Methods and Techniques
63. Approaches and Methods
  63.1 Modelling and simulation
    63.1.1 modelling and human characteristics

Cognitive Walkthrough for the Web

Muneo Kitajima

Human-Computer Interaction Group
Institute for Human Science and Biomedical Engineering
National Institute of Advanced Industrial Science and Technology

Address: 1-1-1 Higashi Tsukuba Ibaraki 305-8566 Japan
Phone Number: +81 29 861 6650
Fax Number: +81 29 861 6012
E-mail: kitajima@ni.aist.go.jp

Manuscript submitted for International Encyclopaedia of Ergonomics and Human Factors 2nd Edition

1. **Introduction**

The Cognitive Walkthrough for the Web (CWW) is a theory-based usability inspection method for detecting and correcting design errors that interfere with finding information on a website (Blackmon, et al. 2002, 2003). CWW, like the original Cognitive Walkthrough (see **2.**), simulates step-by-step user behaviour for a given task and assumes that users perform goal-driven exploration. But CWW is specially tailored to simulate users navigating a website and better fits a realistic website design process, considering three features specific to website design. First, CWW uses realistic narrative descriptions of user goals that incorporate rich information about users' understanding of their tasks and underlying motivation. Second, CWW assumes that generating an action on a webpage (e.g., clicking a link, button, or other widget) is a two-step process. Step one is an attention process that parses a webpage into subregions and attends to the subregion of the page that is semantically most similar to the user goal. Step two is an action selection process that selects and acts on a widget from the attended-to subregion, the widget semantically most similar to the user goal. This two-step CWW web navigation mechanism is derived from a theory of the cognitive processes that control goal driven exploration, CoLiDeS (Kitajima et al., 2000). CoLiDeS, an acronym for Comprehension-based Linked model of Deliberate Search, extends a series of earlier models of performing by exploration based on Kintsch's construction-integration theory of text comprehension and problem solving processes (Kintsch, 1998). CoLiDeS is part of a broad consensus among theorists and website usability experts that problem solving processes, guided by users' goals and information scent, drive users' information-seeking or search behaviours when exploring a new website or carrying out a novel task on a familiar website (see **4.** for models of web navigation). Third, the CWW evaluation process can balance competing constraints by working on web pages in relation to a whole set of representative user goals. The CWW evaluation process can start with only a detailed description of the home page and a rough outline of its immediate successor pages, and can then be applied repeatedly to incrementally design and evaluate each successor page down through the hierarchy.

2. Cognitive Walkthrough Approach

CWW follows the Cognitive Walkthrough (CW) approach developed in the 1990’s (Polson et al., 1992; Wharton et al., 1994). CW is a usability evaluation approach that predicts how easy it will be for people to learn to do particular tasks on a computer-based system. It is crucial to design systems for ease of learning, because people generally learn to use new computer-based systems by exploration. People resort to reading manuals, using help systems, or taking formal training only when unsuccessful in learning to do their tasks by exploration. Published articles have reported applying CW to a wide variety of systems, including ATMs, telephone message and call forwarding systems, websites, computerized patient record systems for physicians to use, programming languages, multimedia authoring tools, and computer-supported cooperative work systems.

3. Cognitive Walkthrough for the Web

Similar to CW, CWW identifies usability problems by simulating step-by-step user behaviour for a given task using a prototype interface, and by having the design team answer the four essential questions of the original CW. These questions include: Q1) Will the users be trying to produce whatever effect the action has? Q2) Will users be able to notice that the correct action is available? Q3) Once users find the correct action at the interface, will they know that it is the right one for the effect they are trying to produce? Q4) After the action is taken, will users understand the feedback they get? However, CWW has transformed the CW approach by relying on Latent Semantic Analysis (LSA) developed by Landauer and Dumais

(1997) – instead of the subjective judgments of usability experts and software engineers – to predict whether users are likely to select the "correct action."

3.1 Objective Evaluation of Likelihood of Selection by Using LSA

LSA is a computer software system that objectively measures semantic similarity – similarity in meaning – between any two passages of text. LSA also assesses how familiar words and phrases are for particular user groups. LSA builds a semantic space representing a given user population's understanding of words, short texts (e.g., sentences, links), and whole texts. The meaning of a word, link, sentence or any text is represented as a vector in a high dimensional space, typically with about 300 dimensions. LSA generates the space by applying singular value decomposition, a mathematical procedure similar to factor analysis, to a huge terms-by-documents co-occurrence matrix. While analyzing the distinctive characteristics of the particular user group, CWW evaluators choose the LSA semantic space whose corpus of documents best represents the background knowledge of the particular user group – the space built from documents that these users are likely to have read. For example, CWW currently offers a college level space for French and five spaces that accurately represent general reading knowledge for English at college level and at third-, sixth-, ninth-, and twelfth-grade levels. So far CWW researchers have tested predictions and repairs only for users with college-level reading knowledge of English, but they expect to prove that CWW gives accurate predictions for other user groups and semantic spaces. LSA semantic spaces can be built to represent users who speak any language in the world at any level of background knowledge.

The degree of semantic relatedness or similarity between any pair of texts, such as the description of a user's goal and a link label on a webpage, is measured by the cosine value between the corresponding two vectors. Cosines are analogous to correlations. Each cosine value lies between +1 (identical) and -1 (opposite). Near-zero values represent two unrelated texts. CWW uses LSA to measure semantic similarity between a user’s information search goal (described in 100-200 words) and the text labels for each and every subregion of the web page and for each and every link appearing on a web page. CWW then ranks all the subregions and link labels in order of decreasing similarity to the user’s goal. CWW predicts success if the “correct action” is the highest-ranking link, if that link is nested within the highest-ranking subregion, and if the “correct action” link label and subregion avoid using words liable to be unfamiliar to members of the user group.

Another important measure provided by LSA is term vector length, a measure that is correlated with word frequency, and that estimates how much knowledge about a word or phrase is embedded in the designated LSA semantic space. A semantic space representing a given user population is generated from a large corpus of written materials (including books, magazines, and newspaper articles) read by typical members of that population. Words not included in the corpus are not represented in the semantic space. Words with low frequency in the corpus (e.g., specialized technical or scientific terms) have short term vector lengths. When a heading/link has a short term vector length, CWW predicts that users modelled by the semantic space will perceive it to be relatively meaningless, reducing the probability that users will attend to or click on them.

Relying on LSA produces the same objective answer every time, and laboratory experiments confirm that actual users almost always encounter serious problems whenever CWW predicts that users will have problems doing a particular task. Furthermore, using CWW to repair the problems produces two-to-one gains in user performance. So far CWW researchers have tested predictions and repairs only for users with college level reading knowledge of English, but they expect to prove that CWW gives comparably accurate predictions for other user groups and semantic spaces.

3.2 Coping with User Diversity

Relying on LSA makes it possible for CWW to do something that even usability experts trained in cognitive psychology can almost never do: objectively predict action selections for user groups whose background knowledge is very different from the background knowledge of the human evaluators. For example, selecting the sixth-grade semantic space enables LSA to “think” like a sixth grader, because the sixth-grade LSA semantic space contains only documents likely to have been read by people who have a sixth-grade education. In contrast, a college-educated analyst cannot forget the words, skills, and technical terms learned since sixth grade and cannot, therefore, think like a sixth grader.

Since the Cognitive Walkthrough was invented in 1990, the number and diversity of people using computers and the Internet have multiplied rapidly. Relying on LSA will enable the CW approach to keep pace with these changes. In cases where none of the existing LSA semantic spaces offers a close match with the background knowledge of the target user group, new semantic spaces can be constructed for CWW (and potentially for CW) analyses – in any language at any level of ability in that language. Specialized semantic spaces can also be created for bilingual and ethnic minority user groups and user groups with advanced background knowledge in a specific domain, such as the domain of medicine for evaluating systems used by health professionals.

4. Studies on User Behaviour on the Web

Studies on user behaviours on the Web are conducted from the theoretical perspective and the observational perspective. This section reviews some of the studies and describes how they related with the CWW approach.

4.1 Cognitive Models of User Search Behaviour

Information foraging theory (Pirolli and Card, 1999) incorporates measures of semantic similarity in a model of user search behaviour that is closely related to CoLiDeS's use of LSA. Both models take actions that are perceived as being close to a user's description of goal. However, information foraging theory has a much broader scope, attempting to characterize users' cost/benefit perceptions in making decisions, like terminating search of one website and searching for another site that contains more information relevant to their goals. However, the critical point is that the two frameworks are complementary, sharing a common model of the search process even though the common models are derived from very different cognitive architectures (ACT-R (Anderson and Lebiere, 1998) for the information foraging theory versus Kintsch's construction-integration theory of text comprehension (Kintsch, 1998) for the CoLiDeS model).

4.2 Approaches to Understanding User Behaviour on the Web

Many research groups are studying user behaviour on the Web, since insights are valuable for designing useful websites as well as efficient web servers. One approach focuses on the global behaviour of web users. Byrne et al. (1999) identified user interaction patterns by analyzing verbal protocols collected during browsing sessions. Tauscher and Greenberg (1997) addressed the same issues with usage data. Huberman et al. (1998) derived distributions of numbers of user clicks in a site by applying statistical analysis to user log data. Pitkow and Pirolli (1999) predicted web pages that users are likely to request by applying data mining technique to user log data. Chi et al. (2001) applied techniques used in information retrieval research to estimate the likelihood of selecting each link in a given website for given information needs. Except for Chi et al. (2001), the common characteristic of most of these studies was that user behaviour was aggregated over the different user goals. These studies used click stream data to uncover properties of typical sequences of page accesses. In most cases, investigators had no information about the content of users' goals.

Thus, these studies did not attempt to show how the content of users' goals controls navigation behaviour, despite widespread agreement that goals control search behaviour.

5. Conclusions

CWW helps designers detect and repair usability problems by simulating user's web navigation process. CWW is psychologically valid because it is based on well-verified cognitive architecture and model's predictions are confirmed to be consistent with what real human users would do. CWW uses LSA semantic spaces to represent users' background knowledge independently, and thereby achieves what designers cannot do unaided – accurately predict behaviour of users unlike themselves.

References

1. Anderson, J. R., & Lebiere, C. The Atomic Components of Thought. Lawrence Erlbaum Associates, 1998.
2. Blackmon, M. H., Kitajima, M., & Polson, P. G. (2003). Repairing usability problems identified by the cognitive walkthrough for the web. CHI 2003: Proceedings of the Conference on Human Factors in Computing Systems (pp. 497–504). New York: ACM. CHI Letters, 5(1), 497–504.
3. Blackmon, M. H., Polson, P. G., Kitajima, M., & Lewis, C. (2002). Cognitive walkthrough for the Web. CHI 2002: Proceedings of the Conference on Human Factors in Computing Systems (pp. 463–470). New York: ACM. CHI Letters, 4(1), 463–470.
4. Byrne, M. D., John, B. E., Wehrle, N. S., & Crow, D. C. The tangled Web we wove: A taskonomy of WWW use. In Proceedings of CH'99, ACM Press, 544-551, 1999.
5. Chi, E., Pirolli, P., Chen, K., & Pitkow, J. Using information scent to model user information needs and actions and the Web. In Proceedings of CHI 2000, ACM Press,

490-497, 2001.

6. Huberman, B. A., Pirolli, P., Pitkow, J., & Lukose, R. Strong regularities in World Wide Web surfing. Science, 280, 95-97, 1998.
7. Kintsch, W. Comprehension: A Paradigm for Cognition, Cambridge University Press, 1998.
8. Kitajima, M., Blackmon, M. H., & Polson, P. G. (2000) A Comprehension-based model of Web navigation and its application to Web usability analysis. In People and Computers XIV, Springer, 357-373.
9. Landauer, T. K. & Dumais, S. T. A solution to Plato's problem: The Latent Semantic Analysis theory of acquisition, induction, and representation of knowledge. Psychological Review, 104, 21-240, 1997.
10. Pirolli, P. & Card, S. Information foraging. Psychological Review, 106, 643-675, 1999.
11. Pitkow, J. & Pirolli, P. Mining longest repeated subsequences to predict World Wide Web surfing. In Proceedings of the USENIX Conference on Internet, 1999.
12. Polson, P., Lewis, C., Rieman, J., & Wharton, C. (1992). Cognitive walkthroughs: A method for theory-based evaluation of user interfaces. International Journal of Man-Machine Studies, 36, 741–773.
13. Tauscher, L. & Greenberg, S. How people revisit web pages: Empirical findings and implication for the design of history systems. International Journal of Human-Computer Studies, 47, 97-137, 1997.
14. Wharton, C., Rieman, J., Lewis, C., & Polson, P. (1994). The cognitive walkthrough method: A practitioner's guide. In J. Nielsen & R. L. Mack (Eds.), Usability inspection methods (pp. 105–140). New York: John Wiley & Sons.