OZCHI'96 Workshop on the Next Generation of CSCW Systems, Hamilton, New Zealand, November 25, 1996, pp. 40-44.

Qualitative Evaluation of a Collaborative Web Browser

Simon Gianoutsos and John Grundy
Department of Computer Science, University of Waikato
Private Bag 3105, Hamilton, New Zealand

Abstract

Many tools have been developed to support cooperative work, but many of these have not been evaluated. This paper describes various experiments conducted to derive a qualitative evaluation of a WWW browser with CSCW support. These experiments are not restricted to this particular domain but can also be applied to other types of CSCW systems.

Introduction

With the growing number of tools to support cooperative work, evaluation of these tools is required to increase our knowledge of user requirements. Many of the tools created react differently and offer differing levels of usefulness under certain circumstances, and modifications may be necessary to make a tool more generic. But how generic can a tool be before it is too basic to be useful? Is it useful for interactions between two users, more than two users, and larger groups? This paper provides the results of experimentation with W4, a WWW browser with CSCW support (described further in [2]), and attempts to answer some of these questions.

Related Work

Evaluation of CSCW systems has been noted as being especially difficult due to the different backgrounds of group members, the administrative or personality dynamics within a group, and by the difficulty in trying to emulate realistic groups within a lab by Jonathan Grudin [3]. He also points out that groupware evaluation "in the field" is difficult due to group composition, and a range of environmental factors that may play a role in determining user acceptance, such as training, management buy-in and vendor follow-through. He sees this lack of suitable evaluation as a contributory factor as to why CSCW systems fail to deliver the benefits intended.

Magnus Ramage argues that existing CSCW evaluation techniques are mostly inadequate, because people have spent a lot of time developing methods that are designed to be the one best way to evaluate or design computer systems, but are often based in a particularly disciplinary background and only consider a certain part of a particular situation [5]. His research suggests that evaluation methods need to take into account issues of individual, group and organisational effects as well as questions of useability.

Evaluation of the MEAD prototype [6], a multi-user interface generator tool for use in the context of Air Traffic Control, has provided an insight into many of the problems of evaluation. This evaluation points out that various people have different views about what evaluation actually is, and the multitude of techniques that can be used to perform evaluation. The researchers concluded that their informal evaluation procedures were a powerful, cost-effective means evaluation, yet raised the question of whether systems for use in cooperative work environments could indeed be evaluated for validity in isolation from the work.

Our work provides another case study from which lessons may be learned. The following sections overview W4, a WWW browser with CSCW support, and contain details of what evaluation was performed, how it was performed and lessons learned from the experience.

W4 Overview

W4 (World Wide Web for Workgroups) is a collaborative tool developed to allows users to add a variety of annotations to Web pages, which include simple URL links, notes, text chats, a brainstorming tool, and a shared whiteboard. The HTML source of Web pages is not modified in any way – the annotations are stored centrally by a GroupKit [4] W4 conference. All cooperating users join this W4 conference, enabling all annotations to be made persistent and allowing users to join and leave conferences, preserving annotations and their contents. In addition to annotation capabilities, W4 supports various group awareness and work coordination facilities, including telepointers, multiple scroll bars, shared page histories and bookmarks, and the ability to "follow" other users' page visitations.

Figure 1 shows a screen dump from W4 in use for a cooperative, possibly geographically distributed, task. In this example two (or possibly more) users are collaborating to determine an EFTS (Effective Full-Time

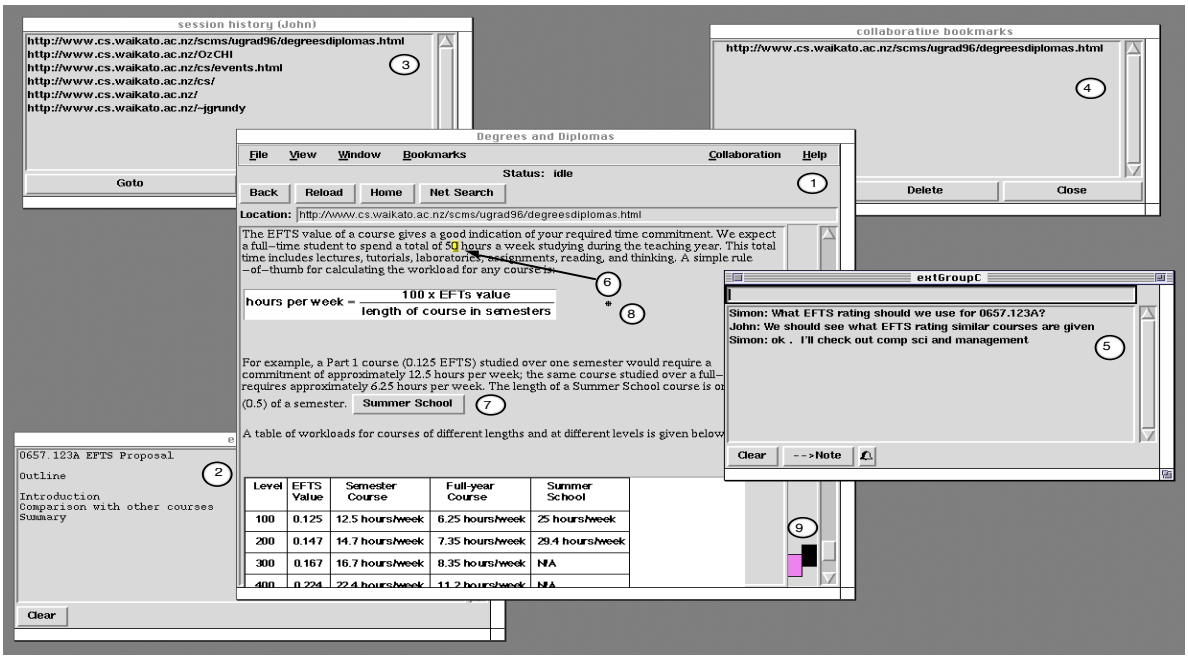

Figure 1: Example of annotated Web page, with collaborative note, text editor, page history and bookmarks

Student) rating proposal for a Part I University course, "0657.123 The Computing Experience". In order to perform this task effectively, the participants need to be able to view the same WWW information as each other, be able to annotate pages of interest with notes or URL links to related pages, be able to collaboratively edit text and diagrams either embedded in the page or separate to it, and be able to send email-like messages to each other or communicate in real time. In addition, they need ways to remain aware of each others work, including pages visited, bookmarked pages of interest, and seeing the focus of attention of their collaborators.

W4 provides a range of facilities to allow collaborators to work together in these ways. As shown in Figure 1, window (1) is the collaborative browser window provided by W4, showing a WWW page from the University of Waikato's Computer Science Department WWW server. The person using this browser is user Simon. The collaborators in this W4 conference also have a text document, in window (2), which they are writing together and which contains the EFTS proposal for the 123 course. This is a collaborative text editor which provides WYSIWIS text editing capabilities. Window (3) shows the session history of Simon's collaborator, John i.e. the WWW pages John has visited while in this conference, and window (4) shows shared bookmarks accessible to all members of the conference. These windows are updated each time John moves to a new page or one conference participant adds a new shared bookmark. Window (5) is a collaborative text chat in which Simon and John have been informally exchanging ideas and discussing the work they are doing. The WWW page has been annotated with a yellow square (a "sticky note" representation), at the position indicated by (6), which when clicked on will display the text associated with this note. Users can also reply to the note, creating further notes, or send context-dependent email-like messages to each other using this notes facility (described further in [1]). A URL link annotation (7) has also been added to the WWW page being viewed, which when clicked on by a collaborator will open the "Summer School" WWW page. Any annotation added to a WWW page is visible by all users, and is shown in other users browsers when they select the appropriate WWW page, or if already selected will appear when they have been added by a collaborator.

Additional group awareness capabilities provided by the browser include telepointers (8), showing the position of collaborators' cursors. Multiple scrollbars (9) indicate the position of other users on the same WWW page. Telepointers and multiple scroll bars are only shown for collaborators who are viewing the same Web page as the user. Users can also click on the scrollbar of another user and request to follow their page browsing.

What we tested

At the commencement of our useability experiments with W4 we were not entirely sure what information we would obtain. Initially, we focussed on trying to obtain a

qualitative measure of usefulness of particular applets under different conditions. This we hoped would guide us in developing tools that user's would find useful. Not only did we want our software to provide useful tools, but also tools that were easy for users to learn how to use, and that facilitated cooperative work.

A variety of projects were offered for user's to undertake including planning trips, discussing updates to Web pages, and collaboratively obtaining information on a certain topic. By giving user's a choice of projects to undertake, user's were not restricted to some abstract topic that they knew nothing about. This also gave user's the feeling that they could govern what direction the project should take, and that they could adventure out on a side-tangent if desired.

To ascertain the usefulness of W4 applets, these tools had to be used for different projects, that lasted for different periods of time, with different numbers of users of varying expertise. Experiments were also conducted with users using the tool at the same time, and at different times.

We wanted the opinions of users with various CSCW and WWW experience, to get a broad perspective of the usefulness of W4 applets. This meant users would have quite different mindsets at looking at a problem, and varying degrees of computer, WWW and CSCW experience.

Tests were conducted with different numbers of users, group members of varying expertise, and different genders. This helped to ensure W4 was not being directed at a certain group of user's and that it was evaluated for single-user browsing and for small group browsing.

How we tested

Qualitative techniques were mainly used to obtain the information we required. Questionnaires before and after W4 tests, observing user's at work, and verbal discussion with user's provided useful qualitative information about W4 applets and W4 as an environment to work. It was also noted during the post-questionnaire that several user's found it easier to communicate their opinions, problems, and suggestions verbally, rather than attempting to put it into written words.

The most useful information was derived from observing users at work, and conversing with them prior, during and after the experiment. A questionnaire was given to all users prior to using W4 to ascertain how familiar they were with the WWW, CSCW systems, and what their expectations were of W4. At the conclusion of the experiment users filled in another questionnaire, and this was focussed on how useful they found the various tools provided for their particular task, and how useful they thought the tools could potentially be (i.e. for other tasks than the one they performed). User's were also asked to comment on their experiences with W4. We did this in order to qualify our judgements of W4 applets and to better understand the responses they had given.

Quantitative techniques built into W4 were used during the early tests of W4. A record was kept of artefact events, when they were made, and by whom. Many different events were recorded, including reading a note, adding text to a whiteboard, viewing a user's session history, and even ringing of the bell. At the conclusion of each experiment these results were analysed. In one experiment that lasted an hour the bell was shown to be used 75 times, which may lead user's to thinking it was being heavily abused. However, further analysis of the event log shows that the bell was often used multiple times in succession (19 times in succession, in one instance) to try and grab a user's attention. Given the extreme "heavy handed" use of the bell by one user, the bell being run 75 times does not give conclusive evidence that user's needed better ways of getting a user's attention. This example, although an extreme case, provides an insight into why we found quantitative analysis ineffective. A more complex quantitative analysis would have possibly provided useful information, but this has been left as future work. Since preliminary quantitative analysis results showed nothing evident, future quantitative tests were abandoned.

Results

After a few experiments it became clear that the number of users using a tool greatly influenced how the tool was used and how effective the tool was. It was also apparent that certain tasks had different requirements for tools that aided communication.

With groups consisting of two users the embedded text chats and whiteboards were not found to be useful. This was due to the fact that users found that communication (for two users at least) was easy enough via an external text chat, a simple tool that has proven very useful for groups of varying sizes. Context-sensitive Notes were very seldomly used within these small groups. If the users were intending to use W4 for much longer periods of time (e.g. a couple of months) then the notes would possibly be used a lot more, largely as a reminder of things that have occurred previously. Longer-term experiments are currently being conducted to validate this. Directed messages were used when users were working at different times and was the main source of communication during this type of asynchronous

interaction, yet the external text chat "took over" as soon as user's were working simultaneously.

Groups of 3-4 users utilised many more of the communication applets, but the predominant applet for synchronous communication was still the simple external textchat. Context-sensitive notes also proved to be a lot more useful, since there was a lot more chance that another user might actually stumble across them.

A collaborative text editor was commonly required by users to compile information retrieved from the WWW, but due to a number of bugs inherent within the text editor provided with W4, it was deemed unusable by users. User's desired a text editor (or even better, a word processor) with group support that they could safely write to, knowing that their text would not accidentally get deleted! User's also found it beneficial that URL links and notes could be embedded within the text editor, since it made the text editor suited to its environment, W4.

Whiteboards were not used commonly by users, although consultation with the users revealed that it was not because they did not need it, but because it did not provide enough functionality. As with the text editor, applets are required that are robust and bundled with features.

A problem that was observed with the experiments was that users with very minimal WWW experience tended to wander off and look at other things on the WWW, and thereby abandoning the project.

Summary

The most basic tools often prove to be the most useful, and from our investigations a simple textchat is useful for groups of varying sizes. Context-sensitive notes and messages are yet another simple idea, yet can also aid user's to work collaboratively. There advantage over conventional email is that they can be associated with work artefacts and are available within the context they are describing.

Collaborative text editors and word processors are important in terms of user's compiling information together. They are also useful if they can have links to the context in which they were created (c.f. URL links in W4).

Our evaluation of W4 shows that if applets do not provide enough functionality or are unusable due to bugs, collaborating workers will not use them. Applets need to be robust and provided lots of appropriate functionality, in order to suit workers requirements.

References

1. Apperley, M.D., Gianoutsos, S., Grundy, J.C., Paynter, G., Reeves, S., and Venable, J.R., "A generic, light-weight collaborative notes and messaging facility for groupware applications" Working Paper, Department of Computer Science, University of Waikato, 1996.

2. Gianoutsos, S. and Grundy, J., Collaborative work with the World Wide Web: Adding CSCW support to a Web Browser. In *Proceedings of Oz-CSCW96,* Brisbane, Australia, August 1996.

3. Grudin, J., Why CSCW applications fail: problems in the design and evaluation of organisational interfaces, in *Proceedings of CSCW'88,* Portland, Sept. 1988, pp 85-93.

4. Roseman, M. and Greenberg, S., Building Real Time Groupware with GroupKit, A Groupware Toolkit. *ACM Transactions on Computer-Human Interaction* (March 1996).

5. Ramage, M., "Evaluation of Cooperative Systems" First Year PhD Report, Computing Department, Lancaster University, 1995. (http://www.comp.lancs.ac.uk/computing/research/cseg/projects/evaluation/1YR_contents.html)

6. Twidale, M., Randall, D., and Bentley, R., Situated evaluation for Cooperative Systems, in *Proceedings of CSCW'94, Chapel Hill*, Oct. 1994, pp 441-452.