Position Paper for CHI 2002 Workshop: Automatic Capture, Representation and Analysis of User Behavior

Anthony Hornof
Department of Computer and Information Science
University of Oregon
<hornof@acm.org>
January 29, 2002

This is a position paper stating my interest and qualification for attending the CHI 2002 Workshop on Automatic Capture, Representation, and Analysis of User Behavior. The workshop will focus on (1) the development of computerized tools for recording, studying and visualizing new forms and quantities of empirical user data, and (2) how these tools and data will contribute to theories and principles for building easier-to-use and easier-to-learn computer interfaces.

Relevant experience

Relevant experience includes having empirically validated detailed research hypotheses about visual search of computer menus (Hornof, 1999; Hornof & Kieras, 1997; Hornof & Kieras, 1999) and layouts (Hornof, in preparation; Hornof, in press) by building cognitive models using the EPIC architecture (Kieras & Meyer, 1997). Research hypotheses have included that people search unfamiliar menus looking for a known target with a somewhat random and somewhat systematic search strategy, considering more than one item at a time. Previous theories proposed either a purely systematic or a purely random search, and that people consider one item at a time. Recent modeling confirms these findings for larger visual layout and also demonstrates that, when a useful visual hierarchy is in place, as when items are in groups with useful group headings, people will conduct a highly efficient search first of the headings and then within the group containing the target. I have collected eye movement data to evaluate these theories and am currently in the process of developing automated tools to represent and analyze this eye movement data with respect to the specific theories of the models.

Issues I have encountered and would like to address

My major research interests relate to how highly visual interfaces can better support human patterns of creativity. For example, how can data visualization tools better support complex decision and problem-solving tasks?

Research problems I would like to solve include:

1. How can eye tracking best serve the field of HCI? The answer will most certainly involve automated capture, representation, and analysis of visual behavior. Numerous researchers are searching for ways that eye tracking can serve HCI, but outstanding opportunities have yet to emerge (Goldberg, 2000; Goldberg & Kotval, 1998). The web foraging work at XeroxPARC seems particularly promising, largely part because it is built on a theory of web usage that can be evaluated and is useful outside of the context of the actual usability analysis (Card et al., 2001).

2. How can eye movement data to validate detailed empirical models? I am currently working with a Ph.D. student to analyze eye movement data that has been collected for a visual search task. A detailed EPIC model that predicts patterns of eye movements has already been constructed to account for the reaction time data (Hornof, in preparation). The eye movement data will be analyzed to evaluate the accuracy of the model. It is challenging to distill the eye movement data down to the cut of the data necessary to validate or rule out some aspect of the strategies.

Automated analysis techniques will be critical. The automated eye movement protocol analyses of Salvucci and Anderson (Salvucci & Anderson, 2001) will help to guide our process, but clearly we will need to further develop the techniques.

3. How can eye movement data contribute to theories about user interaction? Gray and Fu's (2001) is an excellent example of how eye tracking can help build theories about user behavior for a variety of subtly-different task parameters, in this case how likely a user is to commit task knowledge to memory as a function of the ease of accessing the same data during task execution.

But I'd like to move forward to test specific theories about visualization tools such as those coming out of UM College Park, XeroxPARC, and elsewhere (Card, Mackinlay & Shneiderman, 1999). To test theories about how visualizations work, the theory must be stated in sufficient detail such that one pattern of eye movements would support the theory, and another would negate the theory. It would be very interesting to validate the Reference Model for Visualization proposed by Card et al. (1999), but it is not clear whether the model is appropriate for validation with eye movement data.

For each theory to be tested, important questions include: (1) In how much detail must the theory be stated in order to test it with eye movements and (2) How can the theory-testing deal with the inherently noisy nature of eye movement data?

4. What level of detail in a theory is necessary for it to be evaluated with automatically captured and analyzed data? What kind of automated analysis is necessary to validate theories? This is partly but not entirely a chicken-and-the-egg problem. You can only test a theory that is sufficiently well specified such that some outcomes of an experiment would support the theory and all other outcomes would negate the theory. Similarly, an experiment must be carefully designed in order to carefully test a theory. An extended example will illuminate this challenge.

An extended example demonstrating the need for automated capture, representation, and analysis of user behavior

Cognitive models were built to investigate how people search labeled visual hierarchies--in which every group of items in a visual layout has a heading that indicates if the target item is in that group (Hornof, in preparation; Hornof, in press). A two-tiered search strategy was proposed because, given the constraints associated with human visual search processes as captured by the EPIC cognitive architecture (Kieras & Meyer, 1997), the only way the participants could have found the targets as quickly as they did was to conduct a two-tiered search--first searching group labels, and then searching within the target group. A most cursory review eye movement data collected for this task validates this theory--in nearly all cases, participants first moved their eyes among the group labels, and then searched within the target group (unpublished data from work in progress). Figure 1 shows a visualization of eye fixations from the experiment. Numerous similar traces demonstrate that participants conducted a two-tiered search: They first used the group labels to find the target group, and then searched for the target item exclusively within that group.

To arrive at the view of the data that allows us to conclude that people conducted a two-tiered search, substantial automatic capture, representation, and analysis of user behavior was required. First, the eye tracker and computer sampled the gaze location every 17 msec. Second, the gaze data was automatically converted into fixation position and duration data. Third, the fixations and durations were superimposed on the viewed scene in order to reveal the general trends and systematic errors in the data.

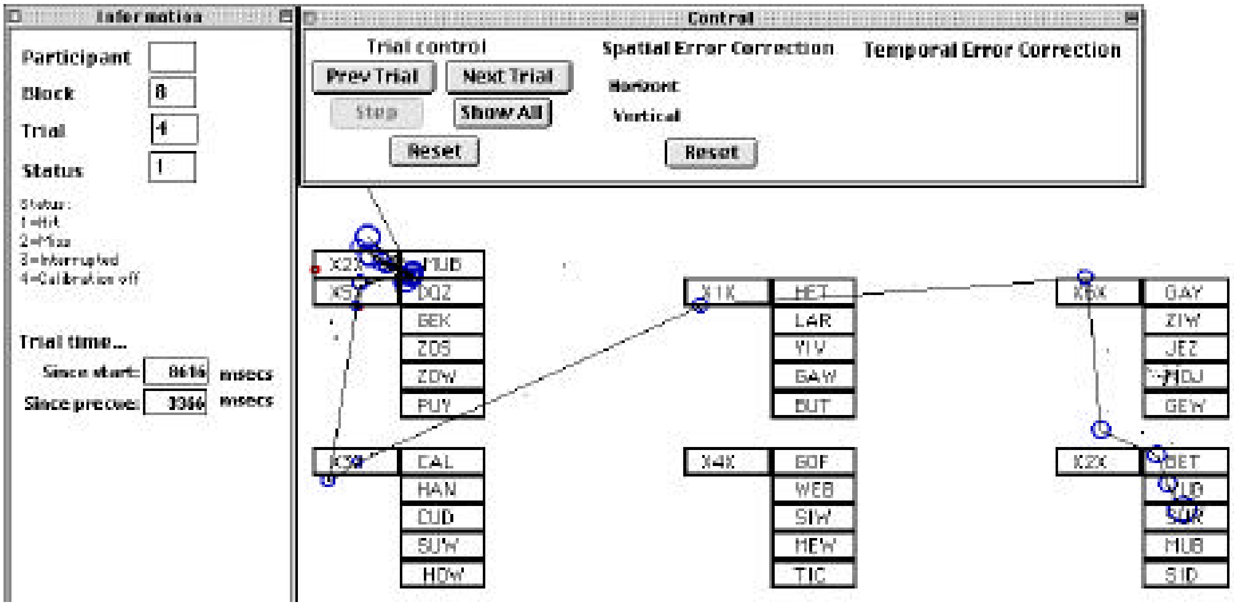

Figure 1. Screenshot from VizFix, the eye fixation visualization and error correction tool currently being developed by Hornof and Halverson, showing a visualization of eye fixations from a visual search experiment. The lines connect consecutive fixations, with the diameter of each circle indicating the duration of the fixation. In the experiment, the participant studied the precue (X2X MUB, in the top left corner), clicked the mouse on the precue to make the layout appear (though both the precue and layout are shown here together), and found and clicked on the target item as quickly as possible. In this trial, the target is the item MUB in the group labeled X2X, at the bottom left of the layout. The trace of the eye movement data shown here demonstrates that, when group labels (X1X, X2X, etc.) were shown, participants used them to first find the target group, and then searched within that group.

These three automatic capture and representation steps were sufficient to reveal the two-tiered search in the data for labeled hierarchies. However, two additional steps are required to measure and automatically analyze and categorize the eye movement patterns for a single trial as following a systematic two-tiered search. We'll continue the enumeration: Fifth, systematic errors in the data must be automatically identified and corrected. Note in Figure 1 that the fixations at the bottom right appear to be roughly one item above the items that were probably fixated (such as the *X2X* and *MUB*). This is probably due to systematic error in the eye tracking data. The data should be thrown out, systematically adjusted, or considered in terms of relative eye movements and not absolute position on the layout.

Sixth, the fixation data must be systematically categorized so that the theory can be tested. To test for a two-tiered search strategy, the automated categorization should be relatively straightforward: Did the eyes first look at group labels and then within the target group? But there are roughly 2,000 trials for which this question needs to be answered. Automated visual search strategy identification tools are needed. Salvucci and Anderson's (2001) automated protocol analysis may be overwhelmed by the many different eye movement patterns that would correspond to a two-tiered search, and may need to be extended.

The theory and the data interact. Data must be collected and analyzed to test the theory, and the theory must lend itself to evaluation with the data. Byrne (2001) and Anderson et al. (1998)

proposed feature-based menu searches, in which the eyes moved to menu items that had a similar shape as the target. The models were evaluated with respect to how well they predicted reaction time data and general patterns of eye movements (Byrne, 2001; Byrne, Anderson, Douglass & Matessa, 1999), but not with respect to the central assumption of the model that the eyes tended to move to items that shared the same shape as the target. Such an analysis would be quite involved, but would ultimately be the best way to test the theory. Each item in the layout would need to be categorized with respect to its similarity to the target, and every fixation would need to be categorized with respect to whether it was to an item that shared features with the target. Existing automated eye movement protocol analysis will be useful only to the extent that a limited number of well-defined eye movement patterns can be identified in advance. It is likely that new automated eye movement protocol analyses will need to be developed.

The five steps involved in the automated capture, representation, and analysis of eye movement behavior can be summarized as thus: (1) Gaze sampling, (2) Fixation determination, (3) Fixation visualization. (4) Fixation error correction. (5) Data categorization or coding that can lead to a direct evaluation of the theory.

Byrne et al.(1999), Byrne (2001), and Gray (2001) are landmark papers in that they use eye movement data to evaluate of two competing models of the visual search of computer menus pave the way for future research in the automated capture, representation, and analysis of user eye movements.

Suggestions for the type of paper they would like to author or co-author.

- An analysis of model-based evaluation tools. A comparison and contrast of existing predictive tools, perhaps with an emphasis on visual layout analysis tools, including but not limited to Faraday's web page critiquing tool (Faraday, 2000), Apex (Freed & Remington, 2000). Special emphasis will be placed on plotting out how each tool has been or could be evaluated with automated capture, representation, and analysis of user behavior. It is anticipated that some tools will be able to be evaluated in greater detail than others, due to varying degrees of fidelity and transparency of the model within the tool. If time and resources are available, perhaps the tools could be acquired from their authors and evaluated, such as with an eye movement study.

- A theoretical discussion of the sort of theories that can be tested with user data, and those which cannot. This would relate to the extended example, above

- Any paper relating to the four research issues discussed above, namely:
    1. How can eye tracking best serve the field of HCI?
    2. How can eye movement data to validate detailed empirical models?
    3. How can eye movement data contribute to theories about user interaction?
    4. What level of detail in a theory is necessary for it to be evaluated with automatically captured and analyzed data?

- How can automated capture, representation, and analysis of user data be used to evaluate cognitive architectures (such as EPIC and ACT-R/PM) and to evaluate models built in those architectures? How can the data be used to evaluate predictive analysis tools based on the GOMS methodology, such as GLEAN (Kieras, Wood, Abotel & Hornof, 1995) and Apex (Freed & Remington, 2000)?

Bibliography

Anderson, J. R., Matessa, M., & Lebiere, C. (1998). The visual interface. In J. R. Anderson & C. Lebiere (Eds.), *The Atomic Components of Thought*. Mahwah, NJ: Lawrence Erlbaum, 143-168.

Byrne, M. D. (2001). ACT-R/PM and menu selection: Applying a cognitive architecture to HCI. *International Journal of Human-Computer Studies*, 55, 41-84.

Byrne, M. D., Anderson, J. R., Douglass, S., & Matessa, M. (1999). Eye tracking the visual search of click-down menus. *Proceedings of CHI 99*, New York: ACM, 402-409.

Card, S. K., Mackinlay, J. D., & Shneiderman, B. (1999). *Readings in Information Visualization: Using Vision to Think*. San Francisco, CA: Morgan Kaufmann.

Card, S. K., Pirolli, P., Wege, M. V. D., Morrison, J. B., Reeder, R. W., Schraedley, P. K., & Boshart, J. (2001). Information Scent as a Driver of Web Behavior Graphs: Results of a Protocol Analysis Method for Web Usability. *Proceedings of CHI 2001*, 498-505.

Faraday, P. (2000). Visually Critiquing Web Pages. *6th Conference on Human Factors & the Web*, available at http://www.tri.sbc.com/hfweb/faraday/faraday.htm.

Freed, M. A., & Remington, R. W. (2000). Making Human-Machine System Simulation a Practical Engineering Tool: An Apex Overview. *Proceedings of the Third International Conference on Cognitive Modeling*.

Goldberg, J. H. (2000). Eye movement-based interface evaluation: What can and cannot be assessed?, Santa Monica, CA: Human Factors Society.

Goldberg, J. H., & Kotval, X. P. (1998). Computer Interface Evaluation Using Eye Movements: Methods and Constructs. *International Journal of Industrial Ergonomics*, 24, 631-645.

Gray, W. D., & Fu, W.-T. (2001). Ignoring perfect knowledge in-the-world for imperfect knowledge in-the-head: Implications of rational analysis for interface design. *Proceedings of ACM CHI 2001: Conference on Human Factors in Computing Systems*, New York: ACM.

Hornof, A. (in preparation). Think global, act local: The strategic coordination of perceptual and motor processes for scanning hierarchical computer screen layouts.

Hornof, A. J. (1999). *Computational models of the perceptual, cognitive, and motor processes involved in the visual search of pull-down menus and computer screens*. Doctoral dissertation in Computer Science and Engineering. The University of Michigan, Ann Arbor, Michigan.

Hornof, A. J. (in press). Visual search and mouse pointing in labeled versus unlabeled two-dimensional visual hierarchies. *ACM Transactions on Computer-Human Interaction*, available at ftp://ftp.cs.uoregon.edu/pub/hornof/Labeled_Hierarchies_Paper.pdf.

Hornof, A. J., & Kieras, D. E. (1997). Cognitive modeling reveals menu search is both random and systematic. *Proceedings of ACM CHI 97: Conference on Human Factors in Computing Systems*, New York: ACM, 107-114.

Hornof, A. J., & Kieras, D. E. (1999). Cognitive modeling demonstrates how people use anticipated location knowledge of menu items. *Proceedings of ACM CHI 99: Conference on Human Factors in Computing Systems*, New York: ACM, 410-417.

Kieras, D. E., & Meyer, D. E. (1997). An overview of the EPIC architecture for cognition and performance with application to human-computer interaction. *Human-Computer Interaction*, 12(4), 391-438.

Kieras, D. E., Wood, S. D., Abotel, K., & Hornof, A. (1995). GLEAN: A computer-based tool for rapid GOMS model usability evaluation of user interface designs. *Proceedings of the ACM Symposium on User Interface Software and Technology, UIST '95*, New York: ACM, 91-100.

Salvucci, D. D., & Anderson, J. R. (2001). Automated eye-movement protocol analysis. *Human-Computer Interaction*, 16, 39-86.