ONR Attention, Perception and Data Visualization Workshop
George Mason University, 22-24 May, 2002

Cognitive modeling, visual search, and eye tracking

Anthony J. Hornof
Department of Computer and Information Science
University of Oregon
<hornof@acm.org>

Cognitive modeling and eye tracking will improve the usefulness and cognitive assistance provided by data visualizations and complex data displays. Initially, basic perceptual and search strategy parameters must be established through experimentation that examines, isolates, and extracts fundamental human performance characteristics and clarifies the details of visual perception, visual search strategies, and ocular-motor behavior. All of these processes and components interact when a person uses a complex visual display. Simpler, smaller experiments identify the parameters, capabilities, and limitations associated with these fundamental processes, which are incorporated into cognitive architectures and explained with models. The results of this research contributes directly to models of more complex tasks and displays.

Past cognitive modeling research has demonstrated that (1) current cognitive architectures are well-suited for modeling visual search tasks, (2) visual search is both random and systematic, (3) people do not always fixate every item--they sometimes examine more than one item at a time, (4) people use anticipatory location information to guide visual search, and (5) visual features sometimes guide the search.

Current research is underway to provide increasingly accurate explanatory cognitive models of visual search tasks, and to move closer to *a priori* predictive cognitive modeling of visual tasks. Current research underway by the author examines (1) *how many* visual objects can be examined with a single ocular fixation, and how this should be incorporated into a predictive model, (2) how existing models can be refined to better account for new eye movement data, and (3) how to categorize and classify eye tracking data to identify the cognitive strategies that likely produced it.

The promise for cognitive modeling in Human-Computer Interaction

Cognitive models are computer programs that behave in some way like humans. In the context of this discussion, and in most cognitive modeling, the models emulate the perceptual, cognitive, and motor processes that a person would use to accomplish a specified piece of work--a *task*--and take the same amount of time that a human would take. Cognitive modeling is useful to the design and analysis of user interfaces because the modeling reveals patterns of behavior at a level of detail not otherwise available to analysts and designers (Gray, John & Atwood, 1993). The ultimate promise for cognitive modeling in design and analysis of user interfaces is that the modeling provides the science base needed for predictive analysis tools and methodologies (Card, Moran & Newell, 1983).

There are two phases of cognitive modeling: (1) *exploratory* and (2) *predictive*. In the exploratory mode, models are constructed to explain empirical data that has already been collected and analyzed. In the predictive (or *confirmatory*) mode, models are constructed to make *a priori* predictions of user performance; that is, predictions before human data has been collected. In both phases, the output from the model is referred to as a “prediction” even though in exploratory modeling it is really a *post*diction.

Recent exploratory modeling has predicted aspects of menu search (Anderson, Matessa & Lebiere, 1998; Byrne, 2001; Hornof & Kieras, 1997; Hornof & Kieras, 1999), icon search (Fleetwood & Byrne, 2001), and how design decisions affect micro-strategies (Gray & Boehm-Davis, 2000). Predictive modeling has informed the design of telephone operator workstations (Gray et al., 1993), commercial workplace ergonomic analysis software (Gong & Kieras, 1994), multi-modal combat information center watchstations (Santoro, Kieras & Campbell, 2000), and watchstations used simultaneously by teams of workers (Pharmer, Freeman, Scott-Nash, Santoro & Kieras, 2001). Many of these predictive models were validated with human data, confirming the accuracy and usefulness of the models and the methodology. Sometimes predictive models are refined to explain the observed data more accurately, in which case the models are re-used in an explanatory mode.

Once a predictive modeling methodology has become sufficiently well-established and well-defined, it can be used in a purely predictive *a priori* mode to guide a design process early in the development life cycle. User observation studies are still necessary before final system deployment. The GOMS methodology (John & Kieras, 1996; Kieras, 1997) is sufficiently refined and specified to be useful in a purely *a priori* mode. GOMS-based predictive analysis tools such as GLEAN (Kieras, Wood, Abotel & Hornof, 1995) and Apex (Freed & Remington, 2000) facilitate the routine use of cognitive modeling for predicting HCI performance.

Exploratory modeling must precede predictive modeling. Exploratory modeling establishes the invariant components of human information processing that will constrain the behavior of the predictive model. These constraints are represented by a modeling language such as GOMS or by a cognitive architecture such as EPIC (Kieras & Meyer, 1997) or ACT-R/PM (Anderson & Lebiere, 1998; Byrne, 2001), along with a set of conventions for applying the modeling language or architecture to simulate a task execution. The modeling language or architecture represents the invariant aspects of human behavior, analogous to the computer hardware and the hardware performance constraints which do not change regardless of the task, and constrain the types of models that can be built.

The Three Goals of Exploratory Cognitive Modeling

The three goals of exploratory cognitive modeling are as follows:

- #1 Provide new explanations of human behavior and performance characteristics. The explanations must be human-interpretable, and meaningful to designers and analysts outside of the context of cognitive modeling.
- #2 Identify and articulate principles, guidance, and wisdom for building predictive models, including details of the cognitive strategies and perceptual and motor parameters appropriate *and not appropriate* for classes of tasks and interfaces.
- #3 Recommend refinements and extensions to the architectures and modeling conventions so that cognitive modeling can more accurately represents the invariant, fundamental human capabilities and limitations, and become better-suited for making accurate predictions.

Eye Tracking and Cognitive Modeling

Eye tracking will play an important role in understanding how people interact with complex displays. Eye tracking data will be of relatively little use, however, if studied and analyzed in the absence of a cognitive theory or model that discusses and explains *why* a person would look here or there at a particular point in time in the context of pursuing a specific task. Eye tracking and cognitive modeling of visual tasks will become increasingly important to each other. Eye tracking data is needed to evaluate and refine cognitive models of visual tasks, and cognitive models are needed to make sense of eye tracking data.

Figure 1 (left frame) shows one of the layouts used in a recent visual search experiment (Hornof, in preparation; Hornof, 2001), annotated with a few details such as the numbers assigned to the thirty layout positions for purposes of analysis. In this experiment, participants were precued with a word (or pseudo-word) and then asked to find the word in a layout such as that shown in Figure 1. The factors that were varied included the number of items that appeared in the layout (10, 20, or 30) and whether the layout included group labels to guide the participant to the target. Labels were *X1X*, *X2X*, etc. The experiment was designed to separate visual search time from target selection time.

A handful of plausible strategies and models were examined to account for the data and the perceptual, cognitive, and motor processing involved in accomplishing the task. Task analysis and the observed data suggest that participants used the group labels to find the target. Participants likely conducted a two-tiered search, in which they first searched the group labels first, and then searched within the target group. Participants could have used a number of different two-tiered search strategies, several of which were written and run using EPIC.

The best-fitting two-tiered search strategy is the *almost-perfect hierarchical search strategy* model, which assumes a mostly-systematic (but occasionally random) search of layout labels until the target group is found, and then a systematic search within the target group. The model accounts for all of the features in the observed data and explains the observed data with an AAE of 3.6%. This good fit suggests that participants (1) almost always examined the group labels first, (2) conducted a highly efficient almost-perfect streamlined search, as was assumed by the model in order to fit the data, and (3) usually searched the group labels systematically in the order assumed here, but sometimes used a slightly different order.

Figure 1 (right frame) shows the pattern of eye movements predicted by the best-fitting model for labeled layouts, the *almost-perfect hierarchical search strategy* model. These predictions were generated before eye movement data was collected.

A cursory review eye movement data collected for this task appears to validate the assumptions of the *almost-perfect hierarchical search strategy* model, that people conduct a highly streamlined two-tiered search, first of the labels and then in

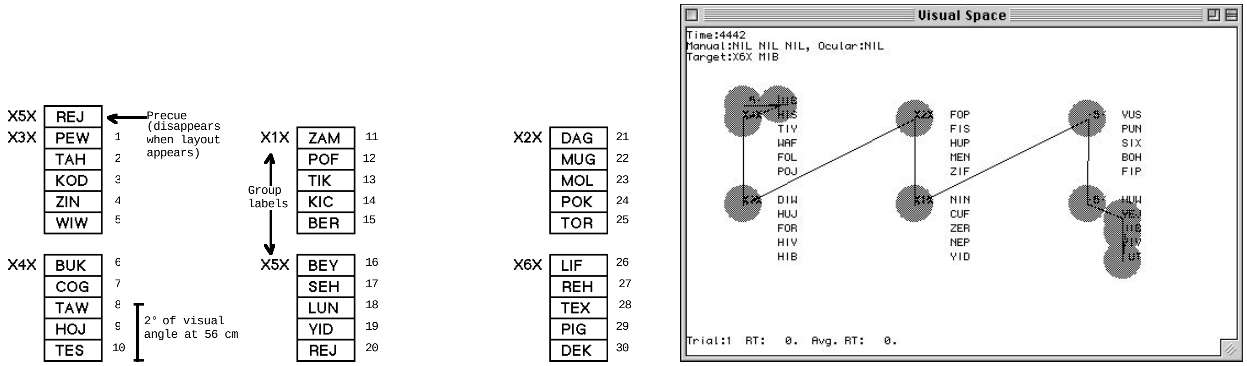

Figure 1. LEFT FRAME: A sample layout used in the visual search experiment. This layout has six labeled groups, drawn to scale and annotated with the number of each target position. The precue is X5X REJ. *The target is in position 20, in the group labeled* X5X. *RIGHT FRAME: Eye movements predicted by the almost-perfect hierarchical search strategy running in EPIC. The lines represent the order of fixations, starting from the top left, and the circles indicate the area of the display that would be visible in the foveal region of the eye as represented by EPIC. The target is in position 28, near the bottom right. The model studies the precue at the top left, clicks the mouse on the precue to make the layout appear, searches the layout for the target, and clicks on the target. The eye movement data shows that the model first searches the group labels and then searches within the target group. These eye movements were predicted prior to the collection of any eye movement data for this task, and were predicted by exploring search strategies that could predict the search time data, given the perceptual, cognitive, and motor processing constraints imposed by the EPIC architecture.*

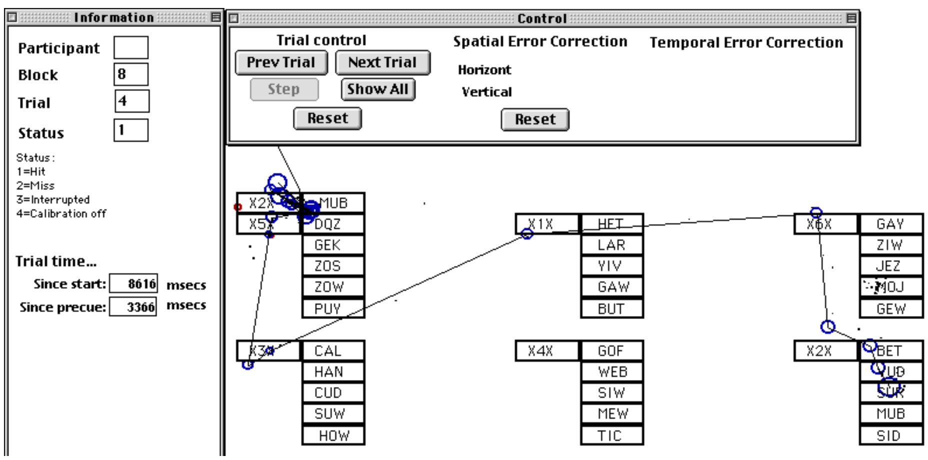

Figure 2. Eye movement data from the visual search experiment conducted by the author. The lines connect consecutive eye fixations, and the diameter of each circle indicates the duration of each fixation. The participant studied the precue (X2X MUB), clicked the mouse on the precue to make the layout appear, searched the layout for the target, and clicked on the target. The eye movement data shows that the participant first searched the group labels (X1X, X2X, etc.) and then searched within the target group, very much as is predicted by the model, as can be seen by comparing this figure to Figure 1 (right frame). This is some of the data that will be analyzed in the proposed work, and used to evaluate the accuracy of the model. Search of the unlabeled layouts generally include many more fixations, and is not as directly predicted by the model.

the target group (unpublished data from work in progress). Figure 2 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.

Though the data appear the support the assumptions of the model, a much more rigorous data analysis is in order to make quantitative claims about the fit betweent the eye movement data and the model. Additional steps are needed to measure and automatically analyze and categorize eye movement patterns from every trial, and determine to what extent each demonstrates a systematic two-tiered search. As well, systematic errors in the data must be automatically identified and corrected. Note in Figure 2 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. Much work needs to be done to make use of this data.

Exploratory cognitive models were built to explain reaction time data. The models predict the patterns eye movements participants likely exhibited. The eye tracking data was also collected by the author and will now be used to evaluate the accuracy of the models, to refine the models, and to provide new insights.

References

Anderson, J. R., & Lebiere, C. (Eds.). (1998). *The Atomic Components of Thought*. Mahwah, NJ: Erlbaum.

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.

Card, S. K., Moran, T. P., & Newell, A. (1983). *The Psychology of Human-Computer Interaction*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Fleetwood, M. D., & Byrne, M. D. (2001). Modeling Icon Search in ACT-R/PM. *Fourth International Conference on Cognitive Modeling*.

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*.

Gong, R., & Kieras, D. E. (1994). A validation of the GOMS model methodology in the development of a specialized, commercial software application. *Proceedings of ACM CHI 94: Conference Companion on Human Factors in Computing Systems*, New York: ACM.

Gray, W. D., & Boehm-Davis, D. A. (2000). Milliseconds matter: An introduction to microstrategies and to their use in describing and predicting interactive behavior. *Journal of Experimental Psychology: Applied*, 6(4), 322-335.

Gray, W. D., John, B. E., & Atwood, M. E. (1993). Project Ernestine: Validating a GOMS analysis for predicting and explaining real-world task performance. *Human-Computer Interaction,* 8, 237-309.

Hornof, A. (in preparation). Exploratory cognitive modeling provides insight into the strategic coordination of the perceptual and motor processes used to scan 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. (2001). Visual search and mouse pointing in labeled versus unlabeled two-dimensional visual hierarchies. *ACM Transactions on Computer-Human Interaction,* 8(3), 171-197.

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.

John, B. E., & Kieras, D. E. (1996). The GOMS family of user interface analysis techniques: Comparison and contrast. *ACM Transactions on Computer-Human Interaction*, 3(4), 320-351.

Kieras, D. E. (1997). A Guide to GOMS model usability evaluation using NGOMSL. In M. Helander, T. K. Landauer, & P. Prabhu (Eds.), *Handbook of Human-Computer Interaction.* (2nd ed.). Amsterdam: North-Holland, 733-766.

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.

Pharmer, J. A., Freeman, J. T., Scott-Nash, S., Santoro, T. P., & Kieras, D. (2001). Complementary methods of modeling team performance. *Proceedings of the 44th Annual Meeting of the Human Factors and Ergonomics Society,* I/ITSEC.

Santoro, T. P., Kieras, D. E., & Campbell, G. E. (2000). GOMS Modeling Application to Watchstation Design Using the GLEAN Tool. *Interservice/Industry Training, Simulation and Education Conference*.