An Empirical Evaluation of an Adaptive Web Site

Cristina Gena

Dipartimento di Informatica, Università di Torino
Corso Svizzera 185, 10149 Torino, Italy
+39 011 6706711
E-mail: cgena@di.unito.it

ABSTRACT

This paper describes the evaluation of an adaptive commercial web site offering a set of utilities tailored on the basis of user needs. We compared the site with the non-adaptive variant in order to study how the adaptivity increases the success in retrieving information and reduces the amount of actions needed to solve the tasks. Moreover, we considered the preference towards the two alternative versions and the user satisfaction.

KEYWORDS: adaptive web sites, empirical evaluation, intelligent user interfaces, web-based services.

INTRODUCTION

After a preliminary questionnaire given to fifty people to know the real needs of potential users, we developed E-tool, a web site offering a wide range of tools for working users such as translations, telephone numbers, converters, calculators, etc [2]. Because of this large amount of offered services, we decided to personalize the interaction in order to tailor them on the basis of user behavior. Following the Kobsa's classification [4], the *adaptation production* has been oriented to the generation of contents and structure adapted on the basis of a given user model. For more information about the other tasks of the personalization process see [2]. The layout of the site has been designed as a interface structured in a set of frames [Fig. 1]. The system adapts the content by presenting, in the top frame, personalized recommendations based on the similarity of the features of suggested items and the features of items the user liked in the past. In this frame we added *i)* bookmarks to the most used utilities inferred by the user behavior; *ii)* personalized commercial offers; *iii)* keyword marketing banners. The adaptation of the structure regards the central frame where the links to the most used utilities are annotated [1] and the left frames where the links to the available utilities grouped by categories are gathered. Over each session, the links are sorted [1] in a more refined way on the basis of usage frequency and user interests. In contrast, the less used or unused links will be lightly faded but not disabled. These adaptations are aimed to improve not only the navigation between pages but also the effect of information presented within a page.

To give the user control over the system and to increase the interactivity of the interface, the system allows the user to correct the system assumptions in the section "Your profile". While most evaluations of adaptive systems are comparisons of the system with and without adaptivity, where the last one has often not been designed optimally for the task [3], the two versions of E-tool are quite similar. Our choices have been also led by the fact that we want to leave the interface predictable after the adaptations.

THE TEST

The test is aimed to measure the task completion time, the amount of within page navigation and the satisfaction of the users. Since our system shows the adaptation after a certain number of interactions in a given period of time, we decided to simulate this scenario instructing the users in a pre-test phase and let them just "click around" to make the interaction more natural. Moreover, we structured our test by a repeating task: at the beginning of the test the user has to accomplish a task in the non-adaptive version (NAV) and at the end of the test s/he has to repeat the same task in the adaptive version (AV) to test the effectiveness of the adaptive changes. We evaluated 14 subjects aged 24-35, all with high knowledge of the Internet and web browser and all using the Internet during job time. These subjects are the kind of users which the site was projected for. The subjects were split into two groups (seven subjects each) and everyone had to solve eight tasks by exploiting utilities given by the site (for e.g. "Look for the telephone number of John Red" or "Translate this word.."). Group 1 had to solve four tasks in the NAV and four tasks in the AV. On the contrary, Group 2 had to solve the last four tasks of Group 1 in the NAV and the first four tasks in AV in order to compare the results of the two versions. The solutions of the four tasks in the AV could be reached by exploiting the supposed *facilities* offered by one of the applied adaptation techniques: (*i*) a sorted menu-item (left frame) or a annotated link (central frame), (*ii*) a annotated link (central frame), (*iii*) a recommendation (top frame), (*iv*) a bookmark (top frame). Every group had a repeated task as described above: the first task in NAV was repeated at the end in the AV to simulate the real running of the system. After the test, we asked the subject eight question about their viewpoint of the system.

RESULTS

The results of the test are shown in the Figure 2 (n. 1 means that Group 1 used the AV, n. 2 means that Group 2 used the AV). The best results in the completion time are given by the annotation (a): the users scan the page content

Figure 1. A portion of the home page of the English version of E-tool (http://www.e-tool.it/OfficeEng/).
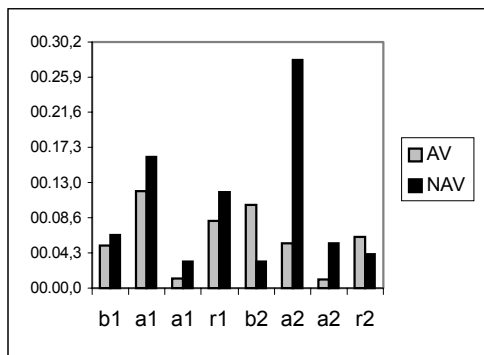

Figure 2. The test results (average seconds per task).

and if some chunk of information is highlighted the attention is focussed on it. This technique is also effective because when users often exploit a utility they tend to remember the general collocation of the link and by emphasizing it improve the retrieval of its position [2]. By observing the users we noticed that the highlighting decreases the within page navigation. The exploitation of bookmarks was related to the repeated task. At the end of the test the users repeated the first task with a difference: the task began from the last viewed page (the other start from Home). This change is due to the fact that the expected utility of bookmarks increases when the user is in another page and instead of going back Home can click on the bookmark in the top frame. Five users of Group 1 (b1) used the bookmarks and the other two went back Home. Instead, two users of Group 2 (b2) exploited the bookmarks, the other four went back Home and only one clicked on the left menu. In fact, Group 1 had better results then 2. Only one user per group used the recommendations (r) to solve the third task and therefore the differences are not relevant. The final interview showed that the most of the users were satisfied with the site and preferred the AV (only two didn't like it). They found adaptation useful to accomplish repeated tasks and to avoid information overload. They mainly preferred the highlight technique and only two voted for the bookmarks which are generally not immediately perceived, as recommendations. All of them but one found the left menu extraneous to the rest of the site (never used in the required AV tasks) and

suggested adding icons and stretch-text containing the subcategory links, and removing the fade. Furthermore, they mainly didn't like the exploitation of the sorting techniques because of the disorientation due to the changes in the presentation order.

DISCUSSIONS
These indications are useful but not exhaustive. We made the changes suggested by the users and we are also studying new possibilities to improve the adaptation perception of the current version of the site. Particularly, we want to highlight the content-based suggestions in order to focus the users attention on new resources and to avoid the negative effects of constraining the users on specific resources. In addition, we want to test the role and the effectiveness of the commercial suggestion and their impact on the users. Therefore, we decided to continue to evaluate the site also by analyzing the on line behavior of its real users to prove its utility in a longer period of time.

REFERENCES
1. Brusylovsky, P. (1996). Methods and Techniques of Adaptive Hypermedia. *User Modeling and User Adapted Interaction*, 6 (2-3): 87-129.

2. Gena, C., Perna, A., Cena, F., Morisano, R. (2001). A personalized hypermedia application for web-based tools, *Poster and Demonstrations Proc. of the 12th Acm Conference on Hypertext and Hypermedia*, ACM Press, Arhus, Denmark.

3. Hook, K. (1997). Evaluating the Utility and Usability of an Adaptive Hypermedia System, *Proc. of 1997 International Conference on Intelligent User Interfaces*, ACM, Orlando, Florida.

4. Kobsa, A., Koenemann, J., Pohl, W. (2001). Personalized Hypermedia Presentation Techniques for Improving Online Customer Relationships. *The Knowledge Engineering Review* 16(2): 111-155.