

DIRECT AND INDIRECT USER-TO-DEVELOPER MESSAGES THROUGH COMMUNICABILITY EVALUATION

Simone D. J. Barbosa

Raquel O. Prates

Clarisse S. de Souza

Departamento de Informática, PUC-Rio
Rua Marquês de São Vicente, 225
Rio de Janeiro, RJ – Brazil – 22453-900
{sim, raquel}@les.inf.puc-rio.br, clarisse@inf.puc-rio.br

INTRODUCTION

The Semiotic Engineering Research Group at PUC-Rio emerged as a spontaneous organization of graduate students, researchers, and professors in 1996. A growing interest of the HCI community in topics related to Semiotics led us to cast HCI problems in semiotic terms. Our first attempts at it date back to 1990, and we have been since exploring the answers to these problems. Research and industrial partners, in Brazil and abroad, have recently joined us, and are contributing to build critical mass in Semiotic Engineering.

Unlike other approaches in Computer Semiotics, our goals thus concentrate on HCI issues only. They are both theoretical and practical. We are trying to provide designers with theoretical foundations based on Semiotic Engineering, for the design, development, and evaluation of computer artifacts. For HCI practice, we are investigating semiotically-based methods and techniques, and developing tools. Our interest in representational support for user-developer communication stems from its counterpart in Semiotic Engineering – the fact that systems can be viewed as one-shot messages from designers to users about how to exchange messages with the system in order to achieve specific goals. We have proposed an evaluation method that allows for direct and indirect communication from user to designers, depending on who performs the different evaluation steps.

In this following we present our communicability evaluation method. Communicability is the property of software that efficiently and effectively conveys to users its underlying design intent and interactive principles. Communicability evaluation is carried out in three major steps, that can be performed by different groups of people (users, designers, and HCI experts). It yields two distinctive types of representations about interaction, which tell us something about user-system interactive patterns, and designer-to-user (intentional or unintentional) communication.

When steps of communicability evaluation are performed by users themselves, they can serve as an insightful means of user-to-developer message. In it, users can express their expectations, attitudes, interpretations, approval or rejection towards HCI design choices implemented in software. When performed by designers or experts, they should be perceived as an inferred message about the same topics, qualified by the analyst's background and expertise in HCI.

Communicative evaluation can be done at different stages of design, and serve different targets. In formative evaluation, it can help designers compare design alternatives and make further design decisions. In summative evaluation, it can inform the changes needed in new releases. Compared to other evaluation methods, ours focuses on the signs, structures and conversational patterns presented to users at the interface level, signalling the immediate interpretations assigned to them, and the role they play in user-system and designer-to-user communication.

Simone Barbosa and Raquel Prates have a PhD in Computer Science from PUC-Rio. They are active researchers at SERG, and have additional professional experience as developers (Simone Barbosa) and HCI analysts and teachers (both). Clarisse de Souza has a PhD in Linguistics, also from PUC-Rio. She is an Associate Professor at the Informatics Department, where she has been carrying out teaching and research activities in HCI since 1988. As part of our research, we have developed the communicability evaluation method presented here, and we are currently assessing its value in different contexts, such as academic research projects, commercial design and redesign projects, as well as in HCI training activities.

A COMMUNICABILITY EVALUATION METHOD

Our method is based on Semiotic Engineering (de Souza,1993), which views systems as one-shot messages (about anticipated users' needs and expectations) from designers to users. These messages can send and receive other messages (with which users try to meet their actual needs and expectations). In this perspective, a system is a kind of "discourse deputy" for the designer in that it encompasses the set of all (and only) the conversational turns and themes she can predict at design time. Users are thus communicating with the designers' deputies, but not with designers themselves. The

communicability evaluation method described in this paper provides a way for users to communicate with designers, although indirectly, about interaction breakdowns (potential and actual).

Communicability evaluation consists of three major steps: tagging, interpretation, and profiling.

TAGGING

Tagging is the association of spans of interactions to the (hypothetical) user utterances, identifying potential conversation breakdowns. The users' actions for accomplishing a pre-defined task must be recorded using software that is able to capture mouse-pointer movements, screen events, etc. (e.g. Lotus® ScreenCam™).

This step may be performed by users, designers, and experts alike. When performed by users, we capture a spontaneous response to interaction patterns, classified as one of the available utterances, as a kind of constrained verbal protocol. When designers do the tagging, they recognize the interactive breakdowns that users experienced and classify them accordingly. Finally, when HCI experts do the tagging, they put words in the users' mouth, in an attempt to recreate a verbal protocol. Designers and experts could also videotape and take notes of users' behavior during test sessions, and use this material for occasional disambiguation during the tagging process.

The following set of utterances can be associated to different sets of breakdown situations and user attitudes that are likely to occur during human-computer interaction (de Souza et al., 1999), and to underlying HCI design problems:

- **Where is? (What now?)**
The user seems to be searching for a specific function, but demonstrates difficulty in locating it. So, he sequentially (worse case) or thematically (better case) browses menus and/or toolbars for that function, without triggering any action. This category includes a special case we have called What now?, applicable when a user does not know what to do and thus searches for the next step (as opposed to the next function).
- **What's this? (Object or action?)**
The user seems to be exploring the possibilities of interaction to gain more (or some) understanding of what a specific function achieves. User lingers on some symbol waiting for a tooltip and/or explicitly calls for help about symbol or hesitates between what he thinks are equivalent options. This category also includes cases in which users are confused about widgets being associated to objects instead of actions and vice-versa (Object or action?).
- **Oops! (Where am I?)**
This category encompasses cases in which users perform some action to cause a specific state of affairs, but the outcome is not the expected one. The user then either immediately corrects (typically by pressing UNDO or by attempting to restore some previous state of interaction otherwise) or completes the task with an additional sequence of actions. Oops includes a special category Where am I? in which the user performs some action that is appropriate in another context, but not in the current one.
- **Why doesn't it? (What happened?)**
This category is related to cases in which users expect some sort of outcome but do not achieve it. One subsequent scenario is that they then insist upon the same path, as if they are so sure that that function should do what they expected that they cannot accept the outcome. They carefully step through the path again and again to check that they are not doing something wrong. The other scenario (What happened?) is when they do not get a feedback from the system and are apparently unable to assign meaning to the function's outcome (halt for a moment).
- **Looks fine to me...**
The user achieves some result he believes is the expected one. At times he misinterprets feedback from the application and does not realize that the result is not the expected one.
- **I can't do it.**
The user is unable to achieve the proposed goal, either because he does not know how to or because he doesn't have enough resources (time, will, knowledge, and the like) to do it.
- **Thanks, but no thanks. (I can do otherwise.)**
The user defers some affordance provided by the application's interface and finds another way around the problem. If it happens after successful attempts but the user switches to other strategy, then it is a case of Thanks, but no thanks. If not (for instance, if it happens after a series of unsuccessful attempts), it is a case of I can do otherwise.

INTERPRETATION

This step consists of tabulating the gathered data and mapping the utterances onto HCI problems. If the mapping is pre-defined, this step may be achieved by a non-expert. For instance, designers can benefit from this step by obtaining a rough diagnosis of interaction problems. We have associated the seven categories to generic classes of interaction and usability problems, as shown in Table I.

	Navigation	Meaning Assignment	Task Accomplishment	Declination of Affordance
Why doesn't it? (What happened?)				
Where is? (What now?)				
What's this? (Object or action?)				
Oops! (Where am I?)				
Looks fine to me...				
I can't do it.				
I can do otherwise. (No, thanks)				

Table I: Mapping conversational categories onto interactive and usability problems

Notice that navigation, meaning assignment and task accomplishment are known usability problems. However, communicability evaluation reveals another class of problems: declination of affordance. Popular sets of design guidelines or usability principles do not address this problem explicitly, neither do the cognitively-based evaluation methods and techniques. Nevertheless, in communicability evaluation, this phenomenon stands out as considerably frequent, and refines HCI problems taxonomies. Users generally decline an affordance when they perceive a disadvantageous cost/benefit ratio for an afforded feature as compared to an alternative way of performing the same task. Among the causes of declination of affordance we point out inconvenient navigational structures, such as deep nesting in menu structures, lack of shortcuts, and the like.

HCI experts, however, may use alternative taxonomies for a more precise user-system interaction diagnosis. For instance, utterances may be mapped to problems of abiding to Nielsen's discount evaluation guidelines (Nielsen,1993), or Shneiderman's 8 golden rules (Shneiderman,1998), or supporting the bridging of Norman's gulfs of execution and evaluation (Norman,1986), or even Sellen and Nicol's taxonomy of contents for building online help (Sellen and Nicol,1990).

PROFILING

Profiling must be done by communicability experts. This step consists of interpreting the tabulation in semiotic terms, in an attempt to retrieve the original designer's meta-communication, i.e., the meaning of the designer-to-user message. What we are dealing with here are the implicit messages conveyed by the choice of signs, structures, and interactive patterns that compose the user interface. These messages may be intentional or unintentional. Unintentional messages are the result of designers' tacit knowledge and assumptions. The role of the expert is to reveal these implicit messages to designers, who may then change or confirm their choices.

Communicability evaluation provides at least two different instances of representation that can be used in user-developer communication: tagged movies and communicability profiles. Tagged movies represent qualified human-computer interaction breakdown instances. When the same task is tagged by different people, biases and/or refinements in interpreting users actions can be captured by the contrasts among taggings. Therefore, the usage of this first representation is twofold: it can serve as a characterization of HCI breakdowns from somebody's point of view, and also as a common ground for discussing intentional, unintentional, and missed communicative acts in HCI.

Communicability profiles are germane to the semiotic engineering approach. They represent elements of the overall designer-to-user communication that influence the perception and reaction of users in face of software artifacts. Compared to taggings, communicative profiles diagnose fine-grained communication phenomenon, desired or undesired. When done by more than one expert, as with taggings, biases and finesses can be accounted for.

CONCLUDING REMARKS

As can be seen from the above, in communicability evaluation user-developer communication is achieved in different ways: (a) designers have access to users' interaction through the movies, and can recognize problems that occurred; (b) designers may have access to users' taggings; and (c) indirect communication, mediated by the HCI expert, who will build the profiles, pointing out taggings and interpretations in terms of HCI design problems. Moreover, taggings provide a common language in which users, designers, and HCI experts can share their knowledge.

Communicability evaluation can be used at different stages of the design process. At early stages, it can serve as a formative evaluation tool, allowing designers to compare different design options or assess the choices they have made. In particular, our method can be used as a instrument for inspection evaluation, since designers and experts may try to put themselves in users' shoes and tag interaction breakdowns and potential breakdowns. At later stages, it can be used as a summative evaluation tool to inform the features to be changed or added in future software releases.

Our method applies basically to single-user interfaces. Multi-user interfaces would probably require other utterances related

to interacting with other users, such as Who are you?, What are you doing?, and Where are you?. The same is true of artificial intelligence applications, for which utterances related to the system's cognitive abilities are likely to occur (e.g. Do you know this?, Can you learn this?).

Compared to other evaluation methods (Preece et al.,1994), ours is clearly focused on what is being said by the interface signs a user is supposed to interpret. Thus, problems related to inadequate task or user modeling are not directly addressed by us, except perhaps as a further inference on why certain communication is conveyed. By the same token, other methods typically do not directly address the problems we deal with. For instance, failure to provide feedback (e.g. Shneiderman,1998 or Nielsen,1993), may cause differentiated taggings to emerge (e.g. Why doesn't it...?, What's this?, I can do otherwise). The effect of differentiation can be noticed in redesign tasks or online help design, for example. Designers can use the tagging to decide on which sign or message they will incorporate to the application so that the problem is solved or minimized.

Sequences of utterances may additionally provide relevant insights about how users interpret the designers' messages. For instance, a Where is...? utterance may be often followed by a I can do otherwise utterance. In this case, it probably indicates that the user fails to perceive some afforded feature and thus that the designer is not getting the message across. Also, a sequence of Thanks, but no thanks utterances is likely to indicate a mismatch between the designer's ideal user and the actual user who is participating in the evaluation.

Some open issues challenge us to proceed with our research about communicability evaluation. Among them, we can single out the following:

- is this set of utterances appropriate? is it technology-dependent? is it culturally determined?

If we allow people to tag movies with *other* utterances (in addition to the ones in our set), we may sense if the set we are working with is satisfactory for the analysis or not. The same applies to specialized technologies, such as multi-user applications or AI-based systems. Finally, it may be the case that different cultures react in different ways to the same communicative acts (even without translation problems). The latter may be particularly interesting for software localization.

- what is the spectrum of taggings that can be done to the same movie by different groups of people (users, designers, and experts)?

We shall soon have the results of a case study in which interaction with a small application will be tagged by different groups of people. By contrasting the taggings, we expect to assess the range of plausible interpretations the same phenomena can yield, which is seldom achieved with other evaluation methods.

- how do utterances change along the users' learning curve?

An ongoing study with 6 participants using the same software over a period of 10 weeks will provide us with data for an appreciation of the types of utterances novices are likely to make, compared to more experienced users. For instance, Where is...? can be expected to disappear, or at least to come down to an insignificant number of occurrences during a task. On the other hand, Thanks, but no thanks might become more frequent over time.

REFERENCES

- de Souza, C. S. (1993). "The Semiotic Engineering of User Interface Languages". *International Journal of Man-Machine Studies*, **39**, 753-773.
- de Souza, C.S.; Prates, R.O.; and Barbosa, S.D.J. (1999). "A Method for Evaluating Software Communicability". In C.J.P. de Lucena (ed.) *Monografias em Ciência da Computação*. PUC-Rio Inf MCC11/99. Computer Science Department, PUC-Rio, Brazil.
- Nielsen, J (1993). *Usability Engineering*. Academic Press.
- Norman, D. (1986). "Cognitive Engineering". In Norman, D. A. and Draper, S. W. (Eds.) *User-Centered System Design*. Lawrence Erlbaum. 411-432
- Preece, J., Rogers, Y., Sharp. H., Benyon, D., Holland, S. and Carey, T. (1994). *Human-Computer Interaction*. Addison-Wesley.
- Sellen, A. and Nicol, A. (1990). "Building User-entered On-line Help". In B.Laurel (Ed.) *The Art of Human-Computer Interface Design*. Addison-Wesley.
- Shneiderman, B. (1998) *Designing the User Interface*. Addison-Wesley.