (VITRA), Bericht Nr. 101, April 1994. In: Proceedings of the 12th National Conference on Artificial Intelligence (AAAI-94), Seattle, Washington, USA, 1994, 1393–1398.

B102 Gapp, K.-P., Maaß, W.: *Spatial Layout Identification and Incremental Descriptions.* SFB 314 (VITRA), Bericht Nr. 102, Mai 1994. In: Proceedings of the Workshop on the Integration of Natural Language and Vision Processing, 12th National Conference on Artificial Intelligence (AAAI-94), Seattle, Washington, USA, August 2-3, 1994, 145–152.

B103 André, E., Herzog, G., Rist, T.: *Multimedia Presentation of Interpreted Visual Data.* SFB 314 (VITRA), Bericht Nr. 103, Juni 1994. In: Proceedings of the Workshop on the Integration of Natural Language and Vision Processing, 12th National Conference on Artificial Intelligence (AAAI-94), Seattle, Washington, USA, August 2-3, 1994, 74–82.

B104 Lueth, T.C., Laengle, Th., Herzog, G., Stopp, E., Rembold, U.: *KANTRA: Human-Machine Interaction for Intelligent Robots Using Natural Language.* SFB 314 (VITRA), Bericht Nr. 104, Juni 1994. In: Proceedings of the 3rd International Workshop on Robot and Human Communication (RO-MAN '94), Nagoya University, Nagoya, Japan, July 18-20, 1994, 106–111.

B105 Kipper, B., Jameson, A.: *Semantics and Pragmatics of Vague Probability Expressions.* SFB 314 (PRACMA), Bericht Nr. 105, Juni 1994. In: Proceedings of the Sixteenth Annual Conference of the Cognitive Science Society, Atlanta, Georgia, USA, August 1994, 496–501.

B106 Jameson, A., Kipper, B., Ndiaye, A., Schäfer, R., Simons, J., Weis, T., Zimmermann, D.: *Cooperating to Be Noncooperative: The Dialog System PRACMA.* SFB 314 (PRACMA), Bericht Nr. 106, Juni 1994. In: Nebel, B., Dreschler-Fischer, L. (eds.): KI-94: Advances in Artificial Intelligence, Proceedings of the Eighteenth German Conference on Artificial Intelligence, Saarbrücken, Germany, September 1994, Berlin, Heidelberg: Springer, 106–117.

B107 Stopp, E., Gapp, K.-P., Herzog, G., Laengle, T., Lueth, T.C.: *Utilizing Spatial Relations for Natural Language Access to an Autonomous Mobile Robot.* SFB 314 (VITRA), Bericht Nr. 107, Juli 1994. In: Nebel, B., Dreschler-Fischer, L. (eds.): KI-94: Advances in Artificial Intelligence, Proceedings of the Eighteenth German Conference on Artificial Intelligence, Saarbrücken, Germany, September 1994, Berlin, Heidelberg: Springer, 39–50.

B108 Maaß, W.: *From Vision to Multimodal Communication: Incremental Route Descriptions.* SFB 314 (VITRA), Bericht Nr. 108, Juli 1994. In: Artificial Intelligence Review Journal, 8(2/3), Special Volume on the Integration of Natural Language and Vision Processing, 1994, 159–174.

B109 Ndiaye, A., Jameson, A.: *Supporting Flexibility and Transmutability: Multi-Agent Processing and Role-Switching in a Pragmatically Oriented Dialog System.* SFB 314 (PRACMA), Bericht Nr. 109, August 1994. In: Jorrand, P., Sgurev, V. (eds.): Proceedings of the Sixth Annual Conference on Artificial Intelligence: Methodology, Systems, Applications (AIMSA '94), Sofia, Bulgaria, September 21-24, 1994, 381–390.

B110 Gapp, K.-P.: *From Vision to Language: A Cognitive Approach to the Computation of Spatial Relations in 3D Space.* SFB 314 (VITRA), Bericht Nr. 110, Oktober 1994. In: Proceedings of the First Conference on Cognitive Science in Industry, Luxembourg, 1994, 339–357.

B111 Gapp, K.-P.: *A Computational Model of the Basic Meanings of Graded Composite Spatial Relations in 3-D Space.* SFB 314 (VITRA), Bericht Nr. 111, Oktober 1994. In: Proceedings of the Advanced Geographic Data Modelling Workshop, Delft, The Netherlands, 1994.

B112 Schäfer, R.: *Multidimensional Probabilistic Assessment of Interest and Knowledge in a Noncooperative Dialog Situation.* SFB 314 (PRACMA), Bericht Nr. 112, Dezember 1994. In: Proceedings of ABIS-94: GI Workshop on Adaptivity and User Modeling in Interative Software Systems, St. Augustin, October 1994, 46–62.

B113 André, E., Herzog, G., Rist, T.: *Von der Bildfolge zur multimedialen Präsentation.* SFB 314 (VITRA), Bericht Nr. 113, Februar 1995. In: Arbeitsgemeinschaft Simulation in der Gesellschaft für Informatik (ASIM), Mitteilungen aus den Arbeitskreisen, Heft Nr. 46, Fachtagung "Integration von Bild, Modell und Text", Magdeburg, 2.-3. März 1995, 129–142.

B114 Laengle, T., Lueth, T.C., Stopp, E., Herzog, G., Kamstrup, G.: *KANTRA - A Natural Language Interface for Intelligent Robots.* SFB 314 (VITRA), Bericht Nr. 114, März 1995. In: Proc. of the 4th International Conference on Intelligent Autonomous Systems, Karlsruhe, Germany, 1995, 357–364.

B115 Gapp, K.-P.: *Angle, Distance, Shape, and their Relationship to Projective Relations.* SFB 314 (VITRA), Bericht Nr. 115, Mai 1995. In: Proc. of the 17th Conference of the Cognitive Science Society, Pittsburgh, PA, 1995.

B116 Maaß, W.: *How Spatial Information Connects Visual Perception and Natural Language Generation in Dynamic Environments: Towards a Computational Model.* SFB 314 (VITRA), Bericht Nr. 116, Juni 1995. In: Proceedings of the 2nd International Conference on Spatial Information Theory (COSIT'95), Vienna, September 21-23, 1995, 223–240.

B117 Maaß, W., Baus, J., Paul, J.: *Visual Grounding of Route Descriptions in Dynamic Environments.* SFB 314 (VITRA), Bericht Nr. 117, Juli 1995. To appear in: Proceedings of the AAAI Fall Symposium on "Computational Models for Integrating Language and Vision", MIT, Cambridge, MA, USA, 1995.

B118 Gapp, K.-P.: *An Empirically Validated Model for Computing Spatial Relations.* SFB 314 (VITRA), Bericht Nr. 118, Juli 1995. In: Wachsmuth, I., Rollinger, C.-R., Brauer, W. (eds.): Advances in Artificial Intelligence, Proceedings of the 19th Annual German Conference on Artificial Intelligence (KI-95), Bielefeld, Germany, September 1995, Berlin, Heidelberg: Springer, 245–256.

B119 Gapp, K.-P.: *Object Localization: Selection of Optimal Reference Objects.* SFB 314 (VITRA), Bericht Nr. 119, Juli 1995. In: Proceedings of the 2nd International Conference on Spatial Information Theory (COSIT'95), Vienna, September 21-23, 1995, 519–536.

B120 Herzog, G.: *Coping with Static and Dynamic Spatial Relations.* SFB 314 (VITRA), Bericht Nr. 120, Juli 1995. In: Amsili, P., Borillo, M., Vieu, L. (eds.): Proceedings of TSM'95, Time, Space, and Movement: Meaning and Knowledge in the Sensible World, Groupe "Langue, Raisonnement, Calcul", Toulouse, Château de Bonas, France, 1995, C 47–59.

B121 Blocher, A., Schirra, J.R.J.: *Optional Deep Case Filling and Focus Control with Mental Images: ANTLIMA-KOREF.* SFB 314 (VITRA), Bericht Nr. 121, Juli 1995. In: Proceedings of the 14th International Joint Conference on Artificial Intelligence (IJCAI-95), Montréal, Canada, August 19-25, 1995, 417–423.

B122 Herzog, G., Rohr, K.: *Integrating Vision and Language: Towards Automatic Description of Human Movements.* SFB 314 (VITRA), Bericht Nr. 122, Juli 1995. In: Wachsmuth, I., Rollinger, C.-R., Brauer, W. (eds.): Advances in Artificial Intelligence, Proceedings of the 19th Annual German Conference on Artificial Intelligence (KI-95), Bielefeld, Germany, September 1995, Berlin, Heidelberg: Springer, 257–268.

B123 Blocher, A., Stopp, E.: *Time-Dependent Generation of Minimal Sets of Spatial Descriptions.* SFB 314 (VITRA), Bericht Nr. 123, Juli 1995. To appear in: Proceedings of the Workshop on the Representation and Processing of Spatial Expressions at the 14th International Joint Conference on Artificial Intelligence (IJCAI-95), Montréal, Canada, August 19, 1995.

B124 Herzog, G.: *From Visual Input to Verbal Output in the Visual Translator.* SFB 314 (VITRA), Bericht Nr. 124, Juli 1995. To appear in: Proceedings of the AAAI Fall Symposium on "Computational Models for Integrating Language and Vision", MIT, Cambridge, MA, USA, 1995.

B125 Jameson, A., Schäfer, R., Simons, J., Weis, T.: *Adaptive Provision of Evaluation-Oriented Information: Tasks and Techniques.* SFB 314 (PRACMA), Bericht Nr. 125, Juli 1995. In: Proceedings of the 14th International Joint Conference on Artificial Intelligence (IJCAI-95), Montréal, Canada, August 19-25, 1995, 1886–1893.

B126 Jameson, A.: *Logic Is Not Enough: Why Reasoning About Another Person's Beliefs Is Reasoning Under Uncertainty.* SFB 314 (PRACMA), Bericht Nr. 126, Juli 1995. In: Laux, A., Wansing, H. (eds.): Knowledge and Belief in Philosophy and Artificial Intelligence, Berlin: Akademie-Verlag, 1995, 199–229.

B127 Zimmermann, D.: *Exploiting Models of Musical Structure for Automatic Intention-Based Composition of Background Music.* KI-Labor, Bericht Nr. 127, Juli 1995. In: Proceedings of the workshop on Artificial Intelligence and Music at the 14th International Joint Conference on Artificial Intelligence (IJCAI-95), Montréal, Canada, official IJCAI-95 workshop proceedings, Menlo Park (CA): AAAI Press, 1995, 55–62.

B128 Ndiaye, A., Jameson, A.: *Predictive Role Taking in Dialog: Global Anticipation Feedback Based on Transmutability.* SFB 314 (PRACMA), Bericht Nr. 128, November 1995. In: Proceedings of the Fifth International Conference on User Modeling, Kailua-Kona, Hawaii, USA, January 1996, 137–144.

B129 Stopp, E., Laengle, T.: *Natürlichsprachliche Instruktionen an einen autonomen Serviceroboter.* SFB 314 (VITRA), Bericht Nr. 129, November 1995. In: Dillmann, R., Rembold, U., Lüth, T.: Autonome Mobile Systeme 1995, Tagungsband der "Autonome Mobile Systeme" (AMS '95), 30. November - 1. Dezember 1995, Berlin, Heidelberg: Springer, 1995, 299–308.

B130 Wahlster, W., Jameson, A., Ndiaye, A., Schäfer, R., Weis, T.: *Ressourcenadaptive Dialogführung: ein interdisziplinärer Forschungsansatz.* SFB 378 (READY), Bericht Nr. 130, November 1995. In: Künstliche Intelligenz, 9(6), 1995, 17–21.

B131 Jameson, A.: *Numerical Uncertainty Management in User and Student Modeling: An Overview of Systems and Issues.* SFB 314 (PRACMA), Bericht Nr. 131, November 1995. To appear in: User Modeling and User-Adapted Interaction, 5, 1995.

B132 Längle, T., Lüth, T.C., Stopp, E., Herzog, G.: *Natural Language Access to Intelligent Robots: Explaining Automatic Error Recovery.* SFB 378 (REAL), Bericht Nr. 132, Oktober 1996. In: Ramsay, A.M. (ed.): Artificial Intelligence: Methodology, Systems, Applications, Proc. of AIMSA'96, Sozopol, Bulgaria, September 1996, Amsterdam: IOS Press, 259–267.

B133 Jameson, A., Weis, T.: *How to Juggle Discourse Obligations.* SFB 378 (READY), Bericht Nr. 133, Oktober 1996. In: Proceedings of the Symposium on Conceptual and Semantic Knowledge in Language Generation, Heidelberg, Germany, 15-17 November, 1995, 171–185.

B134 Weis, T.: *Die Rolle von Diskursverpflichtungen in bewertungsorientierten Informationsdialogen.* SFB 378 (READY), Bericht Nr. 134, Oktober 1996. In: Dafydd Gibbon (ed.): Natural Language Processing and Speech Technology. Results of the 3rd KONVENS Conference, Bielefeld, Germany, October 1996, Berlin: Mouton de Gruyter.

B135 Gapp, K.-P.: *Processing Spatial Relations in Object Localization Tasks.* SFB 378 (REAL), Bericht Nr. 135, Oktober 1996. In: Proceedings of AAAI Fall Symposium on „Computational Models for Integrating Language and Vision", MIT, Cambridge, MA, USA, 1995.

B136 Ndiaye, A.: *Rollenübernahme in einem Dialogsystem.* SFB 378 (READY), Bericht Nr. 136, Januar 1997. In: Künstliche Intelligenz, 10(4), 1996, 34–40.

B137 Stopp, E., Blocher, A.: *Construction of Mental Images and their Use in a Listener Model.* SFB 378 (REAL), Bericht Nr. 137, Januar 1997. In: Proceedings of the Symposium on Conceptual and Semantic Knowledge in Language Generation, Heidelberg, Germany, 15-17 November, 1995, 270–280.

B138 Schäfer, R., Weyrath, T.: *Assessing Temporally Variable User Properties With Dynamic Bayesian Networks.* SFB 378 (READY), Bericht Nr. 138, August 1997. In: Jameson, A., Paris, C., Tasso, C. (Eds.): User Modeling: Proceedings of the Sixth International Conference, UM97. Vienna, New York: Springer, 1997, 377–388.

B139 Weis, T.: *Resource-Adaptive Action Planning in a Dialogue System for Repair Support.* SFB 378 (READY), Bericht Nr. 139, August 1997. To appear in: Nebel, B. (ed.): Proceedings der 21. Deutschen Jahrestagung für Künstliche Intelligenz, Freiburg im Breisgau, Deutschland. Berlin, New York: Springer, 1997.

B140 Herzog, G., Blocher, A., Gapp, K.-P., Stopp, E., Wahlster, W.: *VITRA: Verbalisierung visueller Information.* SFB 378 (REAL), Bericht Nr. 140, Januar 1998. In: Informatik - Forschung und Entwicklung, 11(1), 1996, 12–19.

B141 Wahlster, W., Blocher, A., Baus, J., Stopp, E., Speiser, H.: *Ressourcenadaptierende Objektlokalisation: Sprachliche Raumbeschreibung unter Zeitdruck.* SFB 378 (REAL, VEVIAG), Bericht Nr. 141, April 1998. Erscheint in: Kognitionswissenschaft, Sonderheft zum Sonderforschungsbereich 378, Berlin, Heidelberg: Springer, 1998.

B142 Zimmer, H.D., Speiser, H.R., Baus, J., Blocher, A., Stopp, E.: *The Use of Locative Expressions in Dependence of the Spatial Relation between Target and Reference Object in Two-Dimensional Layouts.* SFB 378 (REAL, VEVIAG), Bericht Nr. 142, April 1998. In: Freksa, C., Habel, C., Wender, K.F. (eds.): Spatial cognition - An interdisciplinary approach to representation and processing of spatial knowledge, Berlin, Heidelberg: Springer, 1998, 223–240.

B143 Wahlster, W., Tack, W.: *SFB 378: Ressourcenadaptive Kognitive Prozesse.* SFB 378 (REAL), Bericht Nr. 143, April 1998. In: Jarke, M., Pasedach, K., Pohl, K. (Hrsg.): Informatik'97 - Informatik als Innovationsmotor, 27. Jahrestagung der Gesellschaft für Informatik, Aachen, 24.-26. September 1997, Berlin, Heidelberg: Springer, 1997, 51–57.

Memos

M1 Baltes, H., Houy, C., Scheller, A., Schifferer, K.: *PORTFIX - Portierung des FRANZ-LISP-Systems von VAX/UNIX nach PC-MX/SINIX.* FB Informatik, III-Projekt, Memo Nr. 1, Juli 1985. #

M2 Grasmück, R., Guldner, A.: *Wissensbasierte Fertigungsplanung in Stanzereien mit FERPLAN: Ein Systemüberblick.* FB Informatik, KI-Labor, Memo Nr. 2, August 1985. #

M3 Baltes, H.: *GABI — Ein wissensbasiertes Geldanlageberatungsprogramm.* FB Informatik, KI-Labor, Memo Nr. 3, November 1985. #

M4 Schmauks, D.: *Formulardeixis und ihre Simulation auf dem Bildschirm. Ein Überblick aus linguistischer Sicht.* SFB 314, (XTRA), Memo Nr. 4, Februar 1986. In: Conceptus 55, 1987, 83–102. #

M5 André, E., Bosch, G., Herzog, G., Rist, T.: *Characterizing Trajectories of Moving Objects Using Natural Language Path Descriptions.* SFB 314, (VITRA), Memo Nr. 5, März 1986. In: ECAI 86. The 7th European Conference on Artificial Intelligence. Proceedings, Vol. 2. Brighton 1986, 1–8. #

M6 Baltes, H.: *GABI: Frame-basierte Wissensrepräsentation in einem Geldanlageberatungssystem.* FB Informatik, KI-Labor, Memo Nr. 6, März 1986. #

M7 Brach, U., Woll, W.: *WILIE — Ein wissensbasiertes System zur Vereinfachung der interaktiven Literaturerfassung.* FB Informatik, KI-Labor, Memo Nr. 7, April 1986. #

M8 Finkler, W., Neumann, G.: *MORPHIX — Ein hochportabler Lemmatisierungsmodul für das Deutsche.* FB Informatik, KI-Labor, Memo Nr. 8, Juli 1986.

M9 Portscheller, R.: *AIDA — Rekursionsbehandlung, Konfliktauflösung und Regelcompilierung in einem Deduktiven Datenbanksystem.* FB Informatik, KI-Labor, Memo Nr.9, Oktober 1986.#

M10 Schirra, J.: *MEGA-ACT — Eine Studie über explizite Kontrollstrukturen und Pseudoparallelität in Aktor-Systemen mit einer Beispielarchitektur in FRL.* FB Informatik, KI-Labor, Memo Nr. 10, August 1986. #

M11 Allgayer, J., Reddig, C.: *Systemkonzeption zur Verarbeitung kombinierter sprachlicher und gestischer Referentenbeschreibungen.* SFB 314 (XTRA), Memo Nr. 11, Oktober 1986. #

M12 Herzog, G.: *Ein Werkzeug zur Visualisierung und Generierung von geometrischen Bildfolgenbeschreibungen.* SFB 314 (VITRA), Memo Nr. 12, Dezember 1986. #

M13 Retz-Schmidt, G.: *Deictic and Intrinsic Use of Spatial Prepositions: A Multidisciplinary Comparision.* SFB 314 (VITRA), Memo Nr. 13, Dezember 1986. In: Kak, A., Chen, S.-S. (eds.): Spatial Reasoning and Multisensor Fusion, Proceedings of the 1987 Workshop. Los Altos, CA: Morgan Kaufmann, 1987, 371–380. #

M14 Harbusch, K.: *A First Snapshot of XTRAGRAM, A Unification Grammar for German Based on PATR.* SFB 314 (XTRA), Memo Nr. 14, Dezember 1986. #

M15 Fendler, M., Wichlacz, R.: *SYCON — Symbolische Constraint-Propagierung auf Netzwerken, Entwurf und Implementierung.* FB Informatik, KI-Labor, Memo Nr. 15, März 1987.

M16 Dengler, D., Gutmann, M., Hector, G., Hecking, M.: *Der Planerkenner REPLIX.* FB Informatik, KI-Labor (SC-Projekt), Memo Nr. 16, September 1987. #

M17 Hecking, M., Harbusch, K.: *Plan Recognition through Attribute Grammars.* FB Informatik, KI-Labor (SC-Projekt), Memo Nr. 17, September 1987.

M18 Nessen, E.: *SC-UM — User Modeling in the SINIX-Consultant.* FB Informatik, KI-Labor (SC-Projekt), Memo Nr. 18, November 1987. In: Applied Artificial Intelligence, 3,1, 1989, 33–44. #

M19 Herzog, G.: *LISPM Miniatures, Part I.* SFB 314 (VITRA), Memo Nr. 19, November 1987.

M20 Blum, E.J.: *ROSY — Menü-basiertes Parsing natürlicher Sprache unter besonderer Berücksichtigung des Deutschen.* FB Informatik, KI-Labor, Memo Nr. 20, Dezember 1987.

M21 Beiche, H.-P.: *Zusammenwirken von LST-1, PLUG, FORMULAR und MINI-XTRA.* SFB 314 (XTRA), Memo Nr. 21, Januar 1988. #

M22 Allgayer, J., Harbusch, K., Kobsa, A., Reddig, C., Reithinger, N., Schmauks, D., Wahlster, W.: *Arbeitsbericht des Projektes NS1: XTRA für den Förderungszeitraum vom 1.4.85 bis 31.12.87.* SFB 314 (XTRA), Memo Nr. 22, Januar 1988. #

M23 Kobsa, A.: *A Bibliography of the Field of User Modeling in Artificial Intelligence Dialog Systems.* SFB 314 (XTRA), Memo Nr. 23, April 1988.

M24 Schmauks, D., Reithinger, N.: *Generierung multimodaler Ausgabe in NL Dialogsystemen — Voraussetzungen, Vorteile und Probleme.* SFB 314 (XTRA), Memo Nr. 24, April 1988. #

M25 Herzog, G., Rist, T.: *Simultane Interpretation und natürlichsprachliche Beschreibung zeitveränderlicher Szenen: Das System SOCCER.* SFB 314 (VITRA), Memo Nr. 25, August 1988. #

M26 André, E.: *Generierung natürlichsprachlicher Äußerungen zur simultanen Beschreibung von zeitveränderlichen Szenen: Das System SOCCER.* SFB 314 (VITRA), Memo Nr. 26, August 1988. #

M27 Kemke, C.: *Die Modellierung neuronaler Verbände basierend auf Netzwerken endlicher Automaten.* FB Informatik, KI-Labor, Memo Nr. 27, August 1988. In: Tagungsband des Workshops "Konnektionismus", St. Augustin 1988.

M28 Hecking, M., Kemke, C., Nessen, E., Dengler, D., Gutmann, M., Hector, G.: *The SINIX Consultant — A Progress Report.* FB Informatik, KI-Labor (SC-Projekt), Memo Nr. 28, August 1988.

M29 Bauer, M., Diewald, G., Merziger, G., Wellner, I.: *REPLAI. Ein System zur inkrementellen Intentionserkennung in Realwelt-Szenen.* SFB 314 (VITRA), Memo Nr. 29, Oktober 1988. #

M30 Kalmes, J.: *SB-Graph User Manual (Release 0.1).* SFB 314 (XTRA), Memo Nr. 30, Dezember 1988.

M31 Kobsa, A.: *The SB-ONE Knowledge Representation Workbench.* SFB 314 (XTRA), Memo Nr. 31, März 1989. #

M32 Aue, D., Heib, S., Ndiaye, A.: *SB-ONE Matcher: Systembeschreibung und Benutzeranleitung.* SFB 314 (XTRA), Memo Nr. 32, März 1989.

M33 Profitlich, H.-J.: *Das SB-ONE Handbuch. Version 1.0* . SFB 314 (XTRA), Memo Nr. 33, April 1989.

M34 Bauer, M., Merziger, G.: *Conditioned Circumscription: Translating Defaults to Circumscription.* FB Informatik, KI-Labor (SC-Projekt), Memo Nr. 34, Mai 1989.

M35 Scheller, A.: *PARTIKO. Kontextsensitive, wissensbasierte Schreibfehleranalyse und -korrektur.* FB Informatik, KI-Labor, Memo Nr. 35, Mai 1989.

M36 Wille, M.: *TACTILUS II. Evaluation und Ausbau einer Komponente zur Simulation und Analyse von Zeigegesten.* SFB 314 (XTRA), Memo Nr. 36, August 1989. #

M37 Müller, S.: *CITYGUIDE — Wegauskünfte vom Computer.* SFB 314 (VITRA), Memo Nr. 37, September 1989.

M38 Reithinger, N.: *Dialogstrukturen und Dialogverarbeitung in XTRA.* SFB 314 (XTRA), Memo Nr. 38, November 1989.

M39 Allgayer, J., Jansen-Winkeln, R., Kobsa, A., Reithinger, N., Reddig, C., Schmauks, D.: *XTRA. Ein natürlichsprachliches Zugangssystem zu Expertensystemen.* SFB 314 (XTRA), Memo Nr. 39, Dezember 1989. #

M40 Beiche, H.-P.: *LST-1. Ein Expertensystem zur Unterstützung des Benutzers bei der Durchführung des Lohnsteuerjahresausgleichs.* SFB 314 (XTRA), Memo Nr. 40, Dezember 1989.

M41 Dengler, D.: *Referenzauflösung in Dialogen mit einem intelligenten Hilfesystem.* FB Informatik (KI-Labor), Memo Nr. 41, Januar 1990. #

M42 Kobsa, A.: *Conceptual Hierachies in Classical and Connectionist Architecture.* SFB 314 (XTRA), Memo Nr. 42, Januar 1990.

M43 Profitlich, H.-J.: *SB-ONE: Ein Wissenrepräsentationssystem basierend auf KL-ONE.* SFB 314 (XTRA), Memo Nr. 43, März 1990.

M44 Kalmes, J.: *SB-Graph. Eine graphische Benutzerschnittstelle für die Wissensrepräsentationswerkbank SB-ONE.* SFB 313 (XTRA), Memo Nr. 44, März 1990.

M45 Jung, J., Kresse, A., Schäfer, R.: *ZORA. Ein Zeigegestengeneratorprogram.* SFB 314 (XTRA), Memo Nr. 45, April 1990.

M46 Harbusch, K., Huwig, W.: *XTRAGRAM. A Unification Grammar for German Based on PATR.* SFB 314 (XTRA), Memo Nr. 46, April 1990. #

M47 Scherer, J.: *SB-PART Handbuch. Version 1.0* SFB 314 (XTRA), Memo Nr. 47, Juli 1990.

M48 Scherer, J.: *SB-PART: ein Partitionsverwaltungssystem für die Wissensrepräsentationssprache SB-ONE.* SFB 314 (XTRA), Memo Nr. 48, September 1990. #

M49 Schirra, J.R.J.: *Zum Nutzen antizipierter Bildvorstellungen bei der sprachlichen Szenenbeschreibung. — Ein Beispiel —.* SFB 314 (VITRA), Memo Nr. 49, Dezember 1991. #

M50 Blocher, A., Stopp, E., Weis, T.: *ANTLIMA-1: Ein System zur Generierung von Bildvorstellungen ausgehend von Propositionen.* SFB 314 (VITRA), Memo Nr. 50, März 1992.

M51 Allgayer, J., Franconi, E.: *Collective Entities and Relations in Concept Languages.* SFB 314 (PRACMA), Memo Nr. 51, Juni 1992.

M52 Allgayer, J., Schmitt, R.: *SB·LITTERS: Die Zugangssprache zu* **SB-ONE^{+}**. SFB 314 (PRACMA), Memo Nr. 52, August 1992.

M53 Herzog, G.: *Visualization Methods for the VITRA Workbench.* SFB 314 (VITRA), Memo Nr. 53, Dezember 1992.

M54 Wazinski, P.: *Graduated Topological Relations.* SFB 314 (VITRA), Memo Nr. 54, Mai 1993.

M55 Jameson, A., Schäfer, R.: *Probabilistische Einschätzung von Wissen und Interessen. Die Anwendung der intuitiven Psychometrik im Dialogsystem PRACMA.* SFB 314 (PRACMA), Memo Nr. 55, Juni 1993.

M56 Schneider, A.: *Konnektionistische Simulation adaptiver Leistungen des Flugsteuersystems der Wanderheuschrecke.* KI-Labor, Memo Nr. 56, Juni 1993.

M57 Kipper, B., Ndiaye, A., Reithinger, N., Reddig, C., Schäfer, R.: *Arbeitsbericht für den Zeitraum 1991–1993: PRACMA* SFB 314 (PRACMA), Memo Nr. 57, Juli 1993.

M58 Herzog, G., Schirra, J., Wazinski, P.: *Arbeitsbericht für den Zeitraum 1991–1993: VITRA* SFB 314 (VITRA), Memo Nr. 58, Juli 1993.

M59 Gapp, K.-P.: *Berechnungsverfahren für räumliche Relationen in 3D-Szenen.* SFB 314 (VITRA), Memo Nr. 59, August 1993.

M60 Stopp, E.: *GEO-ANTLIMA: Konstruktion dreidimensionaler mentaler Bilder aus sprachlichen Szenenbeschreibungen.* SFB 314 (VITRA), Memo Nr. 60, Oktober 1993.

M61 Blocher, A.: *KOREF: Zum Vergleich intendierter und imaginierter Äußerungsgehalte.* SFB 314 (VITRA), Memo Nr. 61, Mai 1994.

M62 Paul, M.: *IBEE: Ein intervallbasierter Ansatz zur Ereigniserkennung für die inkrementelle Szenenfolgenanalyse.* SFB 314 (VITRA), Memo Nr. 62, November 1994.

M63 Stopp, E., Blocher, A.: *Spatial Information in Instructions and Questions to an Autonomous System.* SFB 378 (REAL), Memo Nr. 63, Mai 1997.

M64 Stopp, E.: *Ein Modell für natürlichsprachlichen Zugang zu autonomen Robotern.* SFB 378 (REAL), Memo Nr. 64, Mai 1997.

M65 Schäfer, R., Bauer, M. (Hrsg.): *ABIS-97, 5. GI-Workshop Adaptivität und Benutzermodellierung in interaktiven Softwaresystemen, 30.9. bis 2.10.1997, Saarbrücken.* SFB 378 (READY), Memo Nr. 65, September 1997.

M66 Rupp, U.: *GRATOR - Räumliches Schließen mit GRAdierten TOpologischen Relationen über Punktmengen.* SFB 378 (REAL), Memo Nr. 66, April 1998.

M67 Kray, C.: *Ressourcenadaptierende Verfahren zur Präzisionsbewertung von Lokalisationsausdrücken und zur Generierung von linguistischen Hecken.* SFB 378 (REAL), Memo Nr. 67, April 1998.

M68 Müller, C.: *Das REAL Speech Interface.* SFB 378 (REAL), Memo Nr. 68, Januar 1999.

M69 Wittig, F.: *Ein Java-basiertes System zum Ressourcenmanagement in Anytime-Systemen.* SFB 378 (REAL), Memo Nr. 69, Januar 1999.

M70 Lindmark, K.: *Identifying Symptoms of Time Pressure and Cognitive Load in Manual Input.* SFB 378 (READY), Memo Nr. 70, September 2000.

M71 Beckert, A.: *Kompilation von Anytime-Algorithmen: Konzeption, Implementation und Analyse.* SFB 378 (REAL), Memo Nr. 71, Mai 2001

M72 Beckert, A.: ORCAN*: Laufzeitmessungen von Methoden zur Anytime-Kompilierung.* SFB 378 (REAL), Memo Nr. 72, Mai 2001.

M73 Baus, A., Beckert, A.: ORCAN*: Implementation von Methoden zur Compilierung von Anytime-Algorithmen.* SFB 378 (REAL), Memo Nr. 73, Mai 2001.

M74 Weyrath, T.: *Erkennung von Arbeitsgedächtnisbelastung und Zeitdruck in Dialogen - Empirie und Modellierung mit Bayes'schen Netzen.* SFB 378 (READY), Memo Nr. 74, Mai 2001.

M75 Berthold, A.: *Repräsentation und Verarbeitung sprachlicher Indikatoren für kognitive Ressourcenbeschränkungen.* SFB 378 (READY), Memo Nr. 75, Mai 2001.

M76 Gebhard, P.: ROPLEX*: Natürlichsprachliche Beschreibung von generischen Roboterplandaten.* SFB 378 (REAL), Memo Nr. 76, Mai 2001.

M77 Werner, A.: EBABA*: Probabilistische Einschätzung von Bewertungskriterien aufgrund bewertender Äusserungen.* SFB 378 (READY), Memo Nr. 77, Mai 2001.

M78 Brandherm, B.: *Rollup-Verfahren für komplexe dynamische Bayes'sche Netze.* SFB 378 (READY), Memo Nr. 78, Mai 2001.

M79 Lohse, M.: *Ein System zur Visualisierung von Navigationsauskünften auf stationären und mobilen Systemen.* SFB 378 (REAL), Memo Nr. 79, Mai 2001.

M80 Müller, C.: *Symptome von Zeitdruck und kognitiver Belastung in gesprochener Sprache: eine experimentelle Untersuchung.* SFB 378 (READY), Memo Nr. 80, Mai 2001.

M81 Decker, B.: *Implementation von Lernverfahren für Bayes'sche Netze mit versteckten Variablen.* SFB 378 (READY), Memo Nr. 81, Mai 2001.

M82 Krüger, A., Malaka, R. (Eds.): *Artificial Intelligence in Mobile Systems 2003 (AIMS 2003)* SFB 378 (REAL), Memo Nr. 82, Oktober 2003.

M83 Butz, A., Kray, C., Krüger, A., Schmidt, A. (Eds.): *Workshop on Multi-User and Ubiquitous User Interfaces 2004 (MU3I 2004)* SFB 378 (REAL), Memo Nr. 83, Januar 2004.

M84 Baus, J., Kray, C., Porzel, R. (Eds.): *Artificial Intelligence in Mobile Systems 2004 (AIMS 2004)* SFB 378 (REAL), Memo Nr. 84, September 2004.

M85 Butz, A., Kray, C., Krüger, A., Schmidt, A., Prendinger, H. (Eds.): *2nd Workshop on Multi-User and Ubiquitous User Interfaces 2005 (MU3I 2005)* SFB 378 (REAL), Memo Nr. 85, Januar 2005.

M86 Pederson, T., Pinto, H., Schmitz, M., Stahl, C., Terrenghi, L. (Eds.): *Modelling and Designing User Assistance in Intelligent Environments (MODIE 2006)* SFB 378 (BAIR), Memo Nr. 86, September 2006.