Supporting World-Wide Web Navigation Through History Mechanisms

Linda Tauscher
Computer Science Department, University of Calgary
tauscher@cpsc.ucalgary.ca

Cite as:

Tauscher, L. (1996) **Supporting World Wide Web Navigation Through History Mechanisms.** A position paper in the *CHI 96 Workshop: HCI and the Web*. Vancouver, BC. April.

Abstract

My research concerns navigational history mechanisms within graphical WWW browsers. I believe that improved history support within browsers can reduce the cognitive and physical burdens of navigating to and recalling previously visited Web pages. A major component of my research involved an empirical study in which I collected data from 23 subjects over a period of 6 weeks as they used an instrumented version of Mosaic v2.6. My overall objective was to identify patterns in revisits to pages; similar patterns were identified in previous research into UNIX command reuse, and other hypertext systems.

Introduction

My research concerns navigational history mechanisms within graphical WWW browsers. I believe that improved history support within browsers can reduce the cognitive and physical burdens of navigating to and recalling previously visited Web pages. Cognitive burdens arise when trying to navigate back to a page the user has recently seen, or trying to recall the address for a page that was visited some time ago. Physical burdens arise when the user must repeatedly navigate back through a set of pages, or visit pages that are not of interest but lead to a page of interest. Effective history mechanisms can also make browsing more efficient because fewer unnecessary pages will be accessed and downloaded. This indirectly benefits all Internet users.

Web browsers currently provide four history mechanisms: backtracking, already-visited cues, history lists, and bookmarks or hotlists. While useful, these mechanisms have several deficiencies.

Backtracking is accomplished either by one or more activations of the Back and Forward buttons, or by the selection of the desired URL from the history list. My own study and those of others (Catledge & Pitkow, 1994) show that Back is a heavily used navigation command. However, problems can arise when the user backtracks more than once, or has visited a particular node more than once. In either of these cases, backtracking via the history list would be more effective. Yet history lists are rarely used. Catledge and Pitkow (1994) found that only .1% of all URLs were selected using Mosaic's Window History feature while my research indicates that .7% of URLs were accessed this way. This low usage may be explained by users incorrect mental models of how browser history lists operate. Cockburn and Jones (1996) performed a usability study of hypertext navigation facilities in three popular WWW clients: Mosaic, Netscape and TkWWW. They found that the sessional history list is based upon a stack, whereas users often expected it to be temporal

or linearly incremental. Thus, users found navigation behaviour using Back, Forward, history list selections, and cyclic Web page links unpredictable or "non-deterministic."

Another problem is that history lists on the WWW are not maintained between invocations of a web browser. Still, a limited amount of inter-sessional history is provided by showing that a hyperlink has been previously visited, usually by changing the colour of the link on the page that contains it. This is the major "already-visited" cue in graphical WWW browsers, and is based upon a user preference (visited expiry date) and a global history list. Unfortunately, the colour of hyperlinks do not fade to indicate how long ago a link was followed.

All graphical WWW browsers contain some method for saving interesting URLs in a list. Mosaic refers to this list as the "Hotlist" while Netscape Navigator calls such items "Bookmarks." This feature eases the burden of returning to sites in which one anticipates a future interest. Two problems with the Hotlist are that the user must explicitly add the URL to the list while viewing it on the display (or enter it into a dialog box later), and the user must manage this ever-growing list of sites.

While WWW browsers do include some history features, they tend to be ad-hoc designs. They do not appear to take advantage of previous research into reuse mechanisms within user interfaces. Their designs are not based upon actual studies of how people revisit Web pages. Their actual use has not been examined except superficially.

Data Analysis

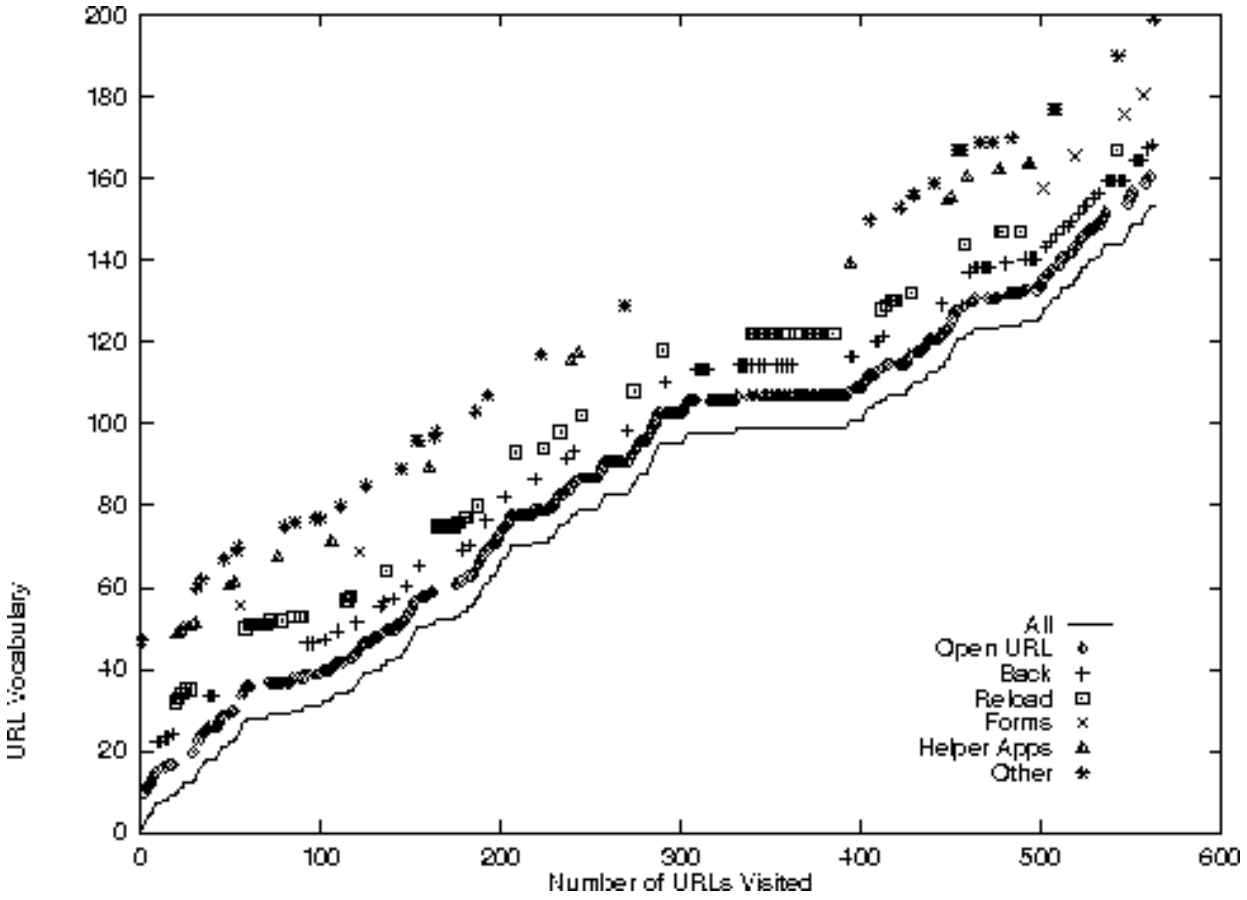
A major component of my research involved an empirical study in which I collected data from 23 subjects over a period of 6 weeks as they used an instrumented version of Mosaic v2.6. My overall objective was to identify patterns in revisits to pages; similar patterns were identified in previous research into UNIX command reuse, and other hypertext systems. Some of my preliminary results include:

1. Recurrence Rate

History systems are only useful if users actually repeat their activities. Greenberg (1993) coined the term "recurrence rate" to refer to the probability that any activity is a repeat of a previous one. My results indicate that 58% of Web pages are revisited. According to this and other criteria, this qualifies Web browsing as a recurrent system for such systems exhibit a recurrence rate between 40-85% (Greenberg 1993).

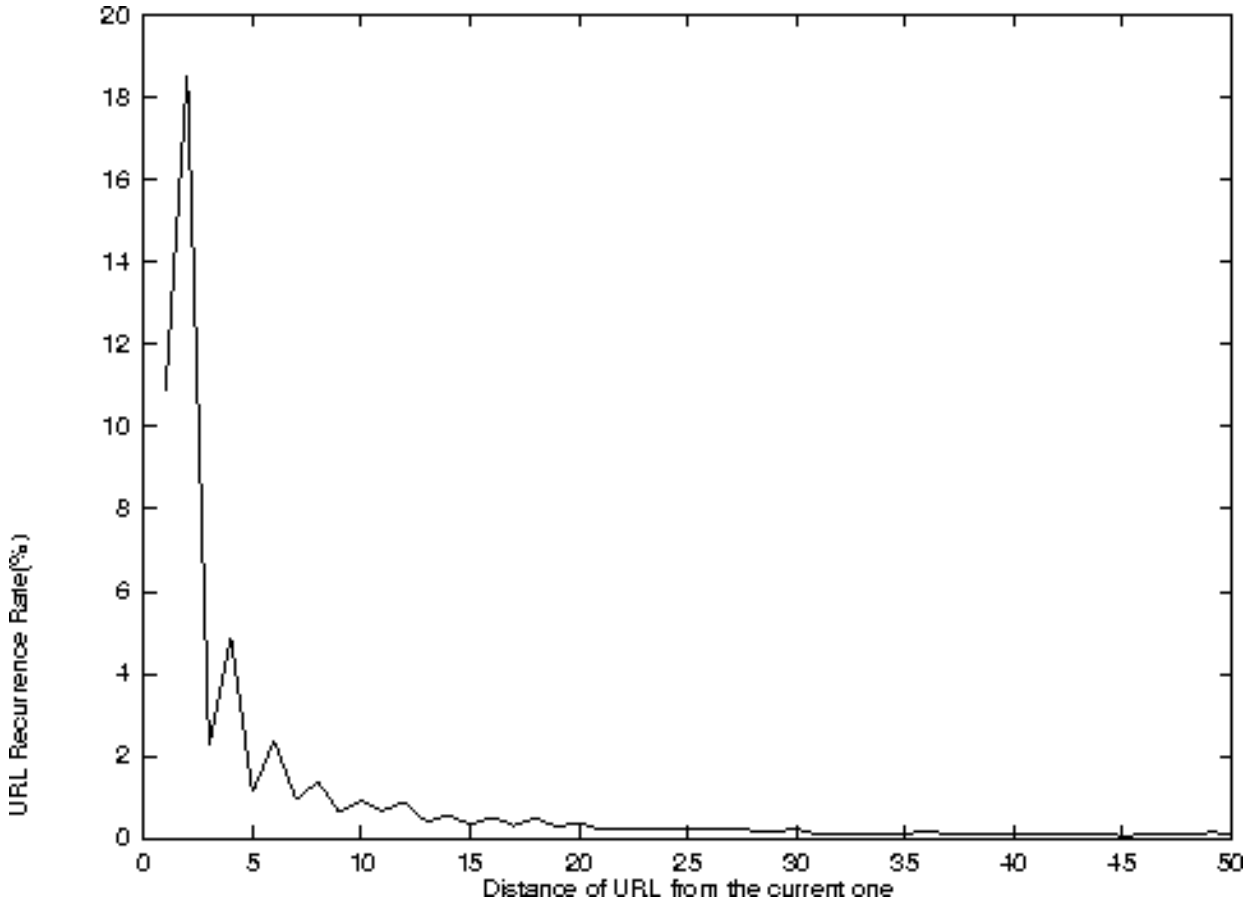
Analysis of subjects' page vocabulary (the rate at which they incorporate new pages into their repertoire of visited pages) shows a remarkably linear slope across all subjects, with short plateau areas where authoring or intense revisits to a site occurred for many subjects (see Figure 1). Thus, while Web pages are recalled, new pages are incorporated at a regular rate. These pages must be accounted for by a history system.

Figure 1



Another interesting question about page revisits is when they occur. My analysis of the distance of the current URL from its last access shows that there is a 40% probability that the current page was visited within the last six URL accesses (see Figure 2). This indicates a strong recency effect in page revisits.

Figure 2



2. Frequency

My analysis of the frequency of page visits indicates that there are few pages that are visited frequently. A frequency plot for all 23 subjects shows that 59% of pages were only visited once, while 20% of pages were visited twice during the 6 week study period. Frequently visited pages tend to be the start-up document for the user (usually their home page), a search engine, or index pages of some sort. Many of the pages that appeared on a frequently visited pages report (that reported the top 15 pages) were only accessed during one particular session. These URLs would be prime candidates for pruning from a history list.

3. Locality

Lee (1992) applied the locality concept from program memory reference research to user interactions in Unix. She found that users repeatedly reference a small group of both Unix command and command lines. I applied this concept to Web browsing activity to determine whether users browse within a cluster of related pages. While locality sets were found in the data, this pattern of presenting history is not recommended for several reasons. First, most locality sets were very small consisting of only one or two unique URLs. Second, these sets lasted for only a short time (usually 2.5 to 4.5 pages). Third, most of the locality sets were never repeated, especially for larger sets.

4. Longest Repeated Subsequences

The concept of paths has been associated with hypertext ever since Vannevar Bush envisioned the

technology in the 1940's. Do WWW users repeatedly follow the same path when browsing? I applied the Pattern Detection Module algorithm (Crow and Smith, 1992) to the data to identify longest repeated subsequences (LRSs) of page visitations. While LRSs do exist, they tend to be short (two or three URLs), and were invoked with an average frequency of two (one of the minimum criteria for being considered a LRS). These results could be improved my making the PDM algorithm more domain-specific. For example, the algorithm could allow the occasional navigation to a side trail if the user then continued along a well-traveled path; hypertext, by nature, encourages such explorations but the current algorithm would consider such deviations as separate LRSs.

5. Use of Browser History Features

My study identified the mechanism used to access a URL. The Back button was responsible for 30% of URL accesses while Forward and Home generated less than 1% of the requests each. Other notable statistics include: less than 1% of accesses were made by choosing the URL from the Window History dialog, and less than 3% of page visits were initiated by a bookmark selection.

Future Work

The next step in my research involves assessing the predictive goodness of various methods of presenting browsing history; I will use my subjects' Web data to accomplish this. Then I plan to build a prototype to demonstrate the viability of the most predictive history method and evaluate it via a small usability study.

The Researcher

I am a Master's student in the Department of Computer Science at the University of Calgary, and am specializing in Human-Computer Interaction. My work now considers HCI issues within the Internet. This interest began with a graduate HCI course project in which I evaluated Internet resource discovery mechanisms. I took this research one step further in a Human Factors graduate course when I designed and conducted a study to elicit experts' search strategies within the WWW. As a 1995 intern student for the Canadian National Research Council, I conducted a feasibility study about how the Canadian Codes Centre committees could take advantage of Internet technology to perform their collaborative work. As part of this work I implemented a prototype system consisting of a Web server, mailing list software, and a mailing list to Web gateway.

References

Catledge, Lara and James Pitkow (1995). "Characterizing Browsing Strategies in the World-Wide Web." *Proceedings of the Third Conference on the World-Wide Web*.

Cockburn, Andy and Steve Jones (1996). "Trails, trials, and tribulations: unravelling navigational problems in the World-Wide Web." *Hypertext '96 Proceedings*.

Crow, Daniel and Barbara Smith (1992). "DB_Habits: Comparing Minimal Knowledge and Knowledge-Based Approaches to Pattern Recognition in the Domain of User-Computer Interactions." in *Neural Networks and Pattern Recognition in Human-Computer Interaction*, edited by Russell Beale and Janet Finlay, Ellis Horwood, New York.

Greenberg, Saul (1993). *The Computer User as Toolsmith: The Use, Reuse, and Organization of Computer-based Tools*. Cambridge University Press.

Lee, Alison (1992). "Investigations into History Tools for User Support." Ph.D. Thesis, University of Toronto.