Pushing Back: Evaluating a New Behaviour for the Back and Forward Buttons in Web Browsers

ANDY COCKBURN, BRUCE MCKENZIE and MICHAEL JASONSMITH

Human-Computer Interaction Lab, Department of Computer Science, University of Canterbury, Christchurch, New Zealand
email: {andy, bruce, mpj17}@cosc.canterbury.ac.nz

The Back button on web browsers is one of the world's most heavily used user interface components, yet its behaviour is commonly misunderstood. This paper describes the evaluation of a 'temporal' alternative to the normal 'stack-based' behaviour of Back and Forward. The main difference of the temporal scheme is that it maintains a complete list of previously visited pages. The evaluation compares the efficiency of the stack and temporal schemes in an 'out of the box' scenario in which participants were asked to use a 'new' version of a commercial browser without any explanation of the presence or absence of new features. This scenario allows us to predict the likely usability impact if commercial browsers were released supporting the temporal scheme. The results showed that the relative efficiency of the two schemes differed across different types of navigational task. In particular, the temporal system poorly supported backtracking to parent pages, but performed better for more distant navigation tasks. The temporal scheme also caused extreme usage patterns, with the subjects either solving tasks very efficiently or very inefficiently, depending on whether they used the Back menu. This observation indicates that adaptations of the temporal system that improve the effectiveness of the Back menu may enhance web navigation.

KEYWORDS: Web navigation, World Wide Web, revisitation, browser design, hypertext.

Pushing Back: Evaluating a New Behaviour for the Back and Forward Buttons in Web Browsers

Andy Cockburn, Bruce McKenzie and Michael JasonSmith

Human-Computer Interaction Lab, Department of Computer Science, University of Canterbury, Christchurch, New Zealand
email: {andy, bruce, mpj17}@cosc.canterbury.ac.nz

The Back button on web browsers is one of the world's most heavily used user interface components, yet its behaviour is commonly misunderstood. This paper describes the evaluation of a 'temporal' alternative to the normal 'stack-based' behaviour of Back and Forward. The main difference of the temporal scheme is that it maintains a complete list of previously visited pages. The evaluation compares the efficiency of the stack and temporal schemes in an 'out of the box' scenario in which participants were asked to use a 'new' version of a commercial browser without any explanation of the presence or absence of new features. This scenario allows us to predict the likely usability impact if commercial browsers were released supporting the temporal scheme. The results showed that the relative efficiency of the two schemes differed across different types of navigational task. In particular, the temporal system poorly supported backtracking to parent pages, but performed better for more distant navigation tasks. The temporal scheme also caused extreme usage patterns, with the subjects either solving tasks very efficiently or very inefficiently, depending on whether they used the Back menu. This observation indicates that adaptations of the temporal system that improve the effectiveness of the Back menu may enhance web navigation.

KEYWORDS: Web navigation, World Wide Web, revisitation, browser design, hypertext.

1   Introduction

Demographics of web use show that daily page visits in the U.S. alone first exceeded one billion in October 1999 (Mack, 1999) and prior studies have shown that the 'Back' button accounts for up to 42% of user actions with web browsers (Catledge and Pitkow, 1995; Tauscher, & Greenberg, 1997). The Back button, therefore, is used many millions of times every day. The economies of scale clearly demand that it is easy to understand and efficient to use.

Prior work has shown that web-users commonly hold incorrect mental models of the stack-based behaviour of Back (Cockburn and Jones, 1996). Even in short navigational paths more than half of the subjects in the study were unable to predict whether Back would allow them to return to pages. Anecdotal evidence assures us that users are often confused about why pages 'disappear' off the set accessible with Back.

The two primary limitations of Back are that many users misunderstand its operation and that all recently seen pages cannot be revisited. Yet, despite its limitations, Back is heavily used. There are several factors contributing to Back's success (Cockburn and Greenberg, 2000):

- It can allow rapid return to recently visited pages (assuming they are still on the stack).
- It is robust. People can use it even with a naïve model of the way it works.
- It is 'ready-to-hand'. Unlike interface features such as the 'History list', which must be explicitly popped up when needed, there is little overhead in accessing Back because it is on constant display.
- It consumes minimal screen real estate.
- It is cognitively undemanding. Users can backtrack through pages using a simple 'click until the desired page is recognized' strategy.

Greenberg and Cockburn (1999) described an alternative 'temporal' behaviour for the Back and Forward buttons. The intention was to develop a behaviour that maintained the successful characteristics of Back's operation, but removed the problems of missing pages.

This paper describes an evaluation that compares the efficiency of the stack and temporal techniques. The implications of the research are clear—if the temporal scheme allows more efficient navigation without detrimentally affecting subjective satisfaction, then the developers of commercial browsers should consider re-implementing Back and Forward.

The following section describes both the traditional stack-based model and the temporal model. The evaluation method is then described, followed by the results, discussion and conclusions.

2 Back and forward behaviours

2.1   THE CURRENT STACK-BASED MODEL

Microsoft Internet Explorer and Netscape Navigator use the same stack-based model for web page navigation. In this model, there are two fundamentally different ways of displaying pages in the browser, which we call 'load' and 'revisit'. Pages are loaded when the user clicks on a link, types a URL, selects a 'Favorite' page, and so on. The effect of load is to add the page to the top of a stack of visited pages. Pages are revisited with the Back and Forward buttons, and the effect is to alter the position within the stack of visited pages. Each Back click displays the next page down the stack until the stack bottom is reached. Forward clicks display pages up the stack until the stack top is reached. The menus associated with the Back and Forward buttons allow users to directly revisit pages within the stack. These menus identify pages by showing their titles.

The critical feature of the stack behaviour is apparent when the user loads a page while within the stack (rather than at its top). The new page is added to the stack above the current position, and all pages above that stack position are removed. We use the term 'pruning' to describe this page removal process.

The 'Stack' column of Table 1 demonstrates this behaviour when navigating between pages **a**, **b**, **c**, and **h** in the simple web-site structure shown in Figure 1. The symbol → indicates navigation by loading (clicking on a link), and the ⇐ symbol indicates revisiting a page by clicking Back. The list on each row of the 'Stack' column shows the state of the stack, with the stack top on the right-hand end of the list. The page displayed in the browser is indicated by the underlined element. Row 5 of Table 1 shows that having navigated through **a→b→c⇐b⇐a**, the user is at the bottom of the stack (page **a** is displayed in the browser), with pages **b** and **c** on the stack above. Row 6 shows that pages **b** and **c** are popped off the stack when the user clicks on the link to page **h**.

2.2   THE TEMPORAL MODEL

The two main aims of the temporal model (Greenberg and Cockburn, 1999) are to ensure that all previously visited pages are accessible with Back and Forward, and that Back navigates through pages in the order in which they were last displayed in the browser.
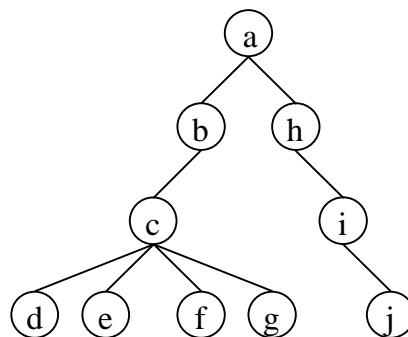
FIGURE 1: Web structure for tasks and examples.

TABLE 1: *Stack and temporal states on navigation.*

Row	Action	Stack	Temporal (main)	Temporal (internal)
1.	a	{<u>a</u>}	{<u>a</u>}	
2.	a→b	{a, <u>b</u>}	{a, <u>b</u>}	
3.	b→c	{a, b, <u>c</u>}	{a, b, <u>c</u>}	
4.	c⇦b	{a, <u>b</u>, c}	{a, <u>b</u>, c}	{b}
5.	b⇦a	{<u>a</u>, b, c}	{<u>a</u>, b, c}	{b, a}
6.	a→h	{a, <u>h</u>}	{~~a, b,~~ c, b, a, <u>h</u>}	

Like the stack-based technique, there are two semantics of page display (loading and revisiting) that are executed through equivalent interface actions. Loading a page causes it to be added to the end of a list that serves a function equivalent to the stack. Like the stack, Back and Forward menus can be used to directly revisit pages on the temporal list.

When the Back and Forward buttons are used, each revisited page is added to the end of an internal secondary list, which is maintained transparently to the user. Rows 4 and 5 of Table 1 show that when the user navigates Back from page **c** to **b**, and then from **b** to **a**, pages **b** then **a** are added to the secondary list (shown in the right hand column). The secondary list is continually updated while the user continues to revisit pages with Back and Forward. When the user next loads a page, the secondary list is concatenated to the end of the main temporal list, and any older duplicate pages are deleted from the new list. Finally, the new page is appended to the end of the list. Row 6 of Table 1 shows that on clicking the link to page **h** from **a**, the main temporal list is updated by concatenating the main and secondary lists. The older duplicates of pages **a** and **b** are then deleted (shown crossed out at the start of the list on Row 6 of Table 1).

2.3   THEORETICAL PROS AND CONS OF THE TEMPORAL SCHEME

There are three reasons for suspecting that the temporal scheme may improve navigation, and two for suspecting that it may make it worse.

The three theoretical advantages are as follows:

Reachability. Unlike the stack system, *all* pages visited during the browsing session can be revisited using the Back or Forward buttons. Furthermore, because all pages are on the Back or Forward menus, all pages can be reached with a single menu selection.

Flexibility. Because the temporal list contains a complete set of pages, there are often more 'sensible' navigation paths to any particular page. For instance, Row 6 of Table 1 shows that having navigated through **a→b→c⇦b⇦a→h**, the stack contains **{a, <u>h</u>}** and the temporal list contains **{c, b, a, <u>h</u>}**. When using the stack to return to page **c**, the only 'sensible' path is **h⇦a→c**. In the temporal system, the user can use the same technique (**h⇦a→c**), or three clicks of the Back button (**h⇦a⇦b⇦c**)**,** or the page can be directly selected from the Back menu (**h⤸c**).

Reduced decision-making. One of the advantages of the Back button, regardless of the algorithm underlying it, is that users can employ a cognitively light-weight 'click until the desired page is recognised' strategy for revisiting pages. With the stack this strategy can fail because the desired page may have been 'pruned'. In the temporal system, however, the strategy will succeed because all pages are accessible through a series of Back clicks.

Despite these theoretical advantages, the temporal system has two potential disadvantages.

Longer paths up the page structure. The pruning behaviour of the stack system favours navigation up the hierarchy of web pages—the parent of the current page can be reached with one click of the Back button. By never removing pages, the temporal system will produce longer backtracking paths to reach parent pages. Eventually backtracking paths will become impractically long. Similarly, the Back menu will eventually have to be truncated in the display (as current browsers do).

Transfer effects. Millions of users have experience with the current stack-based system. The proportion of users that correctly understand the technique is unknown, but there are risks that switching to the temporal technique will confuse many.

The evaluation described in the remainder of this paper aims to determine whether these theoretical advantages of the temporal system outweigh its weaknesses.

3   Evaluation method

The purpose of the evaluation is to compare the efficiency of the two alternative behaviours of the Back and Forward buttons. It also investigates differences in subjective opinions about the two interfaces. The evaluation is based around a series of realistic web-browsing scenarios that test the theoretical pros and cons of each interface. The scenario for the evaluation, as it was described to the subjects, was that they would be using a new release of Netscape Navigator. They were told that the behaviour of the Back and Forward buttons might, or might not, have been modified.

3.1   SUBJECT DETAILS AND TREATMENT

Thirty-four subjects participated in the evaluation. All were Computer Science students who use Netscape Navigator daily in their course work. Each evaluation lasted approximately one hour and was rewarded with five lottery tickets. Subjects were randomly assigned to either the stack or the temporal interface.

The subjects were informed that they would be asked to carry out nineteen short web-navigation tasks as quickly as possible. No training or instruction on the behaviour of the web browser was given, and they were told that the experimenter was unable to respond to questions about the system's behaviour.

A 'cueing interface' controlled the experimental condition and presented successive tasks to the subjects each time they clicked its 'Next Task' button. Each task was described by a short textual description. To help ensure that the subjects understood the task before starting, they were asked to read the task description out-loud twice. They were then asked to repeat the task to the experimenter without reading. They continued to read the task until successfully memorised.

When ready, the subjects were asked to click the "Go" button on the cueing interface. This raised the Netscape window, allowing them to carry out the task. The cueing interface with the textual description of the task remained visible throughout each task, allowing users to confirm the path if needed. The Netscape window was automatically iconified on completing each task.

To familiarize the subjects with this procedure and with the 'new' version of Netscape, five sample tasks were presented to the subjects. These sample tasks used a local web site familiar to the subjects. Data from these familiarization tasks was not analysed.

3.2   PROCEDURE

The experiment consisted of repeating six tasks in three different web sites. The tasks were selected to meet two criteria: first, they should reflect common browsing activities; second, they should expose the fundamental differences between the stack and temporal behaviours.

The finishing state of each task provided the starting state for the subsequent task and the same web-browsing session ran continually throughout the tasks (the browser was not reset or restarted). Having completed the six tasks in one site, the subjects carried out the same tasks in the second and third sites. They were not informed that the structure of the tasks was the same in each site. The web-site structure used in the tasks is shown in Figure 1. A final task (Task 7, third site) involved navigating back to a page visited in the first site. The order of the sites used was rotated per subject.

On completing each task, the subjects were asked to respond to two questions on a five-point Likert scale (1 disagree, 5 agree):
- The interface was efficient for the task.
- I was confused by the browser's bevaviour.

Having completed all tasks, two final questions were asked.

- The Back and Forward buttons in this version work the same way as those of my normal browser (Yes/No).
- If they replied "No" to the previous question, they were asked: "Compared to my normal browser, this interface is: worse, bit worse, same, bit better, better".

Comments were recorded after every task.

3.3 WEB SITES USED

The three sites used were as follows. Site one was a cut-down version of Boeing's website (www.boeing.com). Site two was a similarly cut-down version of Waikato University's Department of Computer Science (www.cs.waikato.ac.nz). Site three described the mammalian branch of the tree-of-life. It was constructed specifically for the experiment, and only provided four textual links per page, with no graphics.

In cutting down sites one and two, we removed navigational aids such as sidebar shortcuts from all pages because we wanted the evaluation to focus on the support provided by the browser rather than that of the pages. We also ensured that no page scrolling was required and that all links were clearly visible towards the middle of the pages. The subjects were informed of these conditions. Despite removing features from sites one and two, the pages included many graphical and textual items typical of most commercial web pages. In contrast, the Spartan pages used in site three were intended to act as a control to ensure that the visual properties of the other two sites did not interact adversely with either of the interfaces under test.

All pages were held on the local file system to ensure rapid and consistent display times.

3.4 APPARATUS

Netscape Navigator version 4.72 under the X-Windows system was used for both the stack and temporal interfaces. The keyboard was moved away from the subject, and a single-buttoned mouse was used. The subjects were therefore unable to create multiple browser windows or use keyboard shortcuts. The intention was to focus the evaluation on single-window 'point-and-click' navigation.

The visual appearance, mouse bindings and performance of the stack-based and temporal interfaces were identical to that of an unmodified version of Netscape Navigator version 4.72. Both interfaces were supported by overlaying unparented windows that provided exact visual replicas of the Back and Forward buttons over Netscape's actual Back and Forward buttons. The overlaid buttons moved seamlessly with the Netscape window when relocated or resized.

When the subjects clicked the Back or Forward buttons our program modified the state of its stack or temporal list and issued appropriate page display requests to the browser. The subjects could pop-up menus that were visually indistinguishable from those normally produced by Netscape by pressing and holding over either the Back or Forward buttons. The program also monitored navigation actions within Netscape, allowing us to log their actions and to update the state of the stack/temporal-list appropriately.

3.5 EVALUATION TASKS

Table 2 describes each task by showing the alternative paths that can be used to accomplish each task in each interface. It also shows the minimum number of navigational actions—link selections, Back/Forward clicks, and menu selections—needed to follow each path.

The nature of the tasks and the reasons for their selection are as follows.

Task 1: *Hub and Spoke Navigation.* This task is based on the common 'hub and spoke' browsing activity (Catledge and Pitkow, 1995) in which users visit a series of 'spoke' links off a 'hub' page, using the Back button to return to the hub after each page. The structure for this task (Figure 1) is from location **a** through **d**, **e**, **f** to **g**.

In the Boeing site the task was presented as follows: "You are interested in Boeing 747 aircraft. From Boeing's home page, follow the links to their 'Commercial aircraft' and onto the main page for the 747. From here there are links to four different types of 747: international, domestic, combi and freighter. Visit each of these pages."

Solving this task with either interface results in the same ideal path (Table 2).

Task	Stack AND Temporal	#	Stack	#	Temporal	#
1.	a→b→c→d⇐c→e⇐c→f⇐c→g	9				
2.	g⇐c→f	2			g⇐c⇐f	2
					g↯f	1
3.	f↯a	1	f⇐c⇐b⇐a	3	f⇐e⇐d⇐b⇐a (if g⇐c→f in Task 2)	4
					f⇐c⇐g⇐e⇐d⇐b⇐a (if g⇐c→f in 2)	6
4.	a→h→i→j⇐i⇐h⇐a	6				
	a→h→i→j ↯a	4				
5.	a→h→i→j or a⇨h⇨i⇨j	3				
	a↬j	1				
6.	j⇐i⇐h⇐a→b→c→g	6			j⇐i⇐h⇐a⇐b⇐ c⇐d⇐e⇐f⇐g	9
	j↯a→b→c→g	4			j↯g	1
7.	g₃↯a₁→h₁→i₁→j₁	4	g₃⇐a₃⇐g₂⇐a₂⇐g₂⇐a₁→ h₁→i₁→j₁	8	g₃↯j₁	1
					g₃⇐ a₂⇐ j₂⇐ i₂⇐ h₂…⇐ a₁⇐ j₁	12+
					Depends on prior path. Up to 27 ⇐s	

Task 2: *Spoke Revisitation.* This task investigates the subjects' ability to revisit a 'spoke' page, starting from a different 'spoke' page. The structure for this task is location **g** to **d** (Figure 1).

In the Boeing site the task was: "You are at the page describing the 747-freighter. Return to the page for the 747-international."

In theory, there should be no difference between the interfaces for this task. Both interfaces allow the user to revisit the page by clicking Back once, followed by selecting the required link. However, if the user assumes that the target page is on the Back list, then multiple clicks of the Back button (or the Back menu) will succeed in the temporal system, but fail in the stack-based system.

Task 3: *Parent Revisitation.* This task examines the subjects' ability to return to a parent (or home) page from a relatively deeply nested page. The structure of this task is from structural location **d** to **a** in Figure 1.

In the Boeing site the task was: "You are at the page describing the 747-international. Return to Boeing's Homepage."

The page-pruning behaviour of the stack system results in shorter paths than the temporal system for this task. The temporal system requires four or six clicks of the Back button, depending on which backtracking technique was used to page **c** in Task 2. The stack system requires only three Back clicks. Similarly, the temporal system will produce a longer Back menu than the stack system, requiring the user to read through more page titles to identify the target page. For these reasons, the Stack system should be more efficient.

Task 4: *Depth-First Search and Back.* This task compares the subjects' ability to follow a depth-first navigation path and return to the original root of their search. The structure for this task is from location **a** to **j** and Back to **a** in Figure 1.

In the Boeing site the task was: "You are at Boeing's homepage. Boeing makes a space-rocket called the Delta IV. Visit this page by following the links to 'Space Systems', 'Delta', and 'Delta IV'. Finally, return to Boeing's Homepage."

Both interfaces have the same paths for this task (Table 2).

Task 5: *Depth-First Forward Revisitation.* This task examines the techniques used to return to a recently viewed page. The structure for this task is from location **a** to **j** (Figure 1). In both interfaces, the most efficient way to solve this task is by using the Forward menu. We suspected, however, that many or most users would backtrack by link-selection rather than by Back/Forward.

In the Boeing site the task was: "Return to the page for the 'Delta IV' rocket."

As for task 4, both interfaces have the same paths for this task.

Task 6: *Within-Site Distant Revisitation.* This task is used to compare the support provided by the two interfaces for revisiting more distant pages within a web site. The structure for this task is from structural location **j** to **g** (Figure 1).

In the Boeing site the task was: "You are at the page for the 'Delta IV' rocket. Return to the page describing Boeing's 747-freighter commercial airliner."

In theory, the temporal system allows more rapid completion of this task by allowing users to select the page from the Back menu. However, in our experience many users are unaware of the Back/Forward menus. If the Back menu is not used, then users of the temporal system can still solve the problem by using the cognitively lightweight "click Back until the desired page is seen" strategy. Users of the stack system must first backtrack to the root page, and then click on appropriate links to reach the destination.

Task 7: *Cross-Site Distant Revisitation.* Tasks 1 to 6 were repeated for the three different web sites. Task 7 was carried out only once, after Task 6 in the third site. It examines the subjects' ability to revisit a temporally distant page. The subjects were required to navigate from the page at the third site's structural location **g** to the page at the first site's location **j** (Figure 1).

An example of the task is: "You are at the page describing Boeing's 747-freighter commercial aircraft. Return to the course outline page for the 'Introduction to Computer Science' course at Waikato University."

In theory, the temporal system allows more efficient task solution than the stack-based system—users can navigate directly to the page by selecting it from the Back menu. The same technique will fail in the stack-based system due to stack pruning. Consequently, users of the stack-system must switch from backtracking with Back to navigation using links. The "click Back until the desired page is displayed" strategy will succeed in the temporal interface, but only after between 12 and 27 clicks of the Back button, depending on prior paths taken.

4 Results

The experimental design depended on the subjects memorizing and following short navigational paths. The method of cueing tasks—with the subjects repeatedly reading and repeating their paths prior to starting—was successful, and almost all participants completed their tasks without checking the path description in the cueing interface.
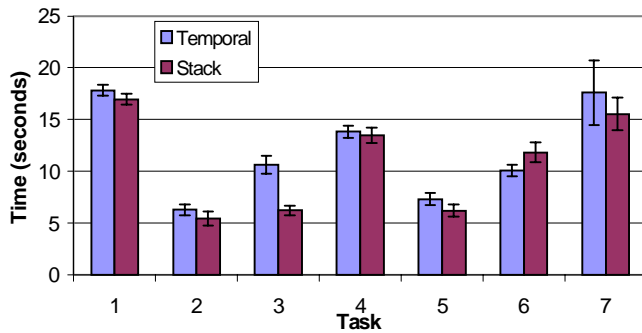
4.1 CHARACTERISING THE BROWSING ACTIVITY

Across both interfaces and all tasks, the mean task completion time was 10.8 seconds (σ 6.3) and the mean number of navigation actions was 4.6 (σ 3.5). The 34 subjects logged a total of 2989 navigational actions. Clicking on links accounted for 50% of actions, the Back button 39%, the Back menu 7.5%, the Forward button 2.5%, and the Forward menu 1.5%. These values are similar to those reported in prior work; Catledge and Pitkow (1995) reported values of 52% and 41% for links and Back, and Tauscher and Greenberg (1997) reported values of 50% and 30%. The similarity between the prior results and our values increases our confidence that the experimental tasks were representative of realistic browsing activities.
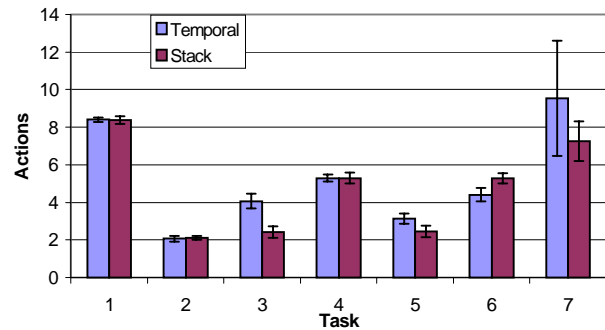
The percentage of Back actions was higher when using the temporal interface (43%) than when using the stack interface (34%). There was a corresponding difference in the percentage of link-selection actions in the two interfaces, with 45.5% and 54.5% for the temporal and stack interfaces. The higher percentage of Back actions in the temporal interface can be explained by its longer backtracking list.

4.2 TASK PERFORMANCE

The evaluation aimed to determine whether the predicted benefits of the temporal scheme would be realized without training. Two performance measures were recorded: the time taken to complete the tasks, and the number of actions taken. Figures 2a and 2b show the mean task completion times and the mean number of actions for each interface across the tasks. Task completion time was strongly correlated with the number of actions (Pearson's r = .76, p<.001), and consequently the results of statistical analysis for the two dependent variables are similar. The results reported below focus on task time, with action count values reported where relevant.

(a) Mean time per task.

(b) Mean actions per task.

FIGURE 2: Time and actions to complete Tasks 1 to 7. Error bars show one standard error above/below the mean.

The task time data from tasks one to six were analysed using a 2×6×3 mixed factor analysis of variance. The first factor, Interface (stack and temporal), was a between-subjects factor. The second and third factors—Task and Web-site—were within-subjects factors.

There was no significant difference between the two interfaces, with stack and temporal mean task completion times of 10.0 (σ 6.0) and 11.0 (σ 5.9) seconds: $F(1,32)=1.7$, $p=.2$. Figure 2 shows the relatively similar performance of the two interfaces for most tasks. There was a strongly significant main effect for Task ($F(5,160)=144.4$, $p<.001$), but this is unsurprising given the widely differing navigational paths followed in the six tasks.

There was a significant difference between the websites: $F(2,64)=26.6$, $p<.001$. The difference is due to rapid navigation through the graphically spartan Site three. Site three's mean task time was approximately 20% faster at 8.9 seconds (σ 5.0), compared to 11.4 (σ 6.3) and 11.2 (σ 6.2) for Sites one and two. The mean action counts used in the three sites were less varied, with Site three taking the same number of moves as Site two (means for Sites one, two and three of 4.6 σ 2.8, 4.3 σ 2.6, and 4.3 σ 2.6): $F(2, 64)=3.13$, $p=.05$. This suggests that the rapid performance in Site three was largely due to the eased visual search for links in its uncluttered pages. The alternative explanation—that the participants' became lost more often in the graphically rich sites—would have resulted in higher action counts as well as performance times in Sites one and two. There was no interaction between Interface type and Site ($F(2, 64)=.37$, $p=.69$), indicating no major difference in the ways the interfaces support navigation through graphically rich and spartan sites.

The absence of a significant main effect for Interface type does not imply that the interfaces were used similarly in all tasks. There was a significant interaction between factors Interface and Task: $F(5, 160)=7.3$, $p<.001$. The interaction is clear in the crossover-effect comparison of Tasks 3 and 6 in Figure 2. In the 'parent revisitation' Task 3, the stack interface out-performed the temporal interface, with time and action-count means of 6.2 seconds and 2.4 actions, compared to 10.6 seconds and 4.1 actions with the temporal interface. The efficiency of the stack is due to its pruning behaviour, which results in shorter backtracking paths up the website structure. The stack system required three actions for Task 3 (**f⇦c⇦b⇦a**), while the temporal system needed four or six Back clicks, depending on the backtracking technique used in Task 2 (**f⇦c⇦g⇦e⇦d⇦b⇦a** or **f⇦e⇦d⇦b⇦a**): see Table 2 for details. In the 'within site distant revisitation' Task 6, however, the efficiency comparison is reversed, with the temporal interface producing slightly lower time and action count means (10.1 seconds and 4.2 actions) than the stack (11.8 seconds and 5.3 actions).

The 'cross-site distant revisitation' Task 7 was conducted by each participant once after completing Task 6 in the third site used. Data from this task was not normally distributed (described further in Section 4.4). A Mann-Whitney U-Test showed no significant difference between the interfaces ($U=135$, $N_1=N_2=17$, $p=.37$), with mean task completion times of 15.5 (σ 6.5) and 17.6 (σ 12.8) seconds for the stack and temporal interfaces. Although the means are similar, the high standard deviations (particularly for the temporal interface) indicate that some users

solved the task quickly, while others struggled. Reasons for these widely distributed data are provided in Section 4.4.

4.3 SUBJECTIVE RESULTS

After all of the tasks were complete, the participants were asked to respond to two five-point Likert scale questions (1 disagree, 5 agree): "I understood the navigation behaviour" and "The interface was efficient for the task." Ratings for understanding were high, with means of 4.6 and 4.4 for the stack and temporal interfaces, but showed no significant difference (Mann-Whitney, U=112, p=.11). Ratings for efficiency were lower (3.8 and 3.7 for stack and temporal), again with no significant difference: U=128, p=.29.

Eight of the temporal subjects replied 'No' to the question "The Back and Forward buttons work the same way as my normal browser". When asked what the difference was, six correctly stated that the set of pages accessible through Back was complete. Two of these stated that the new system was worse because "going back takes too long" and "it's slower"; two stated it was neither better nor worse, and two stated it was better because "it doesn't lose pages".

Three of the stack users stated that the interface was different. Only one of these provided a reason, stating that the 'new' interface was worse because it gave a "different order of pages" and that he "knew where things were supposed to be, but when they weren't there it slowed me down". The fact that fewer than half of the temporal subjects (all of whom were Computer Science students) noticed the modified behaviour of the browser provides further evidence that users' mental models of Back behaviour are weak.

The most frequently occurring comment from users of both interfaces was that identifying pages from their titles in the Back menu was difficult. This problem was worse in the temporal interface because it produces longer menus. There are many problems with representing page titles in the Back menu. These include pages that have no title and truncation of long titles. Title truncation often removes the essential page identification information.

4.4 USE OF THE BACK MENU

Analysing the ways in which participants solved the tasks provides interesting insights into opportunities for enhancing the efficiency of web browsers. The intention here is not to determine whether the two interfaces are statistically different from one another, but rather to observe patterns of behaviour that result in efficient and inefficient navigation.

Use of the Back menu (supported by both the normal stack and temporal interface) was strongly correlated with the efficiency of task completion in Task 7 (Spearman rank correlation between action count and menu use, rho=-.83, p<.01). The mean number of moves used in Task 7 by the 23 participants who used the back menu was 3.0 (σ 2.4), compared with 20.7 (σ 7.8) for the eleven participants that did not. Task times were similarly biased, with means of 11.0 seconds (σ 4.5) and 27.7 (σ 9.3) for menu and non-menu users. The correlation between action count and menu use holds for both the stack (rho=-.89) and temporal (rho=-.86) interfaces.

In comparison to the stack system, temporal interface exaggerated both the efficiency of the skilled users and the inefficiency of the less skilled ones. Use of the Back menu was the key determinant in characterizing 'skilled' versus 'unskilled' navigation. With the temporal interface, seven of the subjects completed the task very quickly with a single menu selection; one subject took two actions, and two subjects took three actions. Of the remaining six temporal users, one took five actions, and the others took more than 25 actions. The range of action counts for the stack users is also bimodal (but with a smaller range), with eleven of the subjects taking four or five actions, and the remaining six taking thirteen or fourteen actions each.

Figures 3a and 3b show the time and number of moves used (y-axis) plotted against a sorted order of participants completing Task 7 (x-axis). It shows that the temporal system allowed most of the participants to solve the task rapidly (with comparison to the stack users), but that the least efficient temporal users were dramatically slower the worst stack users. The figures also show that all eleven efficient users with the temporal system used the back menu, and that none of the inefficient users used it.

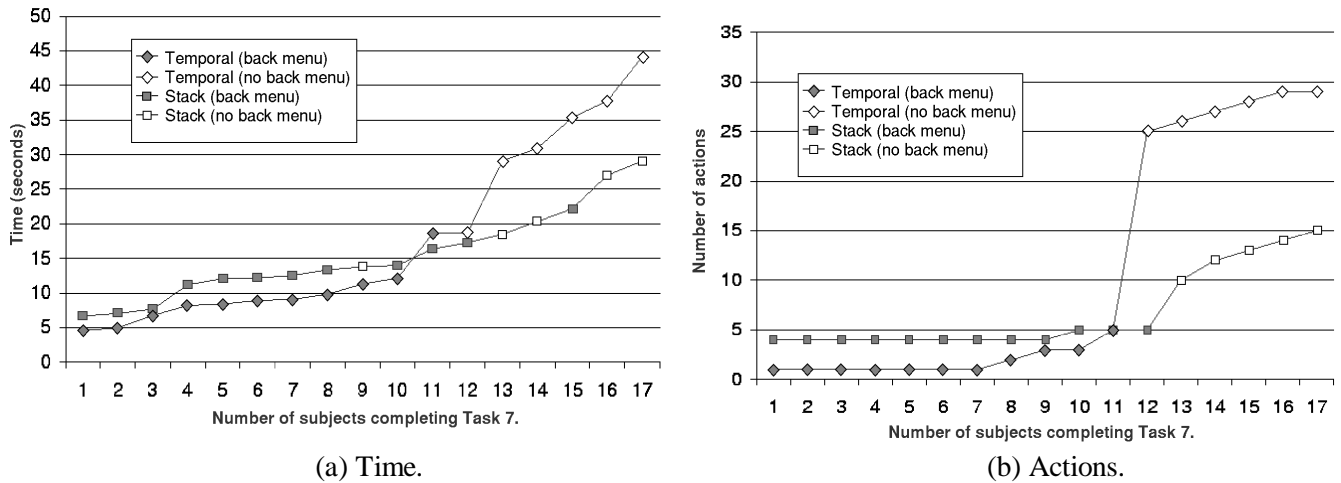|	
(a) Time.	(b) Actions.

FIGURE 3: Time and actions taken by the subjects to complete Task 7 using the stack and temporal interfaces. The x-axis is a sorted order of participants completing the task. Filled shapes show back menu users; unfilled ones show non-back menu users.

4.5 STUDY LIMITATIONS

There are several limitations in this experiment. First, the subject-base of Computer Science students is obviously not representative of all web users, and there is a risk their high level of computer competence may have influenced the results. Second, each evaluation session lasted approximately one hour and visited a total of thirty different web pages. It is not clear how the temporal scheme will scale to prolonged use over a series of days, weeks and months. Third, we did not train the subjects using either interface. The rationale was that we wished to investigate performance with the temporal system as though the user had immediately started to use a newly installed version of the browser. We suspect that we could have substantially improved the subjects' performance with the temporal system by advising them to make heavy use of the Back menu. Similarly, we suspect that with more experience with the temporal interface the subjects may have become more efficient.

5   Discussion

To summarise the results, there was no overall significant difference between the stack and temporal interfaces. There were, however, differences in the relative performance of the interfaces across different types of navigational task (a significant interaction between Interface type and Task type). Most notably, the temporal system was inefficient when returning to parent pages. Analysis of the participants' solution strategies shows that both stack and temporal interface users were efficient when the Back menu was used. For distant navigation tasks, the temporal system was highly efficient when the Back menu was used (in comparison to the stack interface), but highly inefficient when it wasn't.

There are several potentially important implications from these findings. First, considering that the subjects had no training and little experience with the temporal system, it is encouraging that temporal users performed comparably to the stack users in most tasks. It is possible (though not yet tested) that, with more experience, the temporal interface could out perform the stack interface for many navigational tasks. Second, the evaluation has shown that the weakness of the temporal scheme lies in its support for returning to parent pages. The mean task completion time for temporal interface users for Task 3 "parent revisitation" was 71% greater than that of the stack users. Possible strategies for overcoming this limitation are discussed in Section 5.1. Third, the observation that participants solved distant navigation tasks most rapidly when using the Back menu implies that web browsing could be made more efficient if browsers encouraged users to make greater use of the Back menu. Fourth, several participants stated that identifying pages on the Back menu was difficult. Easing this problem would probably further improve the Back menu's efficiency. Finally, the high level of efficiency for distant retrieval with the
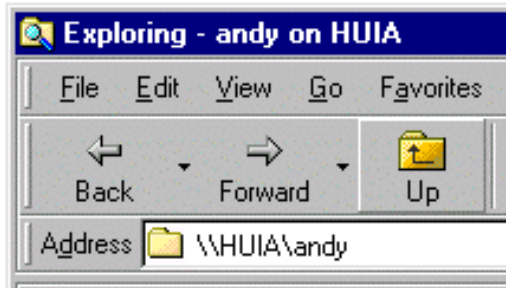
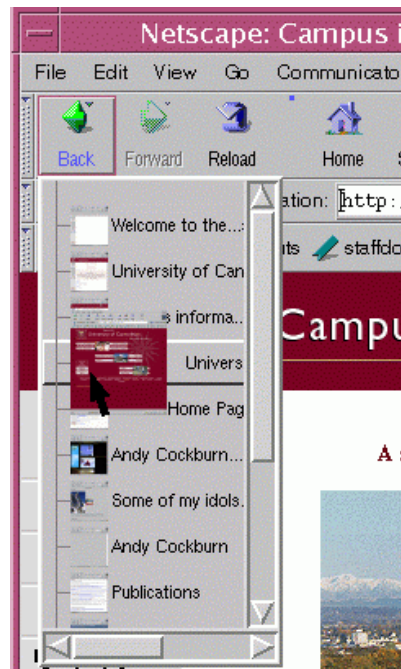FIGURE 4: Back, Forward and Up navigation in Microsoft Explorer.

FIGURE 5: Our enhanced Back-menu, with zooming thumbnail images of each web page visited.

temporal version of the Back menu indicates that enhancements to the temporal system may be able to improve web navigation.

5.1 DESIGN ENHANCEMENTS AND FURTHER WORK

In the introduction to the paper we stated: "if the temporal scheme allows more efficient navigation without detrimentally affecting subjective satisfaction, then the developers of commercial browsers should consider re-implementing Back and Forward." Although the results of this experiment do not support commercial deployment of the current version of the temporal system, the high level of efficiency attained by the temporal Back menu users is interesting, and worth further study.

The biggest weakness of the temporal system was in supporting navigation to parent pages. An additional "Up" button, similar to that used in file browsing systems such as Microsoft Windows Explorer (see Figure 4), could be used to ease this problem. In applying this technique to web browsers, the Back and Forward buttons would support the temporal model described in this paper, while the "Up" button would allow stack-based retrieval of parent pages. The major limitation of this approach is that users would have to consciously choose between different revisitation buttons, removing the non-cognitive "click until the desired page is seen" benefits of the current two-button scheme.

In our ongoing work we are investigating enhancements to the temporal scheme that aim to improve the effectiveness of the Back menu. The temporal Back menu has the potential to improve all web page revisitation, including that to parent pages, because it makes all pages accessible. In order to ease the problem of identifying pages on the menu, our interface (Figure 5) captures a thumbnail image of every page displayed in the browser (Cockburn and Greenberg, 2000). Pointing at any thumbnail in the menu causes it to zoom to an easily recognised size.

In contrast to the extensive research into web site design (for example, see the taxonomy of web site structures Gillenson, Sherrel and Chen's (2000) and Nielsen's (2000) web design guidelines), there has been surprisingly little research into user interaction with web browsers.

In attempting to understand the user's tasks when browsing the World-Wide Web, Byrne, John, Wehrle and Crow (1999) videotaped the day-long browsing behaviour of eight participants who were encouraged to verbalise their actions and tasks. Results from the video analysis show that users spend large proportions of their time waiting for pages to download and scrolling through pages. Other results confirm those of other studies (summarised below) that users make little use of advanced browser components such as history lists.

Another way to understand user action with the browser is to empirically characterise actions by logging low-level events at the browser. The studies by Catledge and Pitkow (1995) and Tauscher and Greenberg (1997) showed that link selection and use of the Back button are the most common page navigation actions (neither study measured scrolling actions within pages). They also showed that other navigation features such as bookmarks, the Forward button, and History lists were all lightly used. A more recent study by Abrahms, Baecker and Chignell (1998) shows that users now maintain large bookmark collections, but does not reveal whether they are often used to access pages. Finally, Cockburn and McKenzie (2001) update the earlier empirical characterisations of web use, revealing that web page revisitation is even more prevalent than earlier reported, with 81% of page visits being to pages previously seen by the user.

The Glabster system (Kaasten and Greenberg, 2001; Kaasten, 2001) demonstrates how page revisitation facilities including Back/Forward, Favorites and History can be integrated in a single sidebar temporal utility. Like our system (Figure 5), Glabster attempts to ease the problems of page identification through a variety of cues, including automatically captured thumbnail images of pages that zoom on mouse-over events.

6   Conclusions

The Back button on web browsers is used many millions of times every day, yet its stack-based behaviour is widely misunderstood. This paper described the evaluation of an alternative temporal scheme for supporting the Back and Forward buttons that has been proposed in prior work. The primary difference of the temporal scheme is that it maintains a complete temporally ordered list of pages that the user visits.

The evaluation was based on an 'out of the box' scenario in which subjects used a 'new' version of a web browser that either supported the current stack-based mechanism or the new temporal scheme. The results showed that the relative efficiency of the interfaces differed across different types of navigational tasks. In particular, the temporal system performed poorly for navigating back to parent pages, but relatively well for more distant navigation tasks. Interestingly, the temporal scheme caused extreme usage patterns, with the subjects either solving tasks very efficiently or very inefficiently, depending on whether they used the Back menu. The combination of the Back menu and the temporal scheme was particularly rapid because, unlike the stack system, the destination page was always available on the list of previously visited pages.

The observation that subjects who used the Back menu with the temporal system were particularly efficient is an important one. It implies that web navigation could be improved by promoting use of temporally ordered complete lists of recently visited pages. We are developing prototype systems that provide this functionality, and further work will focus on evaluating their effectiveness.

Acknowledgements

Many thanks to the anonymous reviewers of this paper for their extremely valuable comments.

References

ABRAMS, D., BAECKER, R. & CHIGNELL, M. (1998). Information Archiving with Bookmarks: Personal Web Space Construction and Organization, *in Proceedings of CHI'98 Conference on Human Factors in Computing Systems*, Los Angeles, April 18–23. 41–48.

BYRNE, M., JOHN, B., WEHRLE, N. & CROW, D. (1999). The Tangled Web We Wove: A Taskonomy of WWW Use, *in Proceedings of CHI'99 Conference on Human Factors in Computing Systems,* Pittsburgh, May 15–20. 544–551.

CATLEDGE, L. & PITKOW, J. (1995). Characterizing Browsing Strategies in the World Wide Web, *in Computer Systems and ISDN Systems: Proc. 3rd International WWW Conference.* Vol. 27. 1065–1073.

COCKBURN, A. & GREENBERG, S. (2000). Issues of Page Representation and Organisation in Web Browser's Revisitation Tools. *Australian Journal of Information Systems 7*, 2, 120–127.

COCKBURN, A. & JONES, S. (1996). Which Way Now? Analysing and Easing Inadequacies in WWW Navigation. *Int. J. Human-Computer Studies 45,* 1, 105–129.

COCKBURN, A. & MCKENZIE, B. (2001). What Do Web Users Do? An Empirical Analysis of Web Use. *Int. J. Human-Computer Studies 54*, 6, 903–922.

GILLENSON, ML., SHERRELL, DL. & CHEN, L. (2000). A Taxonomy of Web Site Traversal Patterns and Structures. *Communications of the Association for Information Sciences, 3*, 4, Article 17.

GREENBERG, S. & COCKBURN, A. (1999). Getting Back to Back: Alternative Behaviors for a Web Browser's Back Button. *in Proceedings 5th Conference on Human Factors on the Web.* http://zing.ncsl.nist.gov/hfweb/.

KAASTEN, S. (2001). Integrating Back, History and Bookmarks in Web Browsers. *M.Sc. Thesis. Department of Computer Science,* University of Calgary, Canada.

KAASTEN, S. & GREENBERG, S. (2001). Integrating Back, History and Bookmarks in Web Browsers. *in Extended Abstracts of CHI'01 Conference on Human Factors in Computing Systems.* 379-380.

MACK, J. (1999) A Net Record: 1 billion page views per day. www.zdnet.com/eweek/stories/general/0,11011,2398564,00.html

NIELSEN, J. (2000) *Designing Web Usability: The Practice of Simplicity.* New Riders Publishing.

TAUSCHER, L. & GREENBERG, S. (1997). How People Revisit Web Pages: Empirical Findings and Implications for the Design of History Systems. *Int. J. Human-Computer Studies, 47*, 1, 97–138.