CEPIS UPGRADE is the European Journal for the Informatics Professional, published bimonthly at <http://cepis.org/upgrade>

Publisher
CEPIS UPGRADE is published by CEPIS (Council of European Professional Informatics Societies, <http://www.cepis.org/>), in cooperation with the Spanish CEPIS society ATI (*Asociación de Técnicos de Informática*, <http://www.ati.es/>) and its journal Novática

CEPIS UPGRADE monographs are published jointly with Novática, that published them in Spanish (full version printed; summary, abstracts and some articles online

CEPIS UPGRADE was created in October 2000 by CEPIS and was first published by Novática and INFORMATIK/INFORMATIQUE, bimonthly journal of SVI/FSI (Swiss Federation of Professional Informatics Societies, <http://www.svifsi.ch/>)

CEPIS UPGRADE is the anchor point for UPENET (UPGRADE European NETwork), the network of CEPIS member societies' publications, that currently includes the following ones:

- inforewiew, magazine from the Serbian CEPIS society JISA
- Informatica, journal from the Slovenian CEPIS society SDI
- Informatik-Spektrum, journal published by Springer Verlag on behalf of the CEPIS societies GI, Germany, and SI, Switzerland
- ITNOW, magazine published by Oxford University Press on behalf of the British CEPIS society BCS
- Mondo Digitale, digital journal from the Italian CEPIS society AICA
- Novática, journal from the Spanish CEPIS society ATI
- OCG Journal, journal from the Austrian CEPIS society OCG
- Pliroforiki, journal from the Cyprus CEPIS society CCS
- Tölvumál, journal from the Icelandic CEPIS society ISIP

Cover page designed by Concha Arias-Pérez
"Devourer of Fantasy" / © ATI 2011
Layout Design: François Louis Nicolet
Composition: Jorge Llácer-Gil de Ramales

Editorial correspondence: Llorenç Pagés-Casas <pages@ati.es>
Advertising correspondence: <info@cepis.org>

Subscriptions
If you wish to subscribe to CEPIS UPGRADE please send an email to info@cepis.org with 'Subscribe to UPGRADE' as the subject of the email or follow the link 'Subscribe to UPGRADE' at <http://www.cepis.org/upgrade>

ISSN 1684-5285

Monograph of next issue (April 2011)
"Software Engineering for e-Learning Projects"
(The full schedule of **CEPIS UP**GRADE is available at our website)

CEPIS UPGRADE
The European Journal for the Informatics Professional
http://cepis.org/upgrade
Vol. XII, issue No. 1, February 2011

Monograph
Internet of Things
(published jointly with Novática*)
Guest Editors: *Germán Montoro-Manrique, Pablo Haya-Coll, and Dirk Schnelle-Walka*

UPENET (UPGRADE European NETwork)

CEPIS NEWS

* This monograph will be also published in Spanish (full version printed; summary, abstracts, and some articles online) by **Novática**, journal of the Spanish CEPIS society ATI (*Asociación de Técnicos de Informática*) at <http://www.ati.es/novatica/>.

Ubiquitous Explanations: Anytime, Anywhere End User Support

Fernando Lyardet and Dirk Schnelle-Walka

In this article we present a new approach for supporting people in their daily encounters with technology. Be it at home, in public spaces or on the road, people often need to solve little puzzles when facing new technology in order to carry out a task. Furthermore, help is seldom available or is hard to understand. In this article we present a new approach for delivering ubiquitous assistance through voice-controlled, multimodal explanations. We take advantage of SIP-based communications on personal, trusted mobile devices such as mobile phones to allow users to access explanations anywhere, regardless of the multimedia capabilities available on-site. We also present a modelling approach and software support to adapt explanations to different user knowledge levels, which enables people to adjust dynamically the granularity of instructions they receive.

Keywords: Computing, End-user Assistance, Prototype, Smart Products, VoIP.

1 Introduction

Assisting people facing new situations and technologies is a complex problem. It involves different aspects such as how relevant and easy to understand the instructions are, the modalities used to deliver the assistance, and limitations in real-world scenarios.

Good explanations are far more than a simple, correlated set of sentences. The content, style, and information required to make instructions effective must meet a series of prerequisites. Also, the granularity of instructional support is key. For assistance to be effective, there must be a balance between the information provided or *"externalized"* and the cognitive process (or *"internalization"*) developed by a user figuring out how to complete a task. The wrong balance of instructional support limits the user's ability to optimize [1], comprehend, and ultimately learn [2].

Traditionally, user assistance has been delivered in the form of printed manuals. However, manuals are seldom used and, when people have to resort to printed materials, they tend not to like them. Instead, many users confronted with a new device will tinker a few minutes with it, and if unsuccessful, they may return a perfectly functioning product as "broken" [3]. A similar phenomenon is known as *"The Paradox of the Active User"*[4]: many people are not interested in learning but rather in using the artefact right away. It is a paradox because reading the instructions would save them time. Printed documentation does have limitations that may render it impractical in many situations:

- it is linear, where tasks are to be carried out in only one sequence,
- it does not address situations where users get stuck and need specific, short guidance on how to continue [5],
- the information available is the same regardless of user expertise.

However, people still need instructions. And people need more instructions today than ever before: embedded electronics increasingly expands and introduce new functions that are either more silent or simply more abstract (such as programming features), that cannot be perceived or understood by simply watching them in operation. This makes machines increasingly opaque to our senses, and it is harder for people to produce an explanation on how a product works and what it is currently doing [3].

In this article we first introduce previous developments that show end user assistance Ubicomp technologies as a key aspect can help solve. We then explain what the requirements for good explanations are, and the current standards in the industry to motivate and introduce a novel modelling approach for adaptive instructional knowledge. This new model is complemented with technology support to allow users to request more detailed information when they need it, or skip information they already know. We then present an architecture and implementation for delivering multimodal instructional assistance to mobile devices such as mobile phones.

Authors

Fernando Lyardet obtained his MSc in Informatics from the *Universidad Nacional de La Plata* (UNLP), Argentina, where he was working as a Research Assistant. After that he joined the Telecooperation group as a PhD candidate at the Darmstadt University of Technology, Germany, and SAP Research, Germany. His research interests include web engineering, smart environments, and ubiquitous computing. He is a member of ACM and AAAI. <fernando@tk.informatik.tu-darmstadt.de>

Dirk Schnelle-Walka received his PhD in 2007 in the area of voice user interface design at Technical University of Darmstadt, Germany. He transferred the idea of design patterns to the design of voice user interfaces. His research focus is on voice-based interaction in pervasive environments. He is also the head of several open source projects on speech technology, e.g. the open source voice browser JVoiceXML.<dirk@tk.informatik.tu-darmstadt.de>

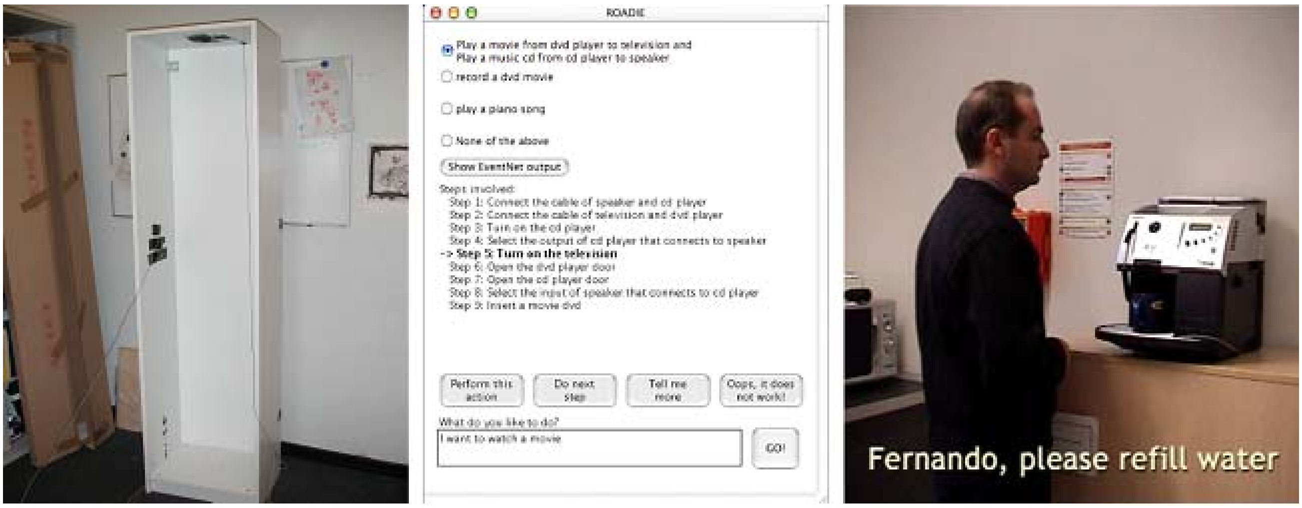

Figure 1: Sample Scenarios of End-user Assistance. (From left to right: proactive furniture assembly [5], home appliances troubleshooting [6], and a smart coffee machine [7].)

2 Related Work

In recent years, several attempts have been made to embed instructions in the objects themselves, together with sensors and software in the tools and products themselves. These ideas seek to provide just-in-time instructions for immediate assistance. One example are printers and photocopiers that display an error status and plausible recovery actions that users can take to solve the problem. In Figure 1 three trailblazing projects are displayed: (i) first, a proactive support for furniture assembly [5]. It takes advantage of embedded sensors to recognize the actions a user takes when building a simple artefact like a piece of Ikea furniture, analyzes the current situation, and automatically generates alternative paths to find a solution taking into account the current building state. (ii) Second, Roadie, which collects information about surrounding appliances and devices, and using AI reasoning techniques is able to figure out possible actions a user could take to solve a problem. (iii) Finally a smart appliance [7] that helps a user through complex maintenance procedures combining voice and visual modalities.

While the focus of [5][6] is on solution-finding based on actual context, [7] explores how to guide the user on a complex process. However, all of these prototypes rely on locally available infrastructure (displays, sound speakers) to deliver the instructional design. This may work well in indoor scenarios such as the home or office, but not so well when on the move or in industrial settings.

> “Assisting people facing new situations and technologies is a complex problem”

3 Requirements for Good Explanations

Writing and designing explanations must meet several requirements in order to be effective. Two kinds of current approaches can be identified: (i) industrial standards such as DIN 8418 [8], and (ii) Linguistics studies [9]. Both approaches stress that user information should be considered as a part of the product, readily available in text or iconic form. As an industrial norm, it also provides some broad recommendations on what explanations should be like when delivered as printed manuals:

- The form and degree of detail can vary depending on the product itself and the expected knowledge of the users. Descriptions should be easy to understand.
- Texts and figures should be ordered as if the product were to be used for the first time.
- Clear structuring of the sections in tune with the way users think.
- They should answer predictable questions such as *why* and *what*.
- Descriptions should be short and precise while focusing on the core issues.
- Any jargon used should be explained.
- The names of the section titles should be problem- and usage-oriented.
- Images can be used to support the text and should be easily related to the corresponding text passages.
- The use of special icons or abstractions should be explained.
- It should also be easy to read which can be achieved by summaries in tables, colours, a clear layout and a suitable font.

Other approaches to good explanations can be found in the domain of linguistics. Schmidt analyzes in [9] the criteria for helpful user information. She demands that user information should be understandable, complete, correct, short, precise, simple and should have a consistent representation. The most important requirement is that users are able to understand the contained information. This also includes to avoid jargon in favour of simple language constructs that can be understood by a wider variety of user groups. Jargon allows people to describe something in a short and precise fashion, but may not be understood by everybody. User information should aim at an active usage of the product.

Furthermore, it should not be absolutely necessary for users to have a detailed understanding of the processes and the reasons for their actions. The increasing functionality of devices leads to a less detailed knowledge of understanding. For example, nearly everybody can drive a car without needing to understand how the engine really works. The level of understanding depends on the sender of the information but also on the recipient. Depending on the background knowledge, different users will have different knowledge about details. Too much information will cause confusion for users with only a little background knowledge.

User information also has to be brief and precise. This implies a trade-off between being brief on the one hand but not to lose information in terms of being complete on the other hand. It is also necessary to repeat information that is important for the safe handling of the product. Images are a suitable means to help in these cases. Images should not repeat the information that is given in the verbal description but enhance it or give an overview.

4 A Model for Adaptive Explanations

In the previous section it was shown that the user's experience and expertise are key factors for designing good explanations. Today, manufacturers need to define which group of users should be supported and how much explicit information is given (externalized) or left to the user to understand (internalize). In both cases, once defined the information is mostly fixed without offering the possibility to adapt to the user and his or her preferences. In this section we describe a model for adaptive explanations. From concurrent task trees [10] we borrow the representation of goals to achieve by means of a tree. The depth of the tree represents the knowledge of the user. The deeper the tree is traversed the less experienced the user is (See Figure 2).

> “When people have to resort to printed materials, they tend not to like them”

5 Prototype

For our demo implementation we limited the experience levels to expert, advanced, intermediate and novice. Technically, there is no limitation for the levels and it can be easily adapted to a custom set of experience levels. Each node represents an activity of a workflow that is being followed to achieve the user's goal. This means that, in an ideal case, a single step is sufficient to guide expert users.

5.1 Profiles and Perception Level

According to Lyardet et al. [11], an explanation is de-

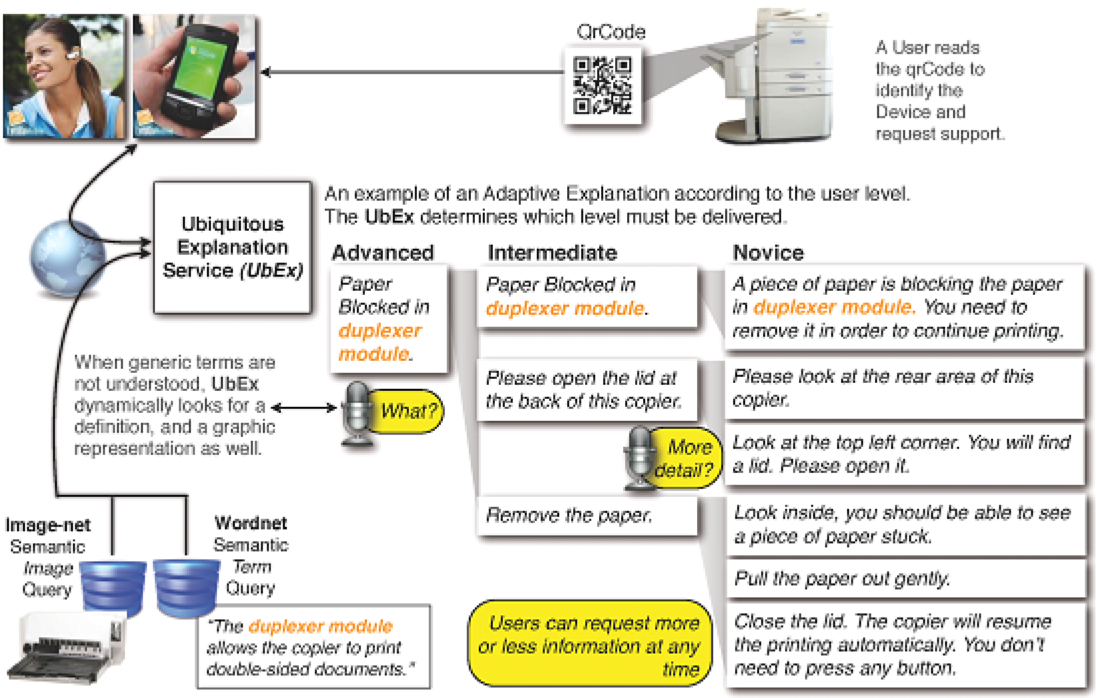

Figure 2: Example of an Explanation for Different Skill Levels.

“Writing and designing explanations must meet several requirements in order to be effective”

scribed by the following criteria: detail level, perception level and visual information. The detail level is directly related to the requirement to have short and precise explanations as described above. A shorter explanation has a higher cognitive load for the user than longer and more detailed explanations. Jargon can also be used to shorten the explanation. This can be stored in a profile. Linguistic factors as they are described above are always important since, for example, a setting of 90% correct information does not make much sense. All factors are rated by a perception level with a degree from 0 to 100. Perception levels are related to a sense: seeing (e.g. colour, shape, volume), hearing (e.g. frequency, duration) and touching (e.g. kind of surface). The types and values of the different perception levels can be used to find representative values that can be used to determine the relevant explanation. This can be done with the help of the Multi-Attribute-Utility-Theory (MAUT) [12]. MAUT transforms the different factors into a multi-dimensional space which is integrated with the help of a weighted sum function.

The profile is used to determine how the system communicates with the user. In an ideal case, there is no further adaption needed. However, the environment may influence how the user really perceives the presented information. For instance, if he or she is tired the user may not be able to concentrate on the given explanation. Hence, skill adaption will sometimes be necessary.

5.2 Skill Adaptation

There are three options to adapt the skill level of the user: direct, time-based and event-based. Direct means an explicit action by the user requiring more or less information. The latter two are dynamically initiated by the system. If the user does not react within a predefined time-span it can be assumed that he or she is either confused or did not understand the given explanation. The first time this happens, the explanation may be repeated. An escalation decreases the skill level at the second occurrence. The third alternative is triggered by sensors detecting that the user performed the wrong actions. This is not always possible since not all devices feature sensors. Skills are adapted event-based using the same escalation strategy as for time-based adaption. This corresponds to an established technique of voice user interface design that Cohen names *Detailed Escalation* [13].

The options are not exclusive but are based on each other and are executed in the described order.

5.3 Dynamic Explanations

The concepts introduced so far rely on a tasktree that is known in advance. In this case the required knowledge has been provided by the manufacturer. However, real world scenarios may prove that the given information is not sufficient to be used, say, by novice users. In this case a more dynamic approach is needed to create adapted user information on the fly. In a mobile setting the additional information can be acquired via, say, a wireless network to adopt additional knowledge. A common way to share knowledge is through the use of ontologies. In the following section we will explain their use by the example of tools that are needed to correct a blockage of a copying machine, for example. Imagine a user who is confronted with a term that is well established in the manufacturers vocabulary and used throughout the explanations because the manufacturer cannot imagine that someone does not know the correct name for the tool to use. On the other side, the user is not able to perform the action because the user neither understands what he or she is supposed to do nor how to handle the tool.

We use WordNet, <http://wordnet.princeton.edu/>, for this purpose which features a good access to information. A search for *screwdriver,* for example, returns both types of information we are interested in: (1) a short description of the tool, (2) a description of how to handle it. Other sources of information, like ImageNet <http://www.image-net.org>, can also be used to acquire images of the search. Searches are performed by means of ontologies. These information sources are integrated into the Ubiquitous Information to provide further help on demand.

6 Ubiquitous Explanations

In this section we introduce *Ubiquitous Explanations*, a new approach for supporting users wherever they interact with technology. This approach takes advantage of current advances in mobile platforms and communications to deliver instructional assistance. It borrows from ubiquitous computing concepts of pervasively available computing power and communications.

However, unlike previous approaches, we focus on developing a mechanism to be as widely available as possible, across different devices. Therefore we tap into the communications and computing power carried by people in the form of mobile phones, and music and game consoles. A centrepiece of this approach is our use of SIP (Session Initiated Protocol [14]) -based communication between the user terminal and the Ubiquitous Explanations Service (UbEx). SIP enables UbEx multimodal communication of

“Too much information will cause confusion for users with only a little background knowledge” voice and data to and from the user. UbEx interprets voice and user actions as commands or queries, and the answers (in this case, the required explanations) are delivered to the user with data (text and graphics) and voice (TTS generated).

In the following sections we introduce the requirements and architectural approach implemented for Ubiquitous Explanations and a detailed description of the architecture.

6.1 Requirements and Assumptions

Providing support to help people with the technology they encounter is based on the following assumptions:

An ME (Minimal Entity): *ME* devices are a representation of their users in the digital world [15]. A small wearable computer easy to carry around that provides a "minimal" key functionality: identity, security, association, interaction, context awareness, and networking. Recent generations of mobile phones enable a true mobile Internet experience and have become today's physical token for our "Digital Me". Most modern mobile phones today, and advanced personal multimedia players like the iPod Touch can fulfil the role of a *ME* entity.

Speech-based: to allow users a simple and natural interaction mechanism and also be useful in situations where eyes-free, hands-free support would be desirable.

Zero Deployment: today mobile devices are pervasive but their execution capabilities and requirements differ greatly. Therefore, to provide a scalable solution, the solution must rely on the capabilities available in the mobile "off-the-shelf". Therefore, we selected a SIP-based solution for delivering voice communication and a standard web browser to deliver visual content.

Sensor Feedback: On many occasions, such as multi-step procedures, it is desirable to receive some kind of confirmation that a prescribed action has been carried out in order to continue. If available, such feedback allows the system providing a more natural guidance, without further request for the user.

Flexible Explanations Support: specific support is required to allow modelling explanations for users with different skill level.

6.2 Architecture

Ubiquitous Explanations Service (UbEx) can support the user in different stages of a smart product's lifecycle, and therefore, there are different ways in which assistance can be triggered. An example of a typical assistance scenario is depicted in Figure 3, where the user encounters a problem and requests further assistance: the user would first associate his or her *"ME"* device (typically a mobile phone), by simply scanning the qr-code available in the product.

The Ubiquitous Explanations software establishes a SIP link that sends and receives commands from both the user terminal and sensor feedback -if available- from the target device. A second link (RTP-based) is also established, in order to provide support for voice communication between the user and UbEx. Whenever the user speaks or triggers a command in his user terminal or by activating the target device, this signals are forwarded to a *Controller* module, that further delivers the request to the *Workflow Adapter*

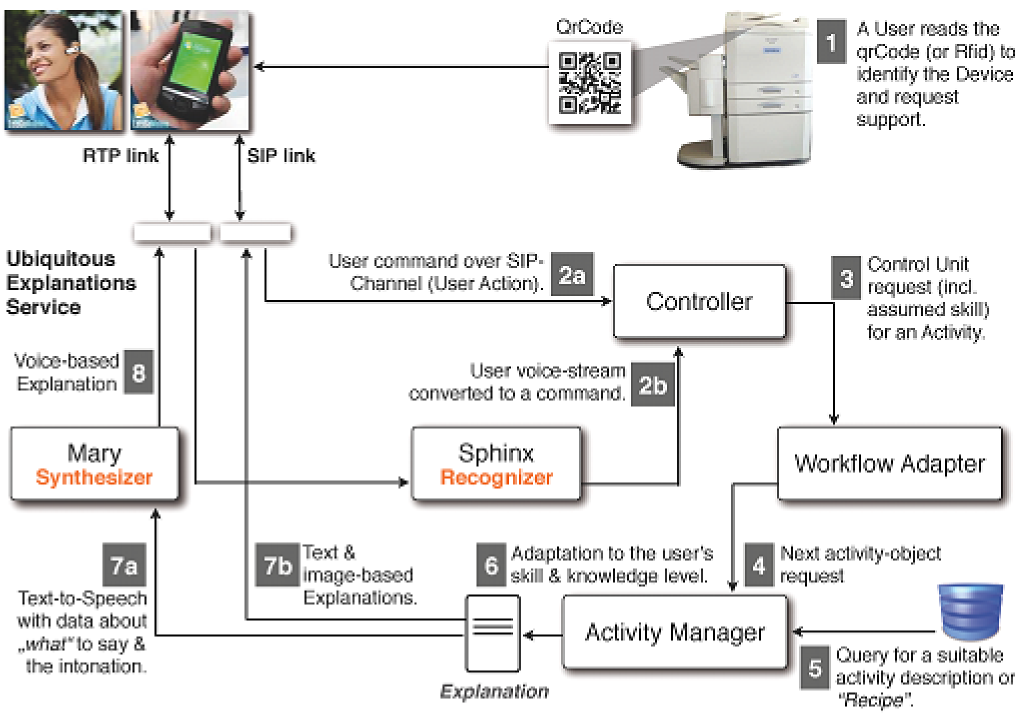

Figure 3: Ubiquitous Explanations Architectural Overview.

> Today mobile devices are pervasive but their execution capabilities and requirements differ greatly

controlling the explanation process. If the user has given a spoken command, the signals will then be delivered through the RTP link and analysed and recognized using an open source speech recognizer called Sphinx [16].

The *Activity Manager* also plays a key role, since it is in charge of determining what level of information should be delivered to carry out the next task (communicated by the workflow module). It fetches the required information from different sources and assembles the explanation that must be delivered to the user. Once the explanation has been defined, it is established which contents are to be delivered via voice, visual modalities or both.

The information to be delivered over voice modality is then forwarded to a *Synthesizer* (Mary TTS [17]) together with extra information regarding the intonation in which the explanation is to be delivered. the intonation information is important, for instance, to help users disambiguate potentially hazardous actions from simple steps. The explanation information to be delivered in visual modality is pushed to the user terminal through the SIP link.

7 Conclusions

While previous approaches have been successful in the proposed scenarios, delivering just-in-time instructions to any product has remained largely unsolved. Providing explanation assistance is challenging and includes several aspects: explanation generation, explanation adjustment to the receiver's knowledge, modalities of explanation delivery, and technology requirements to receive and control the explanation being delivered. On the one hand, conveying different degrees of information according to each user's needs requires special modelling support. On the other, many assistance technologies require some kind of specific hardware to be present in order to deliver the assistance. This requirement limits the locations where explanations can be provided to where such hardware is. Roadie embedded sensors in things can help identity what is going on, and provide better feedback. Current limitations are that such kind of embedded functions, is that the mechanism for delivering such explanations is one-of a kind. We have also experimented with public/open speakers. Such solutions are all right, but require extra local hardware as well.

> *Ubiquitous Explanations* is a new approach for supporting users wherever they interact with technology

Several scenarios are available, most of them in rather similar settings from a Voice User Interface perspective: indoors, TTS delivered over speakers, and custom settings for a given scenario. In this example, we want to provide a mechanism where assistance, not only computing power, can also be ubiquitous, and permit opportunities for opportunistic learning. For instance, when repairing a car or changing a wheel. These are scenarios where such "common speakers" would not be available.

References

[1] S. Payne, A. Howes, W. Reader. Adaptively distributing cognition: A decision making perspective on human-computer interaction. Behaviour and Information Technology 20 (5, 2001), pp. 339-346.

[2] C. Nimwegen, D. van, Burgos, H. Oostendorp, H.J.M. van & Tabachneck-Schijf. The paradox of the assisted user: Guidance can be counterproductive. CHI Proceedings - Awareness and Presence, pp. 917–926, 2006.

[3] Donald A. Norman. Emotional Design: Why We Love (or Hate) Everyday Things. Basic Books, 1st edition, 2003. ISBN-10: 0465051359.

[4] J. Carroll, M. Rosson. The Paradox of the Active User. MIT Press, 1987.

[5] S. Antifakos, F. Michahelles, B. Schiele. Proactive instructions for furniture assembly. In Ubicomp. (2002), pp. 351-360.

[6] H. Lieberman, J. Espinosa. A goal-oriented interface to consumer electronics using planning and commonsense reasoning. In J. Krumm, G.D. Abowd, A. Seneviratne, T. Strang eds.: Proceedings of the 11th international Conference on intelligent User interfaces (Sydney, Australia, 29th January - 1st February 2006) IUI '06. Volume 4717 of LNCS, ACM Press, pp. 226-233.

[7] E. Aitenbichler, F. Lyardet, G. Austaller, J. Kangasharju, M. Mühlhäuser. Engineering Intuitive and Self-Explanatory Smart Products. Proceedings of the 22nd Annual ACM Symposium on Applied Computing, ACM Press, 2007, pp. 1632-1637.

[8] Deutsches Institut für Normung e.V. Norm 8414: *Benutzerinformation – Hinweise für die Erstellung* (February 1988).

[9] U. Schmidt. *Gebrauchsanweisungen - Form und Struktur*. Ph.D.Thesis, *Rheinische Friedrich-Wilhelms-Universität zu Bonn*, 1996.

[10] F. Paternò, C. Mancini, S. Meniconi. Concurtasktrees. A diagrammatic notation for specifying task models. INTERACT '97: Proceedings of the IFIP TC13 International Conference on Human-Computer Interaction, London, UK, Chapman & Hall, Ltd, 1997, pp. 362-369.

[11] F. Lyardet, D.W. Szeto, E. Aitenbichler. Context-aware indoor navigation. AmI '08: Proceedings of the Euro-

pean Conference on Ambient Intelligence, Berlin, Heidelberg, Springer-Verlag, 2008, pp. 290-307.

[12] R. Schäfer. Rules for using multi-attribute utility theory for estimating a user's interest. Technical report, DFKI, 2001.

[13] M.H. Cohen, J.P. Giangola, J. Balogh. Voice User Interface Design. Addison-Wesley, Boston, January 2004. ISBN-10: 0321185765.

[14] Network Working Group. IETF RFC 3261. SIP: Session Initiation Protocol.

[15] E. Aitenbichler. System Support for Ubiquitous Computing. Ph.D.Thesis, Darmstadt University of Technology, 2006.

[16] Carnegie Mellon University. CMU Sphinx: The Open Source Toolkit for Speech Recognition.

[17] DFKI. The German Text-to-Speech Synthesis System MARY.

[18] S. Palmiter, J. Elkerton. An evaluation of animated demonstrations of learning computer-based tasks. CHI '91: Proceedings of the SIGCHI conference on Human factors in computing systems, New York, NY, USA, ACM, 1991, pp. 257-263.

