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.

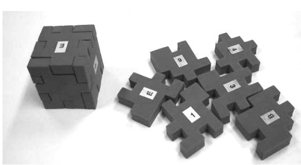

Figure 1: The 3d-Structure Puzzle.

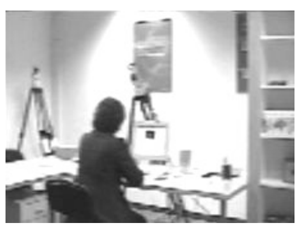

Figure 2: The user being observed by the Wizard of Oz.

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.