Toward Modeling Auditory Information Seeking Strategies on the Web

Shari Trewin, John Richards, Rachel Bellamy
IBM T. J. Watson Research Center
P.O. Box 704, Yorktown, NY 10598 USA
trewin@us.ibm.com
ajtr@us.ibm.com
rachel@us.ibm.com

Bonnie E. John
Human Computer Interaction Institute
Carnegie-Mellon University.
5000 Forbes Avenue
Pittsburgh, PA 15213
bej@cs.cmu.edu

John Thomas, Cal Swart, Jonathan Brezin
IBM T. J. Watson Research Center
P.O. Box 704, Yorktown, NY 10598 USA
jcthomas@us.ibm.com
cals@us.ibm.com
brezin@us.ibm.com

CHI 2010, April 10–15, 2010, Atlanta, Georgia, USA.
ACM 978-1-60558-930-5/10/04.

Abstract
Human performance models based on information foraging theory have proved capable of predicting navigation behavior on the Web. They can therefore provide a useful tool for Web site design. They may also be effective for modeling auditory navigation within a single Web page. Designers often struggle to accommodate this sort of access, different as it is from their own experience. As a step toward realistic simulations based on models of auditory Web access, we describe information seeking strategies observed in people with visual impairment using screen reading software for Web navigation tasks. We outline one example strategy for approaching a new Web page that, guided by information foraging theory, may expose access barriers that current design tools miss.

Keywords
Accessibility, visual impairment, cognitive modeling

ACM Classification Keywords
H.5.2. [Information interfaces and presentation]: User Interfaces---*Evaluation/methodology, Voice I/0*;
H.1.2. [Models and Principles]: User/Machine Systems---*Human information processing*;

General Terms
Design, Human Factors

Introduction

Screen readers are sophisticated software tools that enable people with visual impairment to use personal computer applications such as Web browsers. They provide an auditory presentation that is fundamentally different from the visual experience.

Web page design has a profound impact on the auditory usability of a page, but Web designers struggle to accommodate this mode of access. In part this is due to lack of understanding of the auditory experience, and in part because of the difficulty of using screen readers in the testing process. Developers often have little experience with screen readers and lack strategies for using them effectively. Their deep knowledge of the page contents can also be a disadvantage: they may fail to notice what to a screen reader user is a promising-sounding minor link, because they already know the "obvious" path to the primary content.

A simulation of screen reader access to a Web page that is based on models of human behavior would be additionally valuable since research has shown that the majority of Web accessibility problems cannot be detected by the static content analysis used by current Web-checking tools [4]. By taking a task-oriented, behavioral approach, these simulations could move testing beyond standards compliance towards genuinely usable access.

In this paper we take a first step towards this goal by outlining auditory Web page navigation strategies, describing one such strategy in detail, and suggesting that a simulation of this strategy would expose accessibility problems that current tools do not.

Auditory Access to the Web

Web pages are generally designed to be seen. Layout conventions, size, color and formatting allow viewers to quickly scan the content and focus in on items of interest. Auditory access to the same page is a fundamentally different experience. A screen reader presents Web content as a linear sequence, based on the underlying structure encoded in the HTML markup, and converts each item to a string that is spoken by a speech synthesizer. Figures 1 and 2 illustrate visual and auditory representations of the CHI 2010 home page. In figure 1, the viewer's attention is drawn to the prominent images and the links above them providing direct access to the main content. Figure 2 presents a very different ordering of items, with all of the left navigation links appearing before the main content.

figure 1. CHI 2010 Home Page to a sighted visitor.

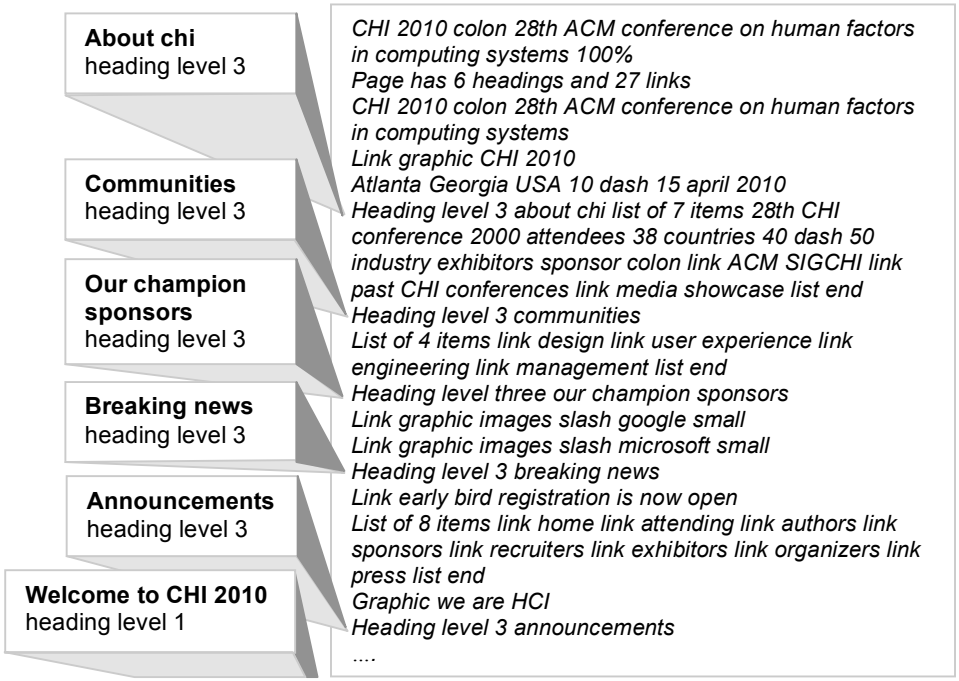

figure 2. CHI 2010 Home Page through a screen reader.

Screen reader users must move through this content without the benefit of the visual clues suggesting its relative lack of importance. Finding a specific section could be time consuming unless there is an easy way to navigate to it.

Screen readers provide a rich set of commands for navigating through this otherwise undifferentiated mélange, and screen reader users often develop effective strategies for using these commands. In addition to simply having the entire page read out, users can move to the next or previous lines, jump over a number of lines, or step through items of a certain type such as headings, links, or form fields. There are also search commands with which a user can go directly to a known phrase within a page.

Understanding Screen Reader Strategies

One Japanese study of five screen reader users performing Web shopping tasks [8] found that 90% of Web page navigation commands were moves to the next line or the next link, and participants made little use of headings or 'skip to main content' links. A British study of seven people performing information search on the Web [11] found that while inexperienced screen reader users relied heavily on item by item scanning, more experienced users took advantage of commands related to the structure of the Web page. In a recent survey of 665 screen reader users carried out by WebAim, 50% of respondents cited navigation through page headings as the first thing they do when trying to find information on a lengthy Web page [10].

Takagi et al [8] hypothesize that information foraging theory [7] can be applied to auditory Web page navigation. This theory states that people search for information by assessing the 'information scent' of the content at their current position, and moving towards areas with higher scent. Their participants were observed to change direction, follow links, or switch from jumping to scanning by lines when encountering items with seemingly high information scent.

Observations of Screen Reader Users

To supplement these reports, and to better understand navigation strategies at the level of detail needed to create behavioral models, three screen reader users were each interviewed and observed by a team of three

researchers at their place of employment. One used their own desktop setup, while two used a lab machine. They used version 9 of JAWS, their usual screen reader, and Internet Explorer 8, their usual Web browser. All had used screen readers for at least 14 years.

Tasks and Approach

We asked the users to perform four tasks, (1) use a website they were familiar with (e.g., a news site), (2) use Google to find the number of people living in a nearby city, (3) find a 'Buy tickets' link on the home page of Edinburgh Castle in Scotland, and (4) perform Blackmon's online encyclopedia search task. This was originally used to validate the AutoCWW model[2], allowing direct comparison to prior work in visual information foraging.

Our procedure was to interview each participant and then to observe and take notes while the participant performed the tasks just described. We asked everyone to browse in their usual manner. They could not always talk to us as they worked, because that would interfere with their listening to the screen reader. Instead, at natural pausing points we interrupted and asked them to describe the set of commands they had just used, and reasons for choosing that strategy.

Strategies for Familiar Web pages

When tackling a familiar Web page, very efficient strategies were observed. Examples include jumping to the end of the page and working backwards when that was quicker, or using block or heading navigation and listening to just enough content to keep track of the current position. As in prior studies [8, 11], *ad hoc* navigation landmarks were important. For example one person searched for the phrase 'results found' to jump to the beginning of an eBay search result list. In addition, people vary in the range of commands they use. One person used many different screen reader commands, e.g. jump to the next graphic, form field, link or visited link. Another relied predominantly on line, link and form field navigation.

Strategies for Unfamiliar Pages

When tackling unfamiliar Web pages, participants asserted in the pre-task interview that they might listen to the whole page first, but were not observed to do this in practice. Instead, all three often listened to the headings to get an overview of the page. If looking for a specific kind of page element such as a form field, all were observed to focus exclusively on items of this type. This strategy allows them to reach the target item more quickly but with a loss of surrounding context.

There was some evidence that strategies were selected based on a mental model of the type of page. While looking for information, the model was refined and strategies updated accordingly. For example, after following a 'Buy tickets' link, one person started exploring the new page by jumping to the next form field. He commented that he was expecting to be able to enter credit card information. Finding no form fields, he listened to the headings instead. All participants also frequently skipped past items at the top of the page without reading them. One commented that he did not expect to find useful content at the top of a page. We observed that alphabetical ordering for long lists of items allowed people to navigate in blocks of lines instead of single lines, since they knew when they had gone too far. Consistent with information foraging theory, participants were observed to follow promising links without first listening to all of the possibilities.

Modeling Auditory Access Strategies

We believe that a simulation tool could encode a set of realistic and appropriate screen reading strategies representing basic and expert approaches to both familiar and unfamiliar pages. These modeled strategies would support a designer in understanding how their site might be traversed and would allow rapid exploration of a number of design questions such as "how long would this task take for an expert user?", or "would navigation be improved if I changed the order of these items?".

Several computational cognitive models of information seeking on the Web, incorporating information foraging theory, have been shown to make good predictions of skilled users' interaction with Web sites [2,3,9]. Of these, CogTool-Explorer [9], using the ACT-R cognitive model [1], is a promising basis for modeling auditory access strategies. It incorporates SNIF-ACT 2.0, a computational cognitive model of how people use information scent cues to make navigation decisions on the Web. In SNIF-ACT, items on a web page are evaluated one by one, and a decision is made either to "satisfice" (choose the best seen so far) or look further. This behavior is, as we described above, consistent with empirical observations of auditory Web navigation.

In CogTool-Explorer, headings, and then items within headings, are evaluated in the order dictated by Halverson and Hornoff's Minimal Model of Visual Search [6]. The auditory access models we are developing can be thought of as alternatives to their approach. Our models differ in that they use a knowledge-based strategy, whereas the minimal model of visual search relies only on proximity.

Example: Navigating Unfamiliar Pages

Consider the problem of auditory information search on an unfamiliar Web page by a moderately experienced Web and screen reader user. Here we illustrate one simple strategy. We anticipate future work will articulate more complex strategies that take into account variables such as users' expertise and content complexity. The simple strategy described below is based on a core set of page navigation actions that are commonly used in practice, does not use prior knowledge of the page, and is driven by decisions that take advantage of information scent:

1. Listen to the page title and the first line of the page.
2. If the page title has strong information scent, and the first line on the page is a 'skip to main content' link, then follow the link to the start of the main content section of the page, else go on to step 3.
3. Traverse page headings to identify and move to the most promising section, perhaps satisficing and not listening to all the headings first. If headings are not present, treat the entire page as a single section.
4. Explore within a section by moving through the lines in 1-line steps and blocks of lines (both forwards and backwards). The decision whether to use line-level or block-level navigation is based on information scent. When scent is low, the model may jump ahead, while the lines around a high scent line will be explored one by one.
5. Repeat from step 3.

A typical user would know a variety of screen reading strategies. Based on their expectations of the page complexity and accessibility, they would select one,

monitor its effectiveness, and switch strategies when necessary. While a complete model would include meta-strategies like this, even models of simple, strategies, embedded in a suitable tool, would allow a designer to better understand the likely navigability of their Web site when presented through a screen reader.

Summary and Next Steps

Screen reader users employ a wide range of strategies for auditory navigation of Web pages. We propose to leverage CogTool-Explorer's implementation of information foraging theory to implement these strategies, and observed meta-strategies. Our goal is to provide a developer tool that simulates screen reader user performance, exposing usability problems that impact important information seeking tasks.

One simple strategy was described, representing a moderately experienced individual approaching an unfamiliar Web page. With multiple strategies built into a modeling tool, developers could explore the ramifications of their designs for screen reader users with different skills and preferences, thereby identifying and prioritizing auditory access problems such as bad alternative text for images, missing headings, and misleading links.

Put another way, while current tools can test for the *presence* of features like headings, auditory information foraging models also have the potential to assess the *quality* of these features.

Acknowledgements

We thank Lighthouse International, and the three screen reader users who spent time demonstrating their Web navigation strategies to us.

References

[1] Anderson, J. and Lebiere, C. The atomic components of thought. Erlbaum, USA, 1998.

[2] Blackmon, M., Kitajama, M. and Polson, P. Tool for accurately predicting website navigation problems, non-problems, problem severity, and effectiveness of repairs. In *Proc. CHI 2005*. ACM Press (2005), 31-40.

[3] Chi, E., Rosien, A., Supattanasiri, G., Williams, A., Royer, C., Chow, C., Robles, E., Dalal, B., Chen, J., & Cousins, S. (2003). The bloodhound project: automating discovery of web usability issues using the InfoScent™ simulator. In *Proc. CHI 2003*. ACM Press (2003), 505-512.

[4] Disability Rights Commission. *The Web: Access and Inclusion for Disabled People*. Disability Rights Commission, London, UK, 2004.

[5] Fu, W. and Pirolli, P. SNIF-ACT: A cognitive model of user navigation on the World Wide Web. *Human-Computer Interaction*, 22, 4 (2007), 355-412.

[6] Halverson, T. and Hornoff, A. A minimal model for predicting visual search in human computer interaction. In *Proc. CHI 2007*, ACM Press (2007), 431-434.

[7] Pirolli, P. and Card, S. Information foraging. *Psychology Review* 106, 4 (1999) 643-675.

[8] Takagi, H., Saito, S., Fukada, K. and Asakawa, C. Analysis of navigability of web applications for improving blind usability. *ACM Transactions on Computer-Human Interaction* 14, 3, (2007), Article 13.

[9] Teo, L. and John, B. E. Towards a tool for predicting goal-directed exploratory behavior. *Proc. Human Factors and Ergonomics Society 52nd Annual Meeting* (2008), 950-954.

[10] WebAIM. Screen Reader User Survey Results, 2009 http://www.webaim.org/projects/screenreadersurvey2/

[11] Yesilada, Y., Stevens, R., Harper, S. and Goble, C. Evaluating DANTE: Semantic transcoding for visually disabled users. *ACM Transactions on Human-Computer Interaction* 14, 3, (2007), Article 14.