Proactivity in Spoken Dialog Systems

Hansjörg Hofmann, Ute Ehrlich, André Berton
Daimler AG
Ulm, Germany
{hansjoerg.hofmann, ute.ehrlich, andre.berton}@daimler.com

ABSTRACT

Proactive speech interfaces have been a hot research topic for many years. However, until today, no precise definition of proactive behavior in spoken dialog systems (SDSs) and its influencing factors has been made. Therefore, this paper aims at defining the characteristics of proactivity with the focus on SDSs. The definitions are derived from other research fields and then transferred to SDSs.

A general proactivity system model, which describes the relevant system components and their interaction is described. A proactive system receives information from a knowledge source and notifies the user about an incoming event without a user request. The system has to act user-friendly and take the current user state and the environment into account. Thus, the proactive behavior can be identified as anticipatory, change-oriented and self-initiated. A proactive human-machine speech dialog can be structured in 3 stages. First, the user has to be notified about an incoming event, then the problem solving process has to be started. Finally, the new task has to be completed and possibly paused tasks have to be resumed.

Author Keywords

proactivity; spoken dialog systems; situation-awareness

ACM Classification Keywords

H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

INTRODUCTION

Today, smartphones are considered as people's companions and are used in various daily situations. People use smartphones to browse the Web, buy things online or communicate via social media, Email and other (instant) messaging applications. Especially, the need to stay "always connected" has increased enormously within the last years and people exchange more and more over-the-top content (OTT) messages[1]. According to Informa Telecoms & Media each OTT messaging user sends an average of 32.6 OTT messages every day [3]. In online communication users take the initiative to interact but also the smartphone triggers the interaction and notifies user about new incoming events. This so-called "proactive" behavior will increasingly demand users as more and more messages are exchanged per day.

In some situation where the manual use of smartphones is not in focus, for instance when driving your car, this increased mental demand can be distractive and dangerous. In so-called dual-task scenarios people perform a secondary task (e.g. reading an email) in parallel to a primary task (e.g. driving), which requires the attention to several sources of information simultaneously. Wickens proposes that performing dual-tasks in parallel is achieved best, when the required user workload is distributed on several resources [15]. As many primary tasks in dual task scenarios tap haptic input and visual output channels speech interfaces are a good means to assist users in a comfortable and safe way [13]. Therefore, spoken dialog systems (SDSs) should be used to perform secondary tasks in a dual-task scenario and to notify users about new information.

Research about proactive SDSs has emerged in the last 10-15 years. However, up to today, no definition of proactivity has been introduced, yet. Related research investigates and characterizes some aspects of proactive behavior of speech interfaces but does not define proactivity and its requirements concretely (e.g., [10, 14, 4]). Therefore, this paper aims at defining the characteristics of proactivity in human-machine interaction (HMI) with the focus on SDSs. In the next Section, a general proactivity system model is presented, which describes the relevant components and influencing factors, which have an impact on proactive systems. In Section 3, a general definition of proactive behavior derived from literature and different research fields is presented, which can also be applied to the field of HMI. Subsequently, the definition of proactivity is transferred to SDSs and an overview about existing proactive SDSs is given. In the final Section, the findings are summarized.

PROACTIVITY SYSTEM MODEL

Imagine, two communication partners (*A* and *B*) would like to exchange information via any system. Therefore, *A* sends the information carrier to the system. When the system has received the information it has two possibilities to interact to handle the information. Either, the system stores the information and waits until *B* requests the information, or it proactively sends the information to *B*. The latter activity would

[1]OTT-messaging applications are downloadable smartphone applications, which enable users to send (instant) text messages for free, using mobile Internet access [3].

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
IUI 2014 Workshop: Interacting with Smart Objects, February 24, 2014, Haifa, Israel
Copyright is held by the author/owner(s).

resemble proactive behavior. Figure 1 illustrates a simple and generalized system model of such a proactive acting system in which *A* would be represented by the source and *B* by the user. The system which manages the information exchange could be a postal service, a smartphone or an in-car SDS. For example, imagine, *A* sends a letter to person *B* via any postal service. The postal service receives the letter from *A*. Instead of waiting until the recipient picks up the letter from the post office, the postman delivers the letter proactively to the addressee [8].

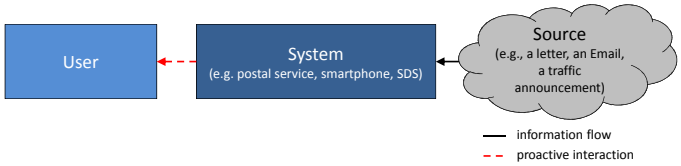

Figure 1. General proactivity system model.

Successful proactive behavior can only be achieved if the proactive system observes the environment in order to act in advance on a future situation and to deliver the content at the right point in time. Therefore, the system model needs to be extended by a context component, which is illustrated in Figure 2 [8]. Context-awareness can relate to the current location, time or situation, knowledge about user preferences, etc. The context knowledge can be gained by observing the environment and from the user himself. For example, if an addressee has changed his residence recently the postal service needs to be aware of the new address. If person *A* would like to send a letter to *B* without knowing the new address the postal service has to take care that the letter arrives at the correct address.

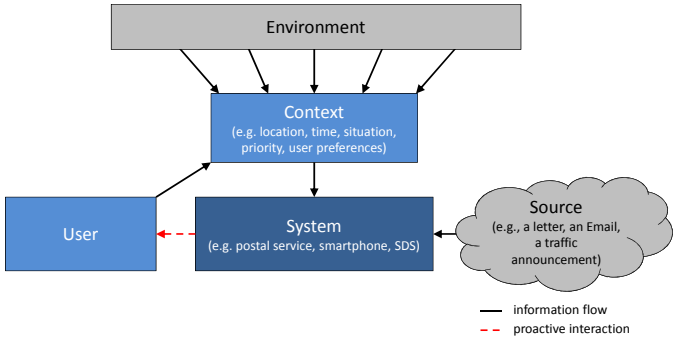

Figure 2. Extended proactivity system model.

If the user is performing a primary task in parallel, such as driving a car the system model needs to be extended (see Figure 3). Depending on the nature of the primary task the task can occupy several input and output channels simultaneously. For instance, by steering and keeping one's eyes on the road driving a car demands a person visually and manually. The state of the primary task needs to be included in the context knowledge. Thereby, the system might know if the user is currently able to process the information, which the system tries to deliver. E.g., if the user is very busy performing the primary task the system should not proactively interact with the user in order to not interfere the primary task.

Proactivity is a relatively new field in HMI and therefore, proactive behavior has not precisely been characterized, yet. Apart from HMI, research in human-human communication

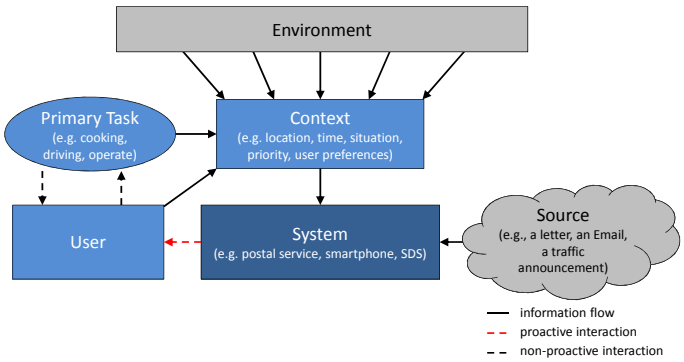

Figure 3. Extended proactivity system model in a dual-task scenario.

and industrial/organizational psychology investigate proactive behavior. The next Section presents a characterization of proactive behavior derived from literature and other research fields, which is applied on SDSs afterwards.

CHARACTERISTICS OF PROACTIVE BEHAVIOR

Dictionary definitions [1, 9, 16] typically contain two key features of proactivity. First, an anticipatory element is emphasized, which involves acting in advance of a future situation, such as "acting in anticipation of future problems, needs or changes" [9]. Second, these definitions highlight taking control and causing change, for example: "controlling a situation by causing something to happen rather than waiting to respond to it after it happens" [16]. These two elements - anticipation and taking control - can be found in most conceptualizations of general proactive behavior. E.g., [12, p. 636] define proactive behavior as "self-initiated anticipatory action that aims to change and improve the situation". In addition, definitions of proactive behavior often emphasize its self-initiative nature, which addresses the attempt to solve problems, which have not yet occurred [5]. Summarizing these definition, proactivity can be described by three key features (as defined in [11]): proactive behavior is

1. **anticipatory** - instead of reacting it involves scanning the environment and acting in advance to a further situation;
2. **change-oriented** - instead of passively adapting to the situation or waiting for something to happen being proactive means to take control or cause something to happen;
3. **self-initiated** - the control is taken on a self-initiative base without being requested to do so.

Most of the definitions are applied in work psychology and used to describe proactive behavior of employees in order to improve individual and organizational effectiveness. For example, a nurse, who is waiting for the doctor sees a patient and prepares the equipment and data the doctor might need. Thereby, the doctor can do his work more effectively. The nurse acts anticipatory by thinking ahead and anticipating the doctor's needs. Instead of waiting for the doctor to come she becomes active and prepares the equipment. The initiative to do so is taken all by herself without being requested by the doctor [11].

PROACTIVITY IN SPOKEN DIALOG SYSTEMS

As the definitions for proactive behavior are formulated in a general manner they can be transferred to other research areas, such as human-human communication or HMI, too. The

focus of this research work is on SDSs. Therefore, the characteristics of proactive behavior and the system model theory are transferred to speech interfaces in the following. Furthermore, the process of a proactive speech interaction is explained in detail.

Proactive Behavior in Spoken Dialog Systems

The proactive behavior of an SDS can also be characterized by the 3 key features proposed by Parker et al. [11]. A proactive SDS should act *anticipatory*, *change-oriented* and *self-initiated*. As the proposed proactivity system model applies for all kind of proactive systems it can be applied to SDSs, too (see Figure 4).

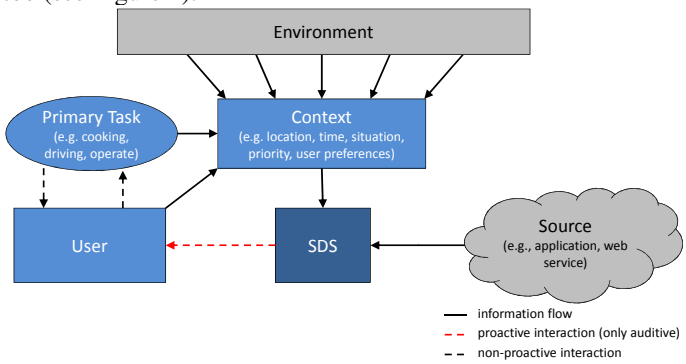

Figure 4. SDS proactivity system model in a dual-task scenario.

A proactive SDS receives the information from an application or a Web Service which is linked to the system. The SDS has to capture the spacial, temporal and user specific context of an interaction in order to take control anticipatory in advance to a further situation, possibly even before the users have become aware of the problem. Furthermore, the system needs to understand the user's current psychological situation, intention and actions and has to keep track of the dialog history. Then, it is able to assist the user in a meaningful way. A proactive SDS initiates the speech interaction itself and not only upon the user's request [10].

Imagine an in-vehicle navigation system, which observes the traffic density on the previously configured route while driving. As the system detects a traffic jam, which would prolong the length of the trip the system speaks up to the driver and suggests to take a different route. Here, the system acts anticipatory by observing the traffic density ahead and preventing the user from a possible traffic jam. Instead of ignoring the pending problem the system suggests to change the route to bypass the traffic jam. The system initiates the dialog itself without a request by the user. The only difference to the proactive behavior of the employee above is that the system only makes suggestions to the driver and does not decide about the new route itself. The driver has the control of changing the route himself.

Proactive Speech Interaction Process

The proactive speech interaction process can be structured in several stages. The dialog flow is illustrated in Figure 5 and the different dialog steps are described in the following.

Before the SDS delivers some new incoming information or informs the user about an upcoming problem two different scenarios are conceivable: Either the user is idle or the user

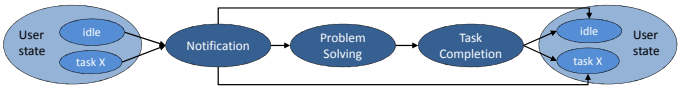

Figure 5. Proactive speech dialog flow.

interacts with the system already. When the SDS initiates the speech interaction, the different speech dialog steps have to be walked through:

1. **Notification**: First, the system has to grab the user's attention to tell him that there is some new information. The manner an SDS interrupts the ongoing dialog or initiates the interaction should be situation-sensitive and user-friendly [4]. If the situation does not allow a speech interaction at the moment, the system should not address the user. E.g., if the user is dictating a sensitive Email he should not be interrupted by the system. When the SDS decides to notify the user, in order to appear user-friendly, the SDS could allow the user to decide if he wants to enter into the new dialog or to reject talking about the newly introduced topic.
2. **Problem Solving**: In the course of the speech dialog the user interacts with the SDS on a regular basis. Depending on the dialog modeling and the competences of the SDS the speech interaction can appear more or less conversational.
3. **Task Completion**: When the problem has been solved or the new information has been delivered, the new task is completed. Depending on the initial state of the user the previous task should be resumed or the SDS should disappear again. Again, by negotiating the desired process the system could leave the decision to the user in a user-friendly manner.

In order to design a proactive SDS the different stages of the presented speech dialog flow should be taken into consideration. Several approaches to proactive speech interaction have been made within the last years. The most advanced projects and products are described in the following.

Overview on Proactive Spoken Dialog Systems

There are only few research projects who incorporate proactive behavior in SDS. The DARPA Communicator program[2] (2000-2001) focused on the improvement of SDSs, which allow for performing complex tasks by using speech as sole input modality. The DARPA projects helped to gain knowledge about proactive dialogue management. In the SmartKom project[3] (1999-2003) complex multimodal dialogs are aspired in which the user as well as the system can initiate interactions. Kwaku, as part of the Neem project [2] was a virtual meeting partner, which performed organisational tasks, such as monitoring the time spent on certain agenda points and reminds participants proactively to go to the next item, if necessary. Strauß et al. [14] envisaged an SDS which listens to multiparty conversations and assists the users proactively in a restaurant search and investigate the user behavior in a Wizard-Of-Oz study.

Today, there already exist products, which notify the user proactively. E.g. most of today's navigation systems employ data retrieved from the Traffic Message Channel (TMC)

[2] Websites offline.
[3] http://www.smartkom.org/

for routing taking into account real-time traffic situations [6]. The speech-enabled navigation systems prompt the driver proactively or play a warning sound if, for example a traffic jam on the route appears. Smartphones alert users about incoming emails or instant messages, upcoming appointments by playing sounds. Location-based information, gained knowledge about the user from the smartphone use and real-time data gathered from the Internet set the basis for successful context-awareness. The smartphone app Google Now[4] uses this context-knowledge to notify the user proactively about relevant information by presenting the information on the screen. E.g., when the user enters a subway platform he can see the schedule of the next trains leaving the station on his smartphone. Another example of a proactive situation-aware system is the Warning and Informationmanagement (WIM) system by Heisterkamp et al. [7], which ranks messages and warnings, which can occur while driving a car. Those warnings and messages are communicated to the user only in appropriate situations.

The overview of proactive systems shows that few projects or products exist, which address the employment of proactive SDSs in different environments. However, up to today, no SDS has been developed, which covers all stages of the presented speech interaction process and which satisfies the demands of proactive behavior in a user-friendly and situation-sensitive way. Research should investigate the desired interaction style of the different stages and the appropriate timing of the proactive interference.

CONCLUSIONS

This paper defined proactive behavior in SDSs and its influencing factors. The definitions are derived from other research fields and then transferred to SDSs. First, a general proactivity system model, which describes the relevant system components and their interaction has been described. A proactive system receives information from a knowledge source and notifies the user about an incoming event without a user request. The system has to act user-friendly and take the current user state and the environment into account. Thus, the proactive behavior can be identified as anticipatory, change-oriented and self-initiated. In a proactive human-machine speech dialog, first, the user has to be notified about an incoming event, then the problem solving process has to be started. Finally, the new task has to be completed and possibly paused tasks have to be resumed.

Future research should focus on investigating user-friendly speech interaction strategies for the 3 different stages of the proactive speech dialog flow. These investigations should take the influencing factors into account and consider different strategies for different situations.

REFERENCES

1. American Heritage Dictionaries Editors. *American Heritage Dictionary of the english language, fifth edition*. Houghton Mifflin Harcourt, 2011.
2. Barthelmess, P., and Ellis, C. A. The neem platform: An evolvable framework for perceptual collaborative applications. *Intelligent Information Systems 25*, 2 (2005), 207–240.
3. Clark-Dickson, P., Talmesio, D., and Sims, G. VoIP and IP messaging: Operator strategies to combat are threat from ott players (revised and updated). Tech. rep., Informa Telecoms & Media, 2013.
4. Edlund, J., C., O., and Gustafson, J. Investigating negotiation for load-time in the gethomesafe project. In *Proc. of IAST Workshop* (2012).
5. Frese, M., and Fay, D. Personal initiative (PI): An active performance concept for work in the 21st century. *Research in Organizational Behavior 23* (2001), 133–187.
6. Hamerich, S. Towards advanced speech driven navigation systems for cars. *IET Conference Publications 2007*, CP531 (2007), 247–250.
7. Heisterkamp, P., and Rothe, S. Daimlerchrysler mobile research environments layout, integration & paths to productization. Augmented Cognition Concept Validation Demonstration and PI Meeting, 2004.
8. Hermanutz, M. Entwicklung und evaluation von sprachdialogkonzepten im fahrzeug bei proaktivem systemverhalten. Master's thesis, Ulm University, 2013.
9. Merriam-Websters online dictionary. Proactive, 2013. Retrieved October, 15 2013, from http://www.merriam-webster.com/dictionary/proactive.
10. Minker, W., Lopez-Cozar, R., and McTear, M. F. The role of spoken language dialogue interaction in intelligent environments. *Journal of Ambient Intelligence and Smart Environments 1* (2009), 31–36.
11. Parker, S. K., and Collins, C. G. Taking stock: Integrating and differentiating multiple proactive behaviors. *Journal of Management 36*, 3 (2010), 633–662.
12. Parker, S. K., Williams, H. M., and Turner, N. Modeling the antecedents of proactive behavior at work. *Journal of Applied Psychology 91*, 3 (2006), 636–652.
13. Peissner, M., Doebler, V., and Metze, F. Can voice interaction help reducing the level of distraction and prevent accidents? meta-study on driver distraction and voice interaction. Tech. rep., Fraunhofer-Institute for Industrial Engineering (IAO) and Carnegie Mellon University, 2011.
14. Strauss, P.-M., and Minker, W. *Proactive Spoken Dialogue Interaction in Multi-Party Environments*, 1st ed. Springer Publishing Company, Incorporated, 2010.
15. Wickens, C. D., and Hollands, J. G. *Engineering Psychology and Human Performance*, 3rd ed. Prentice Hall, Upper Saddle River, NJ, 2000.
16. WordNet 3.1 Princeton University. Proactive, 2013. Retrieved October, 15 2013, from http://wordnetweb.princeton.edu/perl/webwn?s=proactive.

[4] http://www.google.com/landing/now/