Evaluating the Semantic Memory of Web Interactions in the *xMem* Project

Francesca Rizzo°, Florian Daniel°, Maristella Matera°, Sharon Albertario*, Anna Nibioli°

° Politecnico di Milano
Piazza Leonardo da Vinci, 32 – 20133 Milano – Italy
{rizzo,daniel,matera,nibioli}@elet.polimi.it

*Università della Svizzera Italiana
Via Giuseppe Buffi, 13 – 6904 Lugano – Switzerland
albertario.sharon@usilu.net

ABSTRACT

As the amount of information on the World Wide Web continues to grow, efficient hypertext navigation mechanisms are becoming crucial. Among them, effective history mechanisms play an important role. We therefore decided to provide a new method to access users' navigation histories, called *xMem* (Extended Memory Navigation), building on semantic-based and associative accesses, so as to imitate some of the features of the human memory. Such a memory may give users better understanding of the context of their searches, intermixing semantic aspects with the temporal dimension.

The paper presents the experimental study conducted on the *xMem* approach to revisit the Web interaction history. Two controlled experiments have been performed with the aim to evaluate the effectiveness of the *xMem* history mechanism with respect to traditional Web browser histories. The results from the first experiment show a clear advantage, in terms of the time needed to complete a retrieving task, for the subjects that used the *xMem* prototype. Accordingly, users found retrieving previously visited pages with *xMem* more satisfying than using Web interaction histories sorted by the only time dimension. The results from the second experiment show the relevance in the process of information retrieval of clusters and keywords semantically related to the context of the search.

Categories and Subject Descriptors

H.3.3 [**Information storage and retrieval]:** Information storage, information search and retrieval, online information systems. H.5.3 **[Information interfaces and presentation]**: Hypertext / hypermedia.

General Terms

Design, Experimentation, Web History Mechanisms, Human Factors, Hypertext Navigation.

Keywords

World Wide Web, Web Interaction History, Usability, Human Factors, Experimental Evaluation, Information Retrieving.

1. INTRODUCTION

Finding a previously visited page during Web navigation is a very common and important kind of interaction. Many of the results coming from the field of HCI [5][6][4][7] show that more of the 30% of users' activities on the Web are based on the use of the "back" button or of favorites to return to certain pages, but they also show that reverse browsing mechanisms are time consuming and cognitively difficult to use, organize and envision. In fact, even though history tools are common components of browser interfaces there are a variety of human factor issues to be considered [11], mostly dependent from the fact that such mechanisms are typically very simple indexes of visited pages, sorted according to the time dimension.

The literature categorizes retrieving mechanisms as passive or active [3]. Passive history mechanisms, such as the browser history function (see Figure 1) are syntactic in nature, resulting from the navigation actions taken by the user.

Active re-visitation mechanisms, such as bookmarks or the "back" button, have a more semantic quality and are explicitly created by users based on their interest level in the page. All of the passive history mechanisms maintain some kind of information about a particular Web session. The browser stores some subset of the pages visited by the user, typically in a list. However, since most history mechanisms store only the links on the last spoke traversed (i.e., the current path in a depth-first traversal), using this mechanism, the user may only be able to reach a small portion of those pages that have been viewed. If, when deep in a search tree, the user finds an interesting page (i.e., one to which they may wish to return), it may be difficult to preserve that link while continuing with the search.

Bookmarks provide an active method for users to mark interesting Web pages to which they would like to return. For most people, however, bookmarks create a new complication. Because they persist across different sessions, and because users may find a large number of pages interesting, bookmark lists quickly become long and unwieldy. Although most browsers provide mechanisms for editing and otherwise maintaining bookmark lists, they are not an ideal mechanism for maintaining context within a single browsing session.

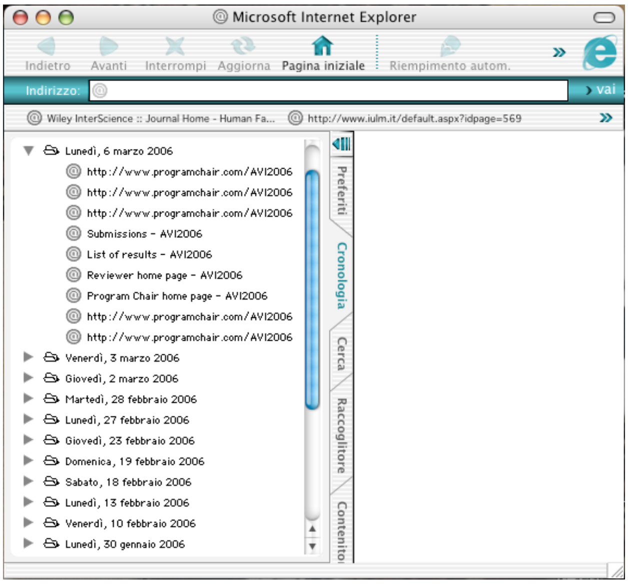


Figure 1. Internet explorer history mechanism. Information is grouped in chunks, hierarchically displayed according to the temporal dimension: at the top level information is sorted per weeks, at the second level per days, at the last level per URL.

The purpose of history mechanisms is to support users to find back to the previously used information. Three facts motivate their introduction on commercial browsers: (1) they help users to navigate through the huge quantity of information provided by the Web, thus providing access to the information they have visited previously; (2) they can substitute search engines for finding old pages and avoid the replication of navigations along intermediate pages to a destination page; (3) they could positively affect users' cognitive activities by reducing cognitive and physical navigation burdens: pages can be retrieved with little effort, and they can show users where they have been in the past.

Even though all commercial browsers incorporate some history mechanisms, people interacting with the Web still experience difficulties to remember where they have been and retrieving already visited page when a certain amount of time is passed from the first visit, since the context in which that page was being viewed is lost. The study discussed in [5] reports a series of results that demonstrate the types of human factor issues that refer to the class of Web history mechanisms. First the 13.4% of subjects engaged in the study affirm not being able to find pages recently visited; only 0.1 of the page accesses occurred through the history list, the 42% of the re-accessed pages were from the back button. So, even though the percentage of the re-accessed pages was high (42%) those re-accessed using the browsers history list were almost 0%. It seems that the path-following method (on which the most part of the history mechanisms are based) for retrieving long term history memories that impose users to traverse in reverse order their previously visited pages is highly inefficient. The method relies on users remembering their navigational behaviour, either because they must recall the page visited and their sequence or because they must realise that they can return to a page by retracing a particular pathway.

We believe that the retrieval of long term information memories from users' navigation history can be improved if the information items are displayed with some other semantic information that can help people to remember the context of their first search and orient themselves [1]. This suggests that better history lists could benefit millions of WWW users.

The *xMem* project offers to users a specialized Web site containing the individual user's memory. The key idea of *xMem* is the transparent transmission to the memory Web site of the URLs of the pages being browsed by the users. Then, suitable rules read the information from the URLs and populate the memory with semantic content. In the light of the previous considerations, the aim of our work becomes twofold:

(i) providing easier and more intuitive navigation-history retrieving mechanisms. Web navigation can be aid by novel navigation history tools providing enriched memory access mechanisms that build on semantics-based search criterions categorizing available pieces of information;

(ii) fostering ubiquitous accessibility of the navigation history by making it accessible over the Internet. The gathered and classified information must be always reachable in order to become an active help facility that can be integrated within the user's browsing environment.

In order to provide users with additional value (with respect to usual history mechanisms), *xMem*, besides chronologically ordered lists of URLs, also supports further semantics in presenting history data. So-called semantic classes or categories are associated to groups of URLs in order to provide high-level meanings for the contents of the related Web sites.

The paper illustrates the controlled study we conducted on two *xMem* prototypes. The adopted method borrowed from the cognitive psychology framework provided us a tool for conducting the users' study to evaluate our approach for the circumstances in which people really don't remember exactly when they met what they are looking for. Note that the paper does not attempt to justify the functional use of the *xMem* tool. The chief contribution of the paper is to emphasize the importance of evaluating the *xMem* mechanism in the context of the related cognitive psychology research. The data presented in the paper focus on the ease of navigation and retrieval for users when they are required to recognize (versus remember) the context of their first search. Our hypothesis is that participants to the experiments should be able to better perform their task if they know some information about the domain of the pages they want to re-access.

The paper is organized as follows: in Section 2 we review research on experimental browser and history mechanisms, including work from HCI and information seeking behavior field. Section 3 describes the *xMem* project in detail. Sections 4 and 5 describe and discuss the experimental methodology, equipment and results obtained from two controlled experiments. Section 6 reports the current *xMem* prototype. The reminder of the paper discusses our conclusion and introduces the future works.

2. RELATED WORK

Few researchers have considered new metaphors for browsing and collecting information on the Web; we next briefly describe their work. Most of the methods introduced in the following use a mix of graphical representations and a sense of context. Some of them place on the end users an extra request of cognitive effort. Most of them are features of experimental browsers and are only able to show syntactic information, not semantic information.

IBM's Web Browser Intelligence (WBI) tool [9] is a browser coupled with personal history functions intended to make Web browsing more efficient by annotating hyperlinks on all Web pages with traffic signals, and this performs well for functions such as: remembering visited pages, providing a keyword search through the text of pages already visited, noticing patterns in the Web browsing behavior and suggesting shortcuts, and automatically checking favorite Web pages for change.

WebTOC [2] is an automated system for creating table of contents (TOC) frames for sets of Web pages. The TOC frame can be quite useful, and having that frame actually run a Java program allows it to more dynamically present the desired information. Its main drawback is that it occupies a large portion of the screen.

WebNet [1] is a browser extension that displays a graphical representation of the hyperspace being explored. It does so dynamically and independent from the content provider. In fact it can do so across many sites. It is a challenge, however, to present the graph in such a way that the contextual information is highlighted.

DeckScape [8] is an experimental browser based on the concept of deck. Each deck is a linear stack of pages that the user can leaf through. As with history mechanisms, if a user starts from page A and goes to B, B is added to that stack or deck. However, unlike history mechanisms, if the user goes back to A and then traverses a link to C, B is not lost; it remains in the deck of pages. Thus, the user can always revisit any page since no page is ever lost. However, users are cognitively loaded with the responsibility of maintaining pages logically in different stacks unless decks are pruned regularly.

Elastic Windows [4] also provides a mechanism that illustrates graphically the hyperspace in which a user is navigating, but it does so more interactively. If the user selects a link using this system, the contents of the corresponding page do not replace the currently displayed page; instead, the new page is displayed alongside its parent. Selecting multiple links from a page results in all the new pages being displayed alongside the parent, but in a smaller size. The same operations may be performed on any window in the browser. Users are provided with functionality to manage the windows by collapsing some sections of the hierarchy, while maximizing the size of others. Since the complete hierarchy is visible at anytime, users can easily move in the hierarchy while not losing their search context. However, the simultaneous display of multiple pages again places a management burden on the user, impacting the cognitive performance.

Recently, [10] the Microsoft research center has worked on the system Stuff I've Seen (SIS), which is able to support users in retrieving and reuse information already seen locally on the PC. The system aims at facilitating the information seeking behavior by providing an index of information that a user has seen (email, Web page, document, appointment) and, in addition a set of rich contextual cues about the searched information, made available from the previous accesses. Personalized Search is an improvement to Google, currently under experimentation. Its aim is to understand how personalization might work for search engine users. In particular, Google offers a search history feature of Personalized Search that enables users to view and manipulate their history of searches. Users might search in their past interactions with Google; they may search history by Web and/or by images; they may pause the history (this means that the services will not collect any history until users choose to resume); they may bookmark the search results displayed in the history list. All solutions described above respond to the need to record users' navigational history (URL lists) for allowing the successive re-access of visited pages. *xMem* works beyond these mechanisms analyzing and interpreting the structure of recorded URLs and making available the results of this process to end users. This is obtained independently from the browser in use.

3. *xMem* PROJECT

Many findings suggest that current browsers lack efficient methods for revisiting Web pages, and that "current interfaces for browsing on the WWW are still primitive... They do not aid users in accessing already visited pages without much cognitive demand" [4].

Users experience frustration in retrieving already visited pages when a certain amount of time is passed from the first visit. The reason is that the context in which page have been viewed is lost. Indeed, this phenomenon may be partially responsible for the lack of efficient mechanisms.

3.1 The *xMem* architecture

xMem consists of several components that share the same data source. The implementation of the correspondent database depends on the expected load at runtime and can consist either in a single database on the *xMem* server itself or in a freely distributed server architecture. The functional architecture of the *xMem* tool is primarily influenced by two goals of our approach: (i) adopting remote logging mechanisms for (ii) providing online access to logged data. Remote logging builds on the client-server paradigm, online log access suggests splitting the overall tool into two logical components, one for each communication direction. Figure 2 graphically depicts the resulting functional architecture, roughly divided into Client PC and *xMem* Server.

The Tracker Client, installed at the client side, is in charge of tracking navigated URLs and of transmitting them to the Tracker Server. The Tracker Client also allows activating/deactivating the tracking mechanism. On the other side, the Tracker Server is responsible for feeding the incoming messages into the URL Repository. For each registered user, such repository contains the actual log data in form of URL strings of the visited Web pages. User data are maintained in a User Repository, which stores psychographic information and user preferences. These data are the basis for managing access rights over history data.

By means of the so-called *xMem* Navigator Web application, users can then access their personalized navigation history over the Internet and browse logged data by means of keywords that are representative for the visited contents. Such keywords are stored in the Semantics Repository. Their extraction is managed by the Page Indexer module, described in the following.

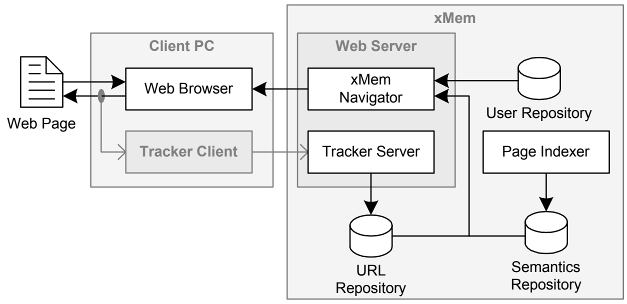


Figure 2. The *xMem* architecture

3.2 The *xMem* model of interaction

In order to provide an alternative paradigm for history navigation trying to solve the previous issues, the *xMem* project offers to users a specialized Web site hosting a repository of the individual user's memory. Besides chronologically ordered lists of URLs (as offered by traditional histories), *xMem* provides further hints in presenting history data. Keywords are associated to groups of navigated pages in order to recall concepts describing the page contents. Keywords may be provided by Web application designers as meta-description of pages, or they can be extracted automatically from re-materialized pages by detecting the main concepts being displayed by the pages themselves. In order to take advantage of the navigation history facilities of *xMem*, users must register as *xMem* users and install an application on their client PCs. While navigating the Web by means of a common Web browser, users can activate this application for tracking the navigated URLs. Such application also manages the transparent transfer of the tracked URLs to the *xMem* system. The tracking mechanism can be enabled/disabled at will. Thus, *xMem* maintains remote log data (with respect to users and 3rd party Web servers) about users' navigation actions. At server side, *xMem* identifies keywords representing concepts seen during navigation, which are used for populating the history memory. *xMem* then provides a Web interface toward the history memory. Besides the chronological lists of URLs, it also offers a semantic organization of history data by exploiting page keywords, for recalling page contents at a comprehensible level of abstraction.

The *xMem* model of interaction relies on the idea that the semantic cuing (i.e., exploiting keywords and clustering) the history list would better support end users in all the situations in which they do not exactly remember when they visited a page. The enrichment of history data is achieved through the extraction of keywords, describing pages, and through the clustering of pages into categories facilitating the retrieval of navigated pages. We define a keyword k with respect to a page p as a natural language term derived from page p, which serves the purpose of indexing a specific content of page p. A page p can thus be characterized by a set of keywords. We then define a cluster C as a collection of pages characterized by similar keywords.

Keyword extraction and similarity definition can be performed according to the semantics of page contents (e.g., by means of ontologies), or using syntactic techniques. In the current version of the *xMem* prototype we use a set of syntactic criteria for both keyword extraction and clustering. Keywords and clusters represent semantic classes, stored in the *Semantics Repository*.

They can be used for achieving a hierarchical organization of visited URLs, where the first level is based on clusters referring to pages with similar keywords, while the second level builds on keywords, used for grouping pages with the same keywords. The leafs of the hierarchy represent the tracked URLs, which allow users to recall the actual page searched. The clustering of history data highlights correlations between visited pages, perhaps not emerged during the original navigation. Such correlations cannot be captured adequately by traditional, time-based histories.

4. *xMem* EVALUATION THROUGH A CONTROLLED EXPERIMENTATION

In order to investigate our assumptions about the semantic enrichment of history, we have conducted two experiments with real users. We wanted to compare the performance of users dealing with a traditional history based on chronological order, with the performance of users adopting an enriched history. In particular, we wanted to verify whether the semantic enrichment supplying additional cues about the page contents significantly improves the user experience. Also, we wanted to verify the effectiveness of the hierarchical organization of the enriched history.

4.1 Experiment 1

We therefore designed a first experiment, comparing three different history organizations (see table 1 for the experimental conditions):

- The Traditional History (TH), based on a chronological sorting of visited URLs (see Figure 3(a)).
- A Hierarchical *xMem* History (HX), based on a hierarchical organization of the enriched history using clusters and keywords for classifying URLs (see Figure 3(b)).
- A Flat *xMem* History (FX), where clusters provide a one-level classification of pages, while keywords just extend the descriptions of URLs (see Figure 3(c)).

Group	Experimental Condition
TH Group	Condition 1: subjects use traditional history.
HX Group	Condition 2: subjects use the hierarchical *xMem* history.
FX Group	Condition 3: Subjects use the Flat *xMem* history

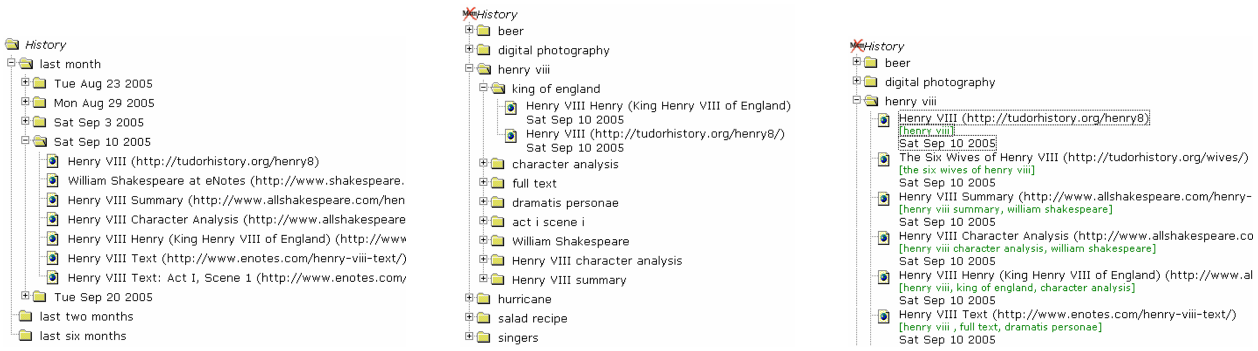
Table 1. The experimental conditions.

4.1.1 The Hypothesis

We started by considering that the organisation used to display contents in browser histories might impose different demands on working memory. This depending on the nature of the cues people hold in mind (about the precise day, week and month he/she visited the page) and on the level of participants' expertise on the knowledge domain which the searched page belongs to.

Our hypotheses were the following:

1. When a user does not exactly remember the time of access to a previously visited page, finding that page, based on a



a) Traditional history b) Hierarchical history c) Flat history

Figure 3. The different history organizations tested.

pure temporal sorting, such as the one used by traditional histories, requires a higher cognitive effort. We therefore expect a better performance by *xMem* users.

2. The hierarchical organization of history data, as also demonstrated in other domains, improves the retrieval task. We therefore expect a better performance of the HX history with respect to the TH and FX history.
3. We finally hypothesize that the enriched (both hierarchical and flat) history enhances user satisfaction. We however also expect a higher satisfaction for HX users with respect to FX users.

4.1.2 The Method

The experiment has been designed in order to identify significant differences in the time taken to retrieve a page, using the three different histories among three groups of participants.

- The TH group was asked to use the chronological history;
- The HX group was asked to use the enriched hierarchical history;
- The FX group was asked to use the enriched flat history.

Subjects' performance was measured on the task completion time. At the end of the experiment, participants were required to fill in a questionnaire combining three different dimensions:

- the previous experience of participants with history mechanisms;
- the knowledge of participants about page contents;
- the subjects' satisfaction with the tool and the keywords.

The Procedure

The experiment was conducted using a classical paradigm based on the manipulation of the independent variable (the history sorting criterion) and on the measurement of the values of the dependent variable (task completion time).

Each subject was assigned to one of the three experimental conditions and asked to retrieve a page about "Henry VIII" using one of the three mock-up interfaces.

Subjects

Participants were selected among the undergraduate students of the USI of Lugano - Swizerland, Università della Svizzera Italiana. 45 subjects were recruited and allotted to the three different groups. The three experimental groups did not differ in relation to the expertise on the task domain they had to perform and their expertise on the use of Web browsers.

Tasks

The experimental task was a simulation of the retrieval of content already visited on the Web. Their task was assigned in form of a written scenario, outlining the presumed past navigation actions. Based on this scenario, subjects were asked to retrieve, by means of the history mechanisms they were assigned to, a page showing contents about "Henry VIII". The scenario purposely did not provide complete indications about the time of visit of the page, so as to simulate the lack of memory along the temporal dimension.

Materials

Three mock-ups were purposely implemented for the study, one for each type of history sorting criterion under test (see Figures 3(a) – 3(c)). Each subject was exposed to the same list of 40 URLs. For the TH group, URLs were structured as they actually are displayed in traditional browser histories. Each of them belongs to one of 6 different knowledge domains: Hurricane Katrina, Beer, Henry VIII, Digital Cameras, Salad Recipes, Singers. The URLs were listed in a random order.

4.1.3 Results

In order to check whether our experimental hypotheses were verified, we performed a cross-comparison of the collected data for the three groups. An ANOVA test performed on the completion times showed a significant difference for the means of the three groups ($F(44)= 3.248$, $p < .049$).

Furthermore:

- A t-test showed a significant difference in the retrieval time between the TH and the HX group in favour of the *xMem* history organization ($t(28) = -3.073$, $p < .005$).
- A t-test comparing the performance of group HX and of group FX did not show any significant difference ($t(28) = -1.521$, $p < .139$).
- A further t-test comparing the performance of TH users and FX users did not produce a significant difference ($t(28) = -.927$, $p < .362$).

Table 2 reports the means of the measured retrieval times for the three groups.

The values show that the average page retrieval time for the TH group is almost twice the value of HX group, while the FX group only slightly improves the retrieval time.

Group	Mean retrieval Time
TH Group	116 sec.
HX Group	64 sec.
FX Group	93 sec.

Table 2. Mean task execution time per group.

Group	Mean satisfaction
TH Group	2,67
HX Group	3,67
FX Group	3,60

Table 3. Means satisfaction per group.

The general satisfaction was expressed by users on a scale of 5 points (1 for very negative, 5 standing for very positive). Table 3 reports the mean values for satisfaction for the three groups. The data show a significant difference between both the TH group and the HX group ($t(28) = 2.617$, $p < .016$) and the TH group and the FX group ($t_{(28)} = 2.514$, $p < .020$).

We also observed that the better performance of the HX group did not depend on the subjects' expertise on the knowledge domain (Henry VIII) or on their expertise on history mechanisms. Indeed, an analysis of correlation between the task completion times and the subjects' expertise on the contents of the visited pages did not show any significant value ($r = -078$; $p < .782$).

Also, a further analysis did not evidence any significant correlation between the performance of the HX group and the expertise on history mechanisms ($r = -.078$; $p < .783$). This can be interpreted as proof of the validity of the enriched history as an efficient tool for history navigation, independently from the user's background.

4.2 Experiment 2

Looking at the results obtained with the first experiment, we have planned a second test to investigate the relevance of the categories used to cluster the history pages in driving the user to reach his/her goals. For this purpose, in particular, we have investigated the difference on recalling the categories and the keywords among the three different experimental groups.

4.2.1 The Hypothesis

We started by considering that the more the user exploits classes and keywords for retrieving the searched URL, the more he/she will remember them in a post-experiment recalling task.

Our hypothesis was the following: when the user is exposed to the hierarchy of classes and keywords, the measure of their recalling in a post-experiment task is directly related to the level of attention they receive during the information elaboration process.

Therefore we expected the best performance on the recalling task for the group that exploits more the classes and keywords, available on the tool, during the retrieval task.

The general satisfaction of the user with the tool is not taken into account in our system of hypothesis for this experiment, because the proposed task could never succeed, laying the user in a sense of frustration that might affect his/her judgments.

4.2.2 The Method

The experiment has been designed in order to identify significant differences in the recalling values of the classes and keywords, using the three different histories (see Figure 3(a) – 3(c))[1]. The experimental design required participants to use one of the history mechanisms to re-access a previous visited page. After the task execution, subjects were asked to recall all of the clusters and keywords they remember to have noticed in the history used. At the end of the experiment, participants were required to fill in a questionnaire combining two different dimensions:

(i) the previous experience of the user with history mechanisms;

(ii) the expertise of the participants on the domains proposed.

The Procedure

The procedure was the same of the Experiment 1. At the end of the experiment subjects were required to list the clusters and keywords they remembered.

Subjects

Participants were selected among the undergraduate students of various Universities (Università della Svizzera Italiana, Università degli studi di Pavia, Università Cattolica Sacro Cuore). 45 subjects were recruited and allotted into three independent groups. The participants were different from the ones recruited in the previous experiment. The three experimental groups did not differ in the expertise with both the experimental task domain and the use of browsers and history mechanisms.

Tasks

Subjects had to retrieve a URL. The task was assigned in form of a written scenario, also outlining a past navigation context in which the URL to be retrieved was supposed to be visited. The scenario purposely did not provide complete indication about the time the first page visit took place, so as to simulate the lack of memory along the temporal dimension. Furthermore, the URL the users had to retrieve was not included in the list proposed, in order to force participants to focus their attention on all the information items available on the prototype they were exposed to.

Materials

We used the same mock-ups implemented for the first experiment (see Figures 3(a) – 3(c)).

4.2.3 Results

In order to check whether our experimental hypotheses were verified, we performed a cross-comparison of the collected data for the three groups. An ANOVA test performed on the recalling variables (clusters, keywords) showed a significant difference among the three experimental groups for clusters ($F_{(44)} = 7.844$, $p < 0.001$) and for keywords' values ($F_{(44)} = 7.535$, $p < 0.002$).

[1] For the traditional history we considered as keywords the date of the visited URLs and as clusters the monthly classification of the dates.

Furthermore:

- A t-test calculated a significant difference on recalling task for clusters ($t(28)= -3.922$, $p< 0.001$) and for keywords ($t(28)= -3.108$, $p< 0.004$) between TH and HX groups, showing a better performance of HX group on both measures;
- A t-test conducted on data of the recalling task of clusters ($t(28)= 2.107$, $p< 0.044$) and keywords ($t(28)=2.694$, $p< 0.012$) between HX and FX showed a superiority in the performance of the HX group;
- A further t-test comparing the TH and the FX groups did not show a significant difference.

Table 4 reports the mean number of recalled keywords and clusters for the three groups:

Group	Clusters' recalling mean	Keywords recalling mean
TH	0.13	0.20
HX	2.00	2.47
FX	0.27	1.00

Table 4: Mean recall values for keywords and clusters

The table shows that the recalling values of the HX group are almost twice those of the other two groups.

Finally, any significant difference has been noticed between the expertise on the task domain and the recalling performance and the expertise on history use and the recalling performance.

5. GENERAL DISCUSSION

From the previous experimental results, it seems possible to claim that the hierarchical order produces a significant effect on the retrieval of pages when users do not exactly remember the time in which the page has been visited.

The experimental hypothesis that a simple semantic enrichment, such as the one provided by the flat history prototype, would improve times for page retrieval has not been totally confirmed. Indeed, the flat history organization, contrarily to the hierarchical *xMem* history, does not introduce any significant improvement with respect to the traditional paradigm. In addition, the mean for the FX group (see Table 2) suggests a deterioration of the users' performance with respect to the HX group.

In more detail, the results of the t-tests reported above are consistent with the experimental hypothesis 1, since they indicate that in general subjects' performance improves when using enriched histories. Results also suggest a clear and significant advantage of the HX group over the TH group.

A possible reason for this result is that, when the page access time is not exactly known, the user needs to scan the whole URL list, while holding in the working memory one of the already scanned items, temporarily considered the most pertaining.

From a cognitive perspective, this implies that a high amount of cognitive resources are spent on scanning, matching and judging the remaining results in the list. In addition, the demand for cognitive resources needed to carry on the comparison and evaluation activities increases as the information elaboration process goes on. This results in a competition between the cognitive resources needed to maintain information and those needed to elaborate it. The reason of the better performance of the HX group lies therefore in the hierarchical classification carried by clusters and keywords, which enables more efficient information processing from the cognitive point of view. In particular, the HX users first analyze and judge the relevance of the clusters for their search. After a cluster is chosen, the subjects can shift their cognitive resources to analyze and judge the keywords of the cluster. Finally, when a keyword is chosen, subjects shift their attention to the small list of URLs associated with that keyword. This process seems to support a more economical management of attention on working memory.

In general, this explanation suggests that the information retrieval problem can be considered essentially a reiterative process of evaluation and remembering item based on the maintenance of the information judged as relevant for the task and the analysis of the new information coming from the environment. When the information to be held in mind increases, it requires more cognitive resources; as a consequence the amount of attention available for the process of analysis decreases. The competition for the cognitive resources between the two processes may negatively impair the retrieving task if the design choice at the interface side does not provide an effective mechanism to classify information.

As a counter example it may be considered data about the recalling performance of keywords, derived from the post-experiment questionnaire in experiment 1, they indicate that none of the subjects in the FX group remembers a keyword. This means that the keywords used to describe the URLs in the FX prototype did not receive attention by subjects. On the contrary, the perception of some additional items to be read (the keywords) interferes with the process of analysis and judgement of the information item impairing the task time completion. As a consequence, it may be said that the performance on the experimental task for the FX group seems not to be driven by the additional semantic characterization of keywords, but rather by the access structure through which this information is made accessible.

Furthermore data coming from the experiment 2 suggested that neither the hierarchical order, like the one used in traditional history tools and based on the time dimension, nor the semantic cuing of the users' history are per se effective, since people do not spend much attention on their clusters and keywords (as showed from recalling data for TH and FX groups). Rather, the combination of the hierarchical order together with a set of clusters and keywords semantically related to the context of the first search (like the one used by the HX group in both experiment 1 and 2) yields to a more effective solution in supporting people during the retrieval task when they don't remember the date of their first search. In relation to the data about the difference between the performance of the HX group and the TH group, it can be argued that the design choice, on the interface side of *xMem*, to display the semantic information in a hierarchical order significantly affects the users' performance. As a consequence, the superiority of the *xMem* tool to support retrieving tasks from a history list not only depends on the semantic information people can exploit during the retrieving activity, but it is strongly affected by the way this information is displayed to the users.

6. THE CURRENT *xMem* PROTOTYPE

Supported by the previous results, we have decided to adopt the hierarchical organization for visualizing the enriched history.

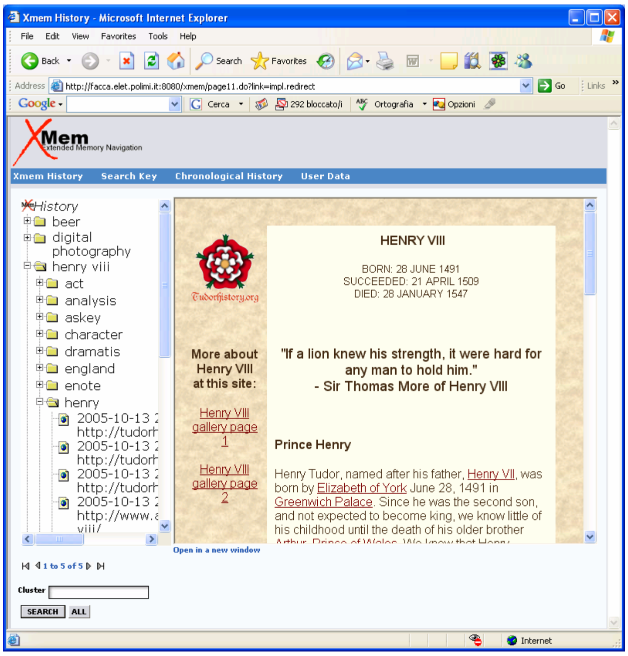


Figure 4. Screenshot of the current interface of the *xMem* navigation tool. At the left hand side users can navigate their personalized history, while at the right hand side they have the possibility to show a preview.

Figure 4 shows the current *xMem* Navigator's Web interface that allows browsing the user's navigation history, according to the enriched history organization discussed in Section 3. This front-end allows registered *xMem* users to access history data by chronological order, as well as by means of a cluster and keywords hierarchy. Also, a keyword-based search is possible.

7. CONCLUSION

We have argued that current browsing functionalities do not adequately support retrieving information from a user's navigation history. *xMem* improves history mechanisms by making use of new criteria to organize and show the navigational history instead of simply exploiting time-sorted history mechanisms that prevail today. As also demonstrated by two controlled experiments, this retrieving strategy makes history navigation easier and more effective, because it provides a characterization of the context in which information has been seen. Context-dependent history mechanisms sustain users' episodic memory of visited Web pages.

As future works, we are planning to investigate the use of ontologies for augmenting the level of abstraction of our keywords and clusters and for highlighting also possible relationships existing among keywords. We will extensively experiment the possibility of interfacing search engines for the provisioning of keywords. We are also planning a further improvement of the user experience by means of collaborative filtering techniques for history sharing among *xMem* users and by means of a cooperative interaction paradigm, in which users could also refine the automatically derived indexing and clustering.

8. REFERENCES

[1] A. Cockburn and S. Jones. Which way now? Analysing and Easing Inadequacies in WWW Navigation. *Int. J. of Human-Computer Studies*, 45(1): 105–129, July 1996.

[2] D. Nation, C. Plaisant, G. Marchionini, and A. Komlodi. Visualizing Websites Using a Hhierarchical Table of Contents Browser: WebTOC. Proc. of Human Factors and the Web'97, Denver, June 1997.

[3] D. Newfield, B. Sethi, and K. Rayll. Scratchpad: Mechanisms for Better Navigation in Directed Web Searching. In Proc. of UIST'98. ACM, 1998.

[4] E. Kandogan and B. Shneiderman. Elastic window: A Hierarchical Multi-window World Wide Web Browser. In Proceedings of UIST'97. ACM, 1997.

[5] GVU's WWW Surveying Team. Gvu's 10th WWW User Survey, 1998.

[6] L. Catledge and J. Pitkow. Characterizing Browsing Strategies in the World-wide Web. Proc. of WWW'95: Technology, Tools and Applications, April 1995.

[7] L. Tauscher and S. Greenberg. How People Revisit Web Pages: Empirical Findings and Implications for the Design of History Systems. *Int. J. Human Computer Studies*, 1997.

[8] M. Brown and R. Shillner. Deckscape: An Experimental Web Browser. Proc. of WWW'95: Technology, Tools and Applications, April 1995.

[9] R. Barrett, P. Maglio, and D. Kellem. Web Browser Intelligence: Opening up the Web. In Proc. of COMPCON'97, 1997.

[10] S. Dumais, E. Cutrell, J. Cadiz, G. Jancke, R. Sarin, and D. Robbins. Stuff I've seen: A system for personal information retrieval and re-use. In Proc. of SIGIR03, July 28 August 1, 2003, Toronto, Canada, 2003.

[11] W. A Kellogg, J. T. Ritchards J. T. "The Human factors of Information on the Internet, in *Advanced in Human Computer Interaction*, Ed. J. Nielsen, Ablex Press, (5), 1-36.