A PROCESS MODEL OF DISPLAY-BASED HUMAN-COMPUTER INTERACTION

Muneo Kitajima

Industrial Products Research Institute
1-1-4 Higashi
Tsukuba Ibaraki 305, Japan
+81 298 54 6731
i8001@ipri.go.jp

Peter G. Polson

University of Colorado
Institute of Cognitive Science
Boulder, Colorado 80309-0345
+1 (303)492-5622
pploson@clipr.colorado.edu

DISPLAY-BASED HCI

Recent empirical studies on display-based human-computer interaction have provided evidence against standard plan-based theories (*e.g.* Card, Moran, and Newell, 1983; Kieras and Polson, 1985) of expertise in HCI. Mayes, Draper, McGregor and Oatley (1988) report that experienced MacWrite users have poor recall for the names of menu-items. In addition, Payne (1991) has shown that experienced users do not have complete knowledge about the effects of commands.

These results provide support for theoretical frameworks that assume that sequences of user actions are not pre-planned. Each action is determined making use of display feedback during the course of generating a sequence of action necessary to complete a task. The display plays a crucial role in successful and smooth interaction; the interaction is *truly* mediated by the display. Howes and Payne (1990), and others have developed theories of skilled performance in which successful interactions are mediated by representations of intermediate states of a task presented in a display. Larkin (1989) called her framework display-based problem solving.

Howes and Payne (1990) extended the task action grammar framework to display-based, menu systems. D-TAG (display-based task-action grammar) is a competence model of users' knowledge of display-based systems used for evaluating the consistency of an interface.

In this paper we outline a process model for display-based HCI based on Kintsch's (1988) construction - integration theory. In a companion paper (Kitajima and Polson, 1992) we describe an early version of the model and the results of three series of simulation experiments. This paper presents an overview of our developing theoretical framework. Our long term goal is to account for both routine, skilled use of a software tool and learning to use a new tool by exploration.

The CI model was originally proposed by Kintsch (1988) as a model of text comprehension. Text comprehension is modeled as a two-phase process of construction and integration. The construction process generates associative networks by using input text as the cues to retrieve relevant knowledge from long-term memory. The construction process is bottom-up; the resultant network includes both contextually relevant knowledge and associatively related knowledge inconsistent with the context. The integration process uses a spreading activation mechanism to extract a contextually relevant consistent interpretation of the text in the form of an activation pattern over the network. The CI model has been generalized to HCI by using the comprehension processes to select an action sequence with the aid of rich knowledge about commands (Mannes and Kintsch, 1991; Doane, Kintsch and Polson, 1990).

A PROCESS MODEL OF DISPLAY-BASED HCI

Kitajima and Polson (1992) have developed a model of display-based HCI based on the CI model. This model simulated the action selection processes when it was given the correct sequence of goals, expectations, and objects to be attended to in order to successfully perform a task.

Figure 1 is an overview of an extension of the Kitajima and Polson (1992) model. The new model improves the old in two respects. First, expectations, which were provided manually in the previous model, are now generated by the model itself. Second, binding of variables in action descriptions are also developed automatically rather than being provided manually.

As the Figure shows, this new model assumes that three cycles of the basic construction-integration process are involved in action selection. Knowledge used in these cycles, not explicitly presented in the figure, are categorized into five classes: *Goals* are equivalent to task descriptions, *expectations* are knowledge about expected display states, *display representations* are the result of parsing current display contents, *domain knowledge* is the general knowledge about task domains, interface objects, and linking knowledge between them, and *action elements* are knowledge about primitive actions at the level of pointing, clicking, typing, and so on.

In the first cycle, expectations stored in long-term memory are retrieved and the most appropriate one for the specific task situation is selected. Here, goals and display representations serve as memory retrieval cues.

The second cycle, the attention formation cycle, determines which of the many objects presented on the display are candidates for being operated on by user actions. The candidate objects are used to change generic action elements into actions that operate on specific attended-to objects; this is done by binding variables in the action elements to candidate objects.

The third cycle, action selection, first combines the bound action elements with the existing network, then links and integrates the network. When activation equilibrates the most highly active action element whose preconditions are satisfied is selected for execution. Executing this action changes the display and a new cycle starts again with expectation formation, continuing until the task is completed.

GIVEN: Task (goals) and Display
PROCESSES:

- DISPLAY REPRESENTATION
- EXPECTATION FORMATION CYCLE
  - Retrieve expectations (cued by goals and display representations)
  - Make links among goals, display representations, and expectations
  - Integrate the network
  - Select the most active expectation
- ATTENTION FORMATION CYCLE
  - Retrieve domain knowledge (cued by goals, expectation, and display)
  - Generate candidate objects
  - Make links among goal, display, expectation, domain knowledge, and candidate objects
  - Integrate the network
  - Select most active objects for variable bindings in action elements
- ACTION SELECTION CYCLE
  - Bind variables in action elements to selected objects
  - Make links among goals, display, expectations, domain knowledge, and action elements
  - Integrate the network
  - Select the most active eligible action

Fig.1 A process model of display-based HCI based on CI model.

CONCLUSION

We have developed a model of display-based HCI in which information contained in the display about the intermediate states of a task plays a critical role in the generating of correct action sequences. In a companion proposal (Polson and Kitajima, 1992), we show how this model accounts for errors made by skilled users. Our model is also consistent with results published by several investigators that call into question the basic assumptions of plan-based theories of skill using a production system framework. We have done simulation experiments validating the assumptions of an earlier version of the model simulating the action selection process. The strength of this model is that it is based on an cognitive architecture, which enables us to understand performance data, such as errors, learning, execution times, forgetting, and the like.

REFERENCES

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

Doane, S. M., Kintsch, W., and Polson, P. G. Modeling UNIX command production: What experts must know. ICS Technical Report #90-1. Institute of Cognitive Science, University of Colorado, Boulder CO., 1990.

Howes, A., and Payne, S. J. Display-based competence: towards user models for menu-driven interfaces. *Int. J. of Man-Machine Studies*, 33 (1990), 637-655.

Kieras, D., and Polson, P. G. An approach to the formal analysis of user complexity. *Int. J. of Man-Machine Studies*, 22 (1985), 365-394.

Kintsch, W. The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, 95 (1988), 163-182.

Kitajima, M., and Polson, P. G. A computational model of skilled use of graphical user interfaces. *CHI '92 Conference Proceedings: Human factors in computing systems*, ACM, New York, 1992.

Larkin, J. H. Display-based problem solving. In D. Klahr and K. Kotovsky (Eds.). *Complex Information Processing: The Impact of Herbert A. Simon*. Lawrence Erlbaum Assoc, Hillsdale, New Jersey, 1989, 319-342,

Mannes, S. M., and Kintsch, W. Planning routine computing tasks: Understanding what to do. *Cognitive Science*, 15 (1991), 305-342.

Mayes, J.T., Draper, S.W., McGregor, M.A., and Oatley, K. Information flow in a user interface: the effect of experience and context on the recall of MacWrite screens. In *People and Computer IV*, D.M. Jones and R. Einder, Eds., Cambridge University Press, Cambridge, UK., 1988.

Payne, S. J. Display-based action at the user interface, *Int. J. of Man-Machine Studies*, 35 (1991), 275-289.