PROTOTYPING OF AN AUTOMATIC PROMPTING SYSTEM FOR A RESIDENTIAL HOME

Christian Peters, Thomas Hermann, Sven Wachsmuth
CITEC, Bielefeld University

ABSTRACT

In this paper, we describe the prototyping process of an automatic prompting system in the healthcare domain. The application area of our system will be a residential home where persons with various cognitive disabilities live together. We describe the specific requirements for an automatic prompting system in a residential home where persons are directly assisted by a caregiver in the execution of Activities of Daily Living (ADLs). We observe real-world trials using qualitative data analysis techniques and conduct a first study using a Wizard-of-Oz (WOz) methodology where we describe the user's reaction behavior when faced with system prompts instead of direct caregiver prompts.

INTRODUCTION AND RELATED WORK

In this paper, we describe our efforts in the prototyping process of an automatic prompting system in the healthcare domain. Mostly, the application area of such systems are the individual user's home with the aim to prolong the user's independence in everyday life [5,4] focussing on user groups with age-specific limitations, e.g. elderly people with Alzheimer's disease or dementia [4,2]. However, we aim to develop an automatic prompting system for a residential home where persons with various cognitive disabilities live together and share the same system. The requirements for a system in a residential home vary compared to a system in a user's individual home: Persons living at home are used to fulfill Activities of Daily Living (ADLs) without the help of a caregiver for their whole life but might not be able to do so anymore. Inhabitants of residential homes are usually assisted by a caregiver in several ADLs over a long period of time. We cooperate with *Haus Bersaba*[1], a residential home belonging to *v. Bodelschwinghsche Stiftungen Bethel*, a care facility in Bielefeld, Germany. Our user group are persons with cognitive disabilities who have problems fulfilling ADLs, in particular brushing their teeth. The cognitive impairments include learning disabilities, obsessiveness, epilepsy, behavioral and autistic spectrum disorders. Usually, these persons are directly assisted by a caregiver standing besides them. Hence, inferring requirements by introspection of regular persons brushing their teeth is not plausible since the caregiver's behavior is as important as the user's behavior. We describe in-situ observations on recorded real-world trials where each video shows one trial of a user brushing his/her teeth while being observed and supported by a caregiver if necessary. We use Interaction Unit (IU) analysis described in [6] as a method of qualitative data analysis which focuses on the analysis of the task and the user's interaction with the objects in the environment. We depict the interaction behavior between caregiver and user by describing the caregiver's way of prompting the user to fulfill the task as well as the user's reaction to the prompts. With the results gathered from the qualitative video analysis we designed a first study using Wizard-of-Oz (WOz) methodology. WOz is widely used in the evaluation of ubiquitous computing systems [3]. In a WOz study, the user interacts with an obviously fully functioning system, but parts of it are operated by a *wizard*. In our scenario, the caregiver is the wizard generating audio/video prompts to assist the user if necessary. Here, we analyze the user's reaction behavior to system prompts in comparison to direct caregiver prompts.

1 The authors would like to thank the inhabitants and caregivers of *Haus Bersaba* for their high motivation to cooperate in the project “Task Assistance for Persons with Cognitive Disabilities” (TAPeD) of the Cognitive Interaction Technology Center of Excellence (CITEC), Bielefeld University.

QUALITATIVE DATA ANALYSIS

Qualitative data analysis describes a variety of procedures and methods to analyze and interpret qualitative data, e.g. videos. We analyze the requirements for an automatic prompting system in a residential home by applying qualitative data analysis to videos recorded in *Haus Bersaba*. Each video shows one trial of a user brushing his/her teeth while being observed and supported by a caregiver if necessary. An intervention of the caregiver is necessary if the user is not able to proceed in the execution of the task. Our data base consists of eight users performing a total number of 23 trials at three different days (seven users conducted three trials, one user only two). The users are supported by two caregivers assisting in 10 and 13 trials, respectively. We recorded each trial with two FullHD-camcorders: one camcorder has a static position with a perspective from the left side of the washstand and the other is manually operated by the first author to observe the scene as shown in figure 1. The analysis of the videos focus on two aspects: Firstly, the examination of an ideal sequence for task execution as well as the mental processes involved. We apply Interaction Unit Analysis [6] and describe the results in section *Interaction Unit Analysis*. Secondly, we analyze the prompting behavior of the caregiver and infer the requirements for an automatic prompting system in a residential home in section *Caregiver's prompting behavior*.

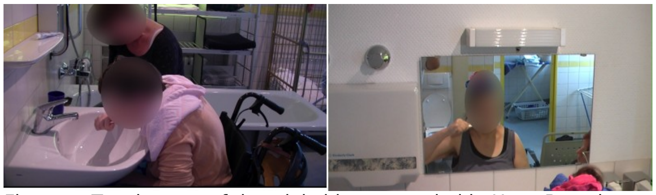

Figure 1: Two images of the trial videos recorded in Haus Bersaba

Interaction Unit analysis

Interaction Unit (IU) analysis proposed by Ryu and Monk [6] is an approach to interaction modeling describing the conjunction of cognitive and environmental pre- and postconditions for individual actions. In this paper, we use an adapted form of IU analysis as stated in [1] by observing the recorded trial videos in an iterative process: We identify an ideal sequence of goals and behaviors to accomplish the task of brushing teeth and describe the mental processes triggering the user's behavior. Table 1 depicts the results: *User behavior* shows the ideal sequence of behaviors and identifies subactions necessary to execute the task. The column *Recognition/Recall/Affordance* describes the mental processes needed to take an action based on the current environment: *Recognition*(Rn) means that the user can directly perceive an objects' state in the environment, e.g. mug empty in IU17. *Recall*(Rl): The user needs to remember a certain state of the environment which is not directly observable, e.g. brush dry in IU8. *Affordance*(Af) describes the recognition of the meaning of an object and the way to use it, e.g. the tap can be altered to *on* which makes the water flow. We analyzed the trial videos with regard to the mental processes involved. In 90 percent of the cases, a caregiver had to prompt a user in actions which involve a recall process, e.g. remembering the next step in the brushing task. Recall processes are not directly observable in the environment, but have to be remembered. For the automatic prompting system, we make recall processes directly observable by prompting the user with a video if they are stuck in task execution. The videos show prerecorded actions of the task performed by the first author, e.g. filling the mug with water or rinsing the mouth. The videos are displayed on a TFT installed below the mirror and are therefore directly integrated into the environment. Additionally to the Interaction Unit analysis, we describe the prompting behavior of the caregiver in the following section.

Caregiver's interaction behavior

The direct assistance of a user by a caregiver is an integral part in a residential home scenario. In this section, we focus on the caregiver's prompting behavior, also obtained by analyzing the recorded trial videos. The caregiver interacts with a user by verbal and non-verbal cues: Most interactions consist of verbal prompts paired with either haptic or visual feedback. Haptic feedback includes touching to attract the user's attention. Visual feedback is provided with both deictic gestures (e.g. caregiver points to an object of interest) and iconic gestures (e.g. caregiver demonstrates a specific movement which can be adapted by the user).

Table 1: Results of the IU analysis for brushing teeth. Column *User behavior* describes one possible ideal sequence of actions. Column *Recognition/Recall/Affordances* depicts the mental processes involved in each step: Rn Recognition, Rl Recall, Af Affordance, TT tube of toothpaste

	IU	User behavior	Recognition/Recall/Affordance
(brush_with_paste)	1	No action	Rn brush, Rn TT on counter, Rl step
	2	Take TT from counter	Af TT
	3	Alter TT to open	Af TT cap closed, Rl contains paste
	4	Take brush from counter	Af brush
	5	Spread paste on brush	Af TT
	6	Alter TT to closed	Rn paste on brush, Af TT cap opened
	7	Give TT to counter	Rn TT closed, Af counter
(brush_wet)	8	No action	Rl brush dry, Rl step
	9	give brush to tap	Af brush
	10	alter tap to on	Af tap off
	11	alter tap to off	Rl brush wet, Af tap on
(teeth_clean)	12	No action	Rl step
	13	give brush to mouth	Af brush
	14	brush all teeth	Af brush
	15	spit out	Rn mouth full of foam, Af sink
(mug_water)	16	No action	Rn mug on counter, Rl step
	17	give mug to tap	Rn mug empty, Af tap
	18	alter tap to on	Af tap off
	19	alter tap to off	Af tap on
(mouth_clean)	20	No action	Rl step
	21	rinse mouth	Af mug water
	22	spit out	Af sink
(mug_clean)	23	No action	Rn mug dirty, Rl step
	24	alter tap to on	Af tap off
	25	give mug under tap	Rn water on, Af tap
	26	alter tap to off	Rn water on, Af tap on
	27	give mug to counter	Af counter
(brush_clean)	28	No action	Rn brush dirty, Rl step
	29	take brush from counter	Rn brush on counter, Af tap
	30	alter tap to on	Af tap off
	31	give brush under tap	Rn water on, Af tap
	32	alter tap to off	Rn water on, Af tap on
	33	give brush to counter	Af counter
(mouth_dry)	34	No action	Af towel
	35	take towel from hook	Rn towel on hook, Af towel
	36	dry mouth	Rn mouth wet, Af towel
	37	give towel to hook	Af towel

We identified verbal communication as the main modality of interaction. We assigned each verbal prompt in the trial videos to one of three categories:

command The caregiver prompts the user to perform a certain action (*take water first* or *put paste on the brush*) which was either forgotten or performed at the wrong time in the execution.

attract attention The caregiver tries to attract the attention of the user by giving verbal prompts like *Look* *here* or directly addressing the person by saying his/her name.

encouragement Encouragement is given from the caregiver in two situations: Firstly, if the caregiver didn't need to help the user for a certain time (reward for a good performance). Secondly, encouragement is a positive reaction to a change in the user's behavior (confirmation that the user has correctly adapted his/her behavior with regard to the given command).

As the main result of the caregiver's prompting behavior, we identified verbal commands paired with either haptic or visual feedback as the main prompting modality. Haptic feedback is used to attract the attention of the user. Since we don't want an automatic prompting system to directly actuate in the user's environment, we avoid haptic feedback and use a verbal *stop* command to attract the user's attention. In our system, we adapted the verbal command and visual feedback given by the caregiver by using prerecorded audio commands in conjunction with video snippets. We evaluate the user's reaction behavior to system prompts in a study described in the following section.

USER STUDY

We conducted a study with the aim to analyze the user's reaction behavior when faced with system prompts instead of direct caregiver prompts. Since we are in the prototyping phase of our project where we have no fully functioning system, yet, we use the WOz methodology: The user gets the task to brush his/her teeth at a washstand equipped with two cameras, a microphone and a TFT display with speakers. He is not aware of being supported by a caregiver but thinks that he is faced with an automatic prompting system. However, the caregiver - the wizard in our scenario - operates the system via a graphical user interface (GUI) and can't observe the washstand directly. Instead, he gets the live-streamed images and audio from the sensors installed at the washstand. The caregiver assists the user in the brushing task by generating prompts via the GUI. The prompts are delivered to the user via the TFT display and speakers in realtime. Figure 2 shows the GUI used by the caregiver for prompting.

Each click on a button generates a single prompt. The GUI is divided into different sections: The middle section displays the live-streamed images of the cameras installed at the washstand. Additionally, the wizard gets the audio stream via headphones. The sections below and above the livestreamed images show the buttons according to different functional phases during the task. The lower section contains buttons according to steps in task execution, e.g. rinse mouth, clean brush, etc. Each of the buttons prompts the user with a verbal command in combination with a video snippet showing the corresponding action. The upper section contains buttons according to the brushing phase. Since the distinction between slight differences, e.g. brush left and brush up, is hard to obtain on a video snippet, we use pure verbal

Figure 2: Caregiver's/Wizard's graphical user interface for generating prerecorded audio and audio/video prompts by clicking on buttons

commands for assistance during the brushing phase. The buttons in the upper left of the GUI trigger a verbal stop and encouragement command, respectively. Encouragement is an important means to motivate the user during the task. The stop command is important to attract the attention of the user which is usually done by haptic feedback of the caregiver in regular trials. Since we don't want our system to directly actuate in the environment, we avoid haptic feedback and use a verbal command *stop* to attract the user's attention. To ease the use of the GUI and ensure contemporary prompting, we use a Wacom touch tablet with a pen-like input device. We trained the caregivers in a preliminary study where we presented prerecorded trial videos showing the first author brushing his teeth. The aim of the prestudy was to ensure that the caregiver is familiar with the operation of the device in the WOz study. We conducted six WOz trials in total with three persons performing two trials each. We are interested in the user's reaction behavior to system prompts in comparison to prompts given by the caregiver directly. Hence, we also conducted five trials performed by five persons where the caregiver directly prompts the user. We will refer to the two scenarios with *WIZ* (for system prompting generated by the wizard) and *CG* (for direct caregiver prompting), respectively. The results are described in the following section.

Results

Column “avg” in the upper part of table 2 shows the user's reaction behavior with respect to the CG (five trials performed by five subjects) and WIZ scenario (six trials performed by three subjects doing two trials each).

Each user (in both the CG as well as the WIZ trials) was able to fulfill the task of brushing teeth properly which was judged by a caregiver after the trial. The average number of prompts is slightly higher in the CG scenario with 9.5 prompts compared to the WIZ scenario with 7.75, given a total number of 38 CG and 46 WIZ prompts. We categorize the reactions to the prompts into three classes: *no* reaction, *false* and *correct* reaction to a prompt. The user's reaction behavior to prompts is comparable in both WIZ and CG scenario. However, the average results hide significant variations amongst users with respect to individual reaction behavior. In the following, we compare three users who conducted nine trials in total (one CG trial and two WIZ trials each). We discard the trials of two users who have conducted no WIZ trials, here. Columns “user i” in the upper part of Table 2 show the individual reaction behavior in the CG and WIZ scenario listed for the three users. We identified two main behaviors: User 1 shows a shift from *false* reactions in the CG scenario to *correct* reactions in the WIZ scenario with similar *no* reaction. *Correct* reactions are significantly increased in WIZ (85%) compared to CG (44%). *False* reactions in CG (44%) are highly decreased in WIZ (5%). We found user 1 having very good abilities to understand system prompts. Users 2 and 3 show a different behavior where the rate of *no* reaction is highly increased in the WIZ scenario: User 2 for example has a similar *false* rate in both scenarios, but *no* reaction is highly increased in the WIZ scenario (42%) compared to CG (29%). User 3 always reacted to the prompts in CG, either correctly or falsely. However, we see a shift from *correct*/*false* reactions to *no* reaction in WIZ. Users 2 and 3 seem to be distracted from task execution by system prompting. For a more detailed description of the user's individual reaction behavior in the WIZ scenario, we take into account the reaction to audio prompts on the one hand and combined audio/video prompts on the other hand shown in the lower part of table 2: We dropped the average results, since we are interested in the individual user's reaction, here. We see significant differences in the reaction to audio prompting compared to combined audio/video prompting: User 2 has a moderate rate of *correct* reactions to audio prompts (44%) and no *false* reactions. However, we see 60% *false* reactions to audio/video prompts. User 2 has severe problems following both task execution and audio/video prompting. We believe that the video snippets distract user 2 in task execution instead of assisting him. User 1, however, has 100% *correct* reactions on audio/video prompts. The video snippets combined with audio seem to be the best way to prompt user 1 in task execution, even better than direct caregiver prompting. The results show that the requirements for an automatic prompting system in a residential home where persons with several disabilities live together are different compared to a system which is applied in an individual user's home: A system in a residential home has to be highly adaptable to the abilities of many different users by providing prompts from different modalities. Furthermore, for each individual user, there is a fine line between assistance and distraction: A video snippet may distract some users from fulfilling the task since they can't concentrate on both task execution and video prompting at the same time. Hence, the modalities of prompting have to be chosen in conjunction with the caregiver and with respect to the user's individual abilities, ideally even adjustable by the system during a trial.

Table 2: Upper part: Average and user's individual reaction behavior to system prompts (WIZ) and direct caregiver prompts (CG) in %. Lower part: User's individual reaction behavior to audio (A) and combined audio/video (A/V) prompts in %.

reaction	**user 1**	**user 2**	**user 3**	**avg**	**scenario**
correct	85 (44)	26 (43)	43 (67)	54 (60)	
false	5 (44)	32 (29)	14 (33)	17 (18)	**WIZ (CG)**
no	10 (11)	42 (29)	43 (0)	28 (21)	
correct	100 (77)	10 (44)	33 (50)	/	
false	0 (8)	60 (0)	33 (50)	/	**A/V (A)**
no	0 (15)	30 (56)	33 (0)	/	

CONCLUSION

In this paper, we described the system prototyping process of an automatic prompting system which we aim to apply in a residential home where persons with several disabilities live together. We analyzed the specific scenario in a residential home where users are directly assisted in the task of brushing their teeth by a caregiver standing besides them. We found verbal commands paired with gestures as the main modality of prompting the user and adapted the caregiver's behavior in a WOz study. We analyzed the user's reaction behavior when faced with system prompting instead of direct caregiver prompting. Our results show that an automatic prompting system has to be highly adaptable to the user's needs and abilities in terms of prompting behavior: Some users showed an increased rate of *correct* reactions to video prompts whereas others were highly distracted and were not able to adapt their behavior according to the prompt. However, all users were able to fulfill the brushing task while assisted with system prompts.

REFERENCES

[1] J. Hoey et al., “Rapid specification and automated generation of prompting systems to assist people with dementia”, Pervasive and Mobile Computing, 2011

[2] H. Kautz, D. Fox, G. Borriello, O. Etzioni and L. Arnstein, “An overview of the assisted cognition project”, Proc. AAAI-2002

[3] K. Mäkelä, E.-P. Salonen, M. Turunen, J. Hakulinen and R. Raisamo, “Conducting a Wizard of Oz experiment on a ubiquitous computing system doorman”, Proc. IPNMD-2001

[4] A. Mihailidis, J. Boger, T. Craig and J. Hoey, “The COACH prompting system to assist older adults with dementia through handwashing: An efficacy study”, BMC Geriatrics, 8(1), 2008

[5] C. Ní Scanaill et al., “A review of approaches to mobility telemonitoring of the elderly in their living environment”, Annals of Biomedical Engineering, 34(4), 2006

[6] H. Ryu and A. Monk, “Interaction Unit Analysis: A new interaction design framework”, Human-Computer Interaction, 24(4), 2009