CHI 2002 Workshop:
Automatically Evaluating the Usability of Web Sites
Workshop Date: April 21-22, 2002

Automated Cognitive Walkthrough for the Web (AutoCWW)

Position Paper by Marilyn Hughes Blackmon

Marilyn Hughes Blackmon, Peter G. Polson, and Clayton Lewis, University of Colorado at Boulder, USA
Muneo Kitajima, National Institute of Advanced Industrial Science and Technology (AIST), Japan

Role of Marilyn Hughes Blackmon in our research group. Over the past two years my role in the collaborative research has been to carry out the cognitive psychology research required to develop and parameterize the Cognitive Walkthrough for the Web (CWW) (Blackmon, Polson, Kitajima, & Lewis, 2002). The CWW is a usability inspection method designed to mesh well with web development procedures and evaluate how well websites support users' navigation and information search tasks. Our goal for the CWW is to transform it into the AutoCWW. My role in the collaborative research is to test the predictions of CWW and AutoCWW against the behavior of diverse groups of actual users, parameterizing the CWW/AutoCWW to improve the accuracy of its predictions until they are highly reliable.

How AutoCWW Can Promote Universal Usability of Websites

Press reports and government studies describe the enormous potential of the World Wide Web to serve the information and learning needs of a broad spectrum of citizenry, including educational outreach in such areas as basic skills tutoring or medical/health information. These hopeful visions, however, are compromised by evidence suggesting that relatively high levels of education are a prerequisite for navigating and comprehending informational/instructional texts on current-generation complex websites. The Pew Internet Project (2001) survey found a dramatic, two-to-one disparity in Internet use between persons who never attended college and those with less education. Among American citizens with a high school education or less, only 37% use the Internet. In stark contrast, 71% of Americans who have completed some college use the Internet, and that percentage rises to 82% for the subset who have completed a college or graduate/professional degree.

Three goals for closing the digital divide emerged from the ACM Conference on Universal Usability 2000 Conference, and our research group is addressed the goal of "bridging the gap between what users know and what they need to know" (http://www.UniversalUsability.org/definition/introduction.html). Many citizens have difficulty navigating websites because of low reading ability, insufficient background knowledge, and/or a cultural heritage differing significantly from the dominant culture. For the next three years I will be funded half time by NSF to conduct research that contributes to closing these knowledge and skill gaps by extending the Cognitive Walkthrough for the Web (CWW). Web developers have trouble predicting the behavior of such users, because developers do not themselves experience difficulties navigating complex websites to find information. Thus, the AutoCWW fills a role not filled by other available automated usability evaluation tools.

Both CWW and the AutoCWW rely on Latent Semantic Analysis (LSA) to (1) objectively estimate the degree of semantic similarity (information scent) between representative user goal statements and heading/link texts on each web page, and (2) assess the probability that a particular user group can effectively learn from texts they discovered by navigating the website. Under NSF funding I will conduct a series of experiments that will first parameterize the CWW for college-educated users navigating informational/instructional websites and then parameterize the CWW to validly simulate navigation behaviors of diverse user groups, including users with 6th-, 9th-, and 12th-grade reading levels in English. This is feasible because LSA currently has semantic spaces representing these levels of educational attainment for American-educated individuals.

Carrying out CWW analyses by hand is time-consuming, particularly if a web developer wants to predict the navigation behavior of several diverse user groups for a wide variety of user goals. There is only one realistic solution to this: automating the CWW to create the AutoCWW. Latent Semantic Analysis (LSA) is a key component of the AutoCWW, because LSA already a computer-based and because it is possible to program submission of texts to LSA, automatically generating predictions of which heading a user will focus on and which link within that heading a user will select when pursuing a given goal.

The AutoCWW Meshes Well with Web Development Procedures

There are several contrasting ways for users to interact with websites – surfing to visit various websites and pages that focuses on a topic, retrieving pages using a search engine by entering queries, or browsing pages on a specific topic to gather information. The CWW deals with a situation where a user starts navigating from the homepage of a website with the intention of locating a page that satisfies the user's goal. For example, a user goal might be to search on a

medical website for information about all reliable treatments for a specific disease.

The first task of the development team is to identify the various groups of intended users and apply ethnographic methods to understand the users likely to visit the website and the broad range of goals they are likely to bring to the site. The developer needs to compile a diverse, representative set of detailed user goal statements of at least 100-200 words each. AutoCWW will submit each of these detailed goal statements to LSA to estimate the similarity between each goal and the link label texts on the page. This step produces predictions of which link the user would probably select in pursuit of his/her goal.

What is unique about the AutoCWW is its capacity to determine if the link label system on the web page under development will correctly route each and every user goal in the set of test goals to the correct subordinate-level web page. The CWW uses LSA to operationalize and quantify "scent" or "information scent": cues that discriminate choices that are more likely to lead to a given goal from choices that are less likely. As the user moves from the home page toward the target page, the user must be able to perceive that the "information scent" is growing stronger, not weaker. On any given web page en route to the destination page, the information scent should be highest for the link label that will actually carry the user closer to accomplishing his/her goal.

Ideally, the web developer applies the AutoCWW during the initial development process. The AutoCWW fits in with the step-by-step procedures used by a typical web development team to design a complex site, starting with the home page and working successively down through lower and lower levels of a hierarchically organized website. The developer would use the AutoCWW to inspect and guide revisions of one web page at a time, the page currently under development. The developer's first task is to apply the AutoCWW to the home page by asking, Will the current iteration of the home page correctly route each and every user goal to the correct subsite main page? At each successively deeper level of the hierarchy, the AutoCWW repeats the same process. It determines if the link label system on each web page at that level will take each and every user goal input to its subset of user goals (the user goals routed to it from other web pages) and correctly route that goal onward to the correct web page.

A key advantage of the AutoCWW, however, is that it breaks down the overall usability inspection of the website into the manageable task of detecting and repairing the usability problems for just one web page at a time, the page currently under development. The AutoCWW test for each web page bears some similarity to testing an Internet router to determine if it accurately sends each packet on towards the correct destination, but the decision-making process is

very different. It is based on semantic similarity – the degree of "information scent" – of the user goal to the links presented on the web page, taking into consideration how the links are clustered into subregions on the page and whether the subregion containing the correct link has a higher scent than other subregions on the page.

When the AutoCWW predicts that the home page link label system will send one or more user goals to the wrong destination, the web developer can usually solve the problem by revising the link label system, and the AutoCWW helps by suggesting possible repairs. The process may continue through several iterations of the home page until the final iteration of the link label system routes all the user goals to the correct subsite main pages, as intended by the developer.

Selecting the correct link on a web page is the initial action evaluated during the walkthrough, but not the only action. When a user clicks a link and moves to the linked-to web page, s/he will evaluate the consequences of the previous action by deciding whether the action resulted in becoming closer to or farther from accomplishing the user's goal. The content and links on the linked-to web page must enable the user to accurately judge whether to press the back button, navigate elsewhere in the website, leave the site, or remain on the linked-to web page to continue working towards the user's goal. Therefore, except for the home page, the first task of the CWW is to examine each user goal routed to it from another web page (usually from a web page higher in the hierarchy) and determine the probability that the user will make the correct decision about staying on or leaving the newly visited page.

At each lower level of the hierarchy, the mean number of user goals routed to each web page diminishes. The home page allocates the initial large set of goals to various subsite main pages, so the CWW for each subsite main page uses only a fraction of the user goals, the ones routed to it from the home page. Thus, the CWW of the home page is the most time consuming, walkthroughs of subsite main pages are tested against fewer goals and take less time, and low-level content page may be tested against only one or a few user goals and take very little time. One can view the process as a flow of goals through a website. Starting at the home page, the subregion and link labels sort the complete set of goals by similarity to the user's goal. Each page directly linked to the home page will "receive" a subset of the goals that in turn will be sorted into sub subsets and "passed" on two second-level pages. Goals continue to "flow" until they "reach" the page that enables the user to accomplish his/her goal.

CWW Accurately Identifies Problems Encountered by College-Educated Users

The question is, *How well do CWW's predictions of problems or lack of problems stack up against user behavior?* Blackmon, Polson, Kitajima, & Lewis (2002) reported data from three experiments using laboratory data from college-educated experimental participants. The experimental task was simulated search of an online encyclopedia to find an article on a specified topic. On each trial, the participant was presented with a web page containing the target topic and a collection of category links. Clicking on a category link lead to a page with a list of target articles. We used CWW to divide the trials into four basic types: unfamiliar category link labels, confusable category link labels, goal specific problems, and no problems. The three experiments tested CWW predictions of users' success rates in accomplishing goals, verifying the value of CWW for identifying these usability problems. CWW accurately distinguished goals that would be easy for people to accomplish on a web page from goals that would be difficult.

In conclusion, we are well on the way towards building an AutoCWW that will be extremely useful to web developers.

References

Blackmon, M. H., Polson, P.G., Kitajima, M., & Lewis, C. (2002). Cognitive walkthrough for the web. *CHI 2002 Conference on Human Factors in Computing Systems.* ACM Press.

Kitajima, M., Blackmon, M. H., & Polson, P. G. (2000). A comprehension-based model of web navigation and its application to web usability analysis. In S. McDonald, Y. Waern & G. Cockton (Eds.), *People and Computers XIV - Usability or Else! (Proceedings of HCI 2000,* pp.357-373). Springer.