Opportunities and Challenges for HCI Design and Research*

Jim Hollan

Distributed Cognition and Human-Computer Interaction Lab
Department of Cognitive Science, University of California, San Diego
(Email: hollan@cogsci.ucsd.edu Web: http://hci.ucsd.edu/hollan)

The decreasing cost, increasing power, and miniaturization of commodity computing devices continues to enable ever wider application of computation. The history of computing hardware can be summarized as a progression from a focus on low-level components towards integration on larger and larger scales, from vacuum tubes and transistors to LSI, VLSI, chipsets, personal computers, LANs, WANs, and now the global internet. Today there is also a move in the opposite direction as the monolithic "computer" is being unbundled into fragmentary components. Currently these components are reemerging coalesced in a multitude of forms, ranging from the rapidly evolving cell phone to novel embeddings in an expanding array of everyday objects.

But this machine-oriented view is far too narrow, because progression on the human side has been at least as dramatic and important, from isolated single users, to timesharing, to groupware and support for community activities, to the frontier where ubiquitous, wireless, context-aware, multi-modal, mobile computing enables currently unknown social possibilities. These radical changes and associated further imbuing of our professional and personal activities with computation present enormous new challenges for HCI design and research.

These challenges make this a most propitious time to convene a workshop to discuss the next generation of HCI and, I would argue, to take stock of where the field is, how we got here, and how best to proceed. Understanding computationally-based systems and ensuring their future design respects human needs and abilities are intellectual challenges of singular importance. These challenges involve not only complex theoretical and methodological issues of how to design effective computationally-based representations and mechanisms of interaction and communication but also confront complex social, cultural, and political issues such as those of privacy, control of attention, and ownership of information.

As with many challenges there is also opportunity. For example, the same forces leading to the unbundling of the monolithic computer and ubiquitous computing are also changing the nature and richness of data we can collect about human activities. In the history of science, the appearance of new technologies for collecting or analyzing data has frequently spawned rapid scientific advancement. The human genome project, for example, would have been unfathomably complex without automatic DNA sequencing technology. In the present case, a new generation of inexpensive digital recording devices is revolutionizing data collection for studying human activity, extending it to situations that have not typically been accessible and enabling examination of the fine detail of action captured in meaningful settings.

There are myriad important issues that would benefit from workshop discussion. Topics range from research strategy questions (e.g., tradeoffs between small-science and big-science approaches, the importance of open source for HCI, etc.) to how to better bridge the cultural chasms between the disciplines involved in HCI to support the interdisciplinary graduate programs so urgently needed to train the next generation of human-computer interaction researchers.

In this position paper I briefly comment on the proposed discussion starting point for the workshop, describe opportunities arising from new facilities to capture and share detailed records of activity, and discuss the importance of capitalizing on these opportunities as one component in elaborating a research agenda for next generation HCI research.

Proposed Starting Point: Reality-Based Interaction

The proposed starting point for the workshop is the notion of "natural or realistic interfaces or reality-based interfaces" that "increasingly draw their strength from ex-

*Position paper for CHI 2006 Workshop: What is the Next Generation of Human-Computer Interaction?

ploiting users' pre-existing skills and expectations from the real world rather than trained computer skills."

It is argued that

> *A unifying characteristic for much next generation HCI may thus be that it increasingly taps into the users' abilities and pre-existing knowledge. Direct manipulation moved user interfaces toward more realistic interaction with the computer; next generation reality-based interfaces push further in this direction, increasing the realism of the interface objects and allowing the user to interact even more directly with them. ...*
>
> *Finally, work that helped define the GUI generation is a model for us. Shneiderman took a set of disparate new user interfaces and unified them through their common characteristics [12]. Hutchins, Hollan, and Norman then explained their power and success of these interfaces with a theoretical framework [9]. Our hope is to take a first step in that direction for the emerging generation.*

Although our paper is now over two decades old, I still think the distinction it makes between two fundamentally different metaphors (conversation versus model world) for interface design has sufficient currency to warrant further discussion. I would also note that as we argued then, and as Rob reiterates in the call, simple mimicking of the real world is obviously insufficient. It would be valuable to discuss how to draw on users' knowledge of interacting with the world without being constrained to merely mimicking those interactions. This was certainly a central motivation for our early Pad++ zoomable interface work [2] and continues to motivate our current Dynapad work [1]. The goal has been to exploit people's wonderful ability to imbue space with meaning but remove or relax some of the restrictions of physical space.

A Permeable Boundary between Physical and Digital

A related idea worthy of workshop attention is that of making the boundary between the physical and the digital permeable. One example is Rekimoto's work on augmented surfaces [11] in which one can move windows from the virtual world of the display onto a physical table or from the physical world into the virtual world. Another particularly compelling example is Guimbretière's innovative work on Paper Augmented Digital Documents (PADD) [6]. In this approach, the digital and paper world are put on equal footing: paper and computers are simply different ways to interact with documents during their lifecycle. When paper affordances are needed, a document is retrieved from the database and printed. The printer acts as a normal printer but adds a pen-readable pattern to each document. Using a digital pen, the document can now be marked like a normal paper document. The strokes collected by the pen are combined with the digital version of the document. The resulting augmented document can then be edited, shared, archived, or participate in further cycles between the paper and digital worlds.

Distributed Cognition and Embodied Interaction

In terms of theory, it is important to note that there is currently a shift in cognitive science away from the notion of cognition as a property of isolated individuals and toward a view of cognition as a property of larger social and technical systems [8, 3, 7]. This extends the reach of cognition to encompass interactions between people as well as interactions with resources in the environment. As a consequence, the human body and the material world take on central rather than peripheral roles. As Andy Clark put it, "To thus take the body and world seriously is to invite an emergentist perspective on many key phenomena - to see adaptive success as inhering as much in the complex interactions among body, world, and brain as in the inner processes bounded by the skin and skull." [3] This new perspective on cognition is emerging from the fields of distributed cognition [8, 5] and embodied interaction [3, 10, 4]. I think it promises to provide an intellectual basis for a paradigm shift in thinking about the dynamics of interaction; one that takes material and social structures to be elements of cognitive systems and views on-going activity as a continually renegotiated emergent product of interaction. This paradigm shift promises to be vitally important for the next generation of HCI.

A Critical Opportunity

In addition to developments in theory, recent advances in digital technology present unprecedented opportunities for the capture, storage, and sharing of activity data. This is important because to understand the dynamics of human activity in sufficient detail to be able to make informed design choices, we must first understand the full context of those activities and this can only be accomplished by recording and analyzing detailed data of real-world behavior. I argue that the ability to capture and share such data has created a critical moment in the practice and scope of research. I think our field has an exciting opportunity to capitalize on this by developing a shared infrastructure to assist in integrating the theoretical and analytic frameworks required to build a stronger scientific base for HCI research.

While detailed activity data are certainly needed to advance science, more data cannot be the whole answer, since many researchers already feel that they are drowning in data. Data without appropriate theoretical and analytical frameworks do not lead to scientific advances. Another obstacle to fully capitalizing on the opportunity provided by digital recording devices is the huge time investment required for analysis using current methods. In addition, we need to understand how to coordinate analyses focused on different levels so as to profit from the theoretical perspectives of multiple disciplines.

Reducing the Cost of Analysis

Today the high labor cost of analyzing rich activity data leads to haphazard and incomplete analyses or, all too commonly, to no analysis at all of much of the data. Even dataset navigation is cumbersome. Data records are chosen for analysis because of recording quality, interesting phenomena, and interaction density—producing a haphazard sampling of the recorded set. Good researchers have a nose for good data, but with a tendency to focus on small segments of the record that contain "interesting" behavior, analyze them intensively, and then move on to the next project.

When analysis is so costly, few analyses can be done—so datasets are severely underutilized—and researchers come to have a large investment in the chosen data segments. Since each analysis may appear as an isolated case study, it can be difficult to know how common the observed phenomena may be. Larger patterns and contradictory cases can easily go unnoticed. Well-known human confirmation biases can affect the quality of the science when each analysis requires so much effort.

Increasing the Power of Analysis

Researchers studying activity in each of the disciplines involved in HCI are beginning to appreciate the importance of understanding patterns that emerge from the interactions of multiple dynamically linked elements. Such interactive patterns may be invisible to approaches that decompose activity into the more or less independent components created by historical distinctions among disciplines. This is why a multidisciplinary approach is necessary. But tools that match this multidisciplinary vision are also needed. Without them we cannot address questions such as: How shall we come to see patterns in interactions among the many modes of multimodal activity? How can we approach the long-standing problem of the relations between patterns that emerge on larger time scales with the short time scale mechanisms that give rise to them?

The richly multimodal nature of real-world activity makes analysis difficult. A common strategy has been to focus on a single aspect of behavior or a single modality, and to look for patterns there. However, the causal factors that explain the patterns seen in any one modality may lie in the patterns of other modalities. In fact, recent theorizing suggests that activity unfolds in a complex system of mutual causality. Analysis may still be based on decomposition of the activity, as long as there is a way to put the pieces back together again. That is, as long as there is a way to visualize the relations among the many components of multimodal activity.

In addition, as activity unfolds in time, describable patterns that take place on the scale of milliseconds are located in the context of other describable patterns that display regularities on the scale of seconds. Those patterns in turn are typically embedded in culturally meaningful activities whose structure is described on the scale of minutes or hours. Patterns at larger time scales are created by and form the context for patterns at shorter time scales.

The high cost of performing analyses on data that represent real-world activity means not only that too few analyses are conducted, but that analyses tend not to be shared. Most often when results of an analysis are published, neither the activity of doing the analysis, nor the procedure that was used are shared. This creates a situation in which most analyses are idiosyncratic and possibly non-replicable.

I will argue that a primary component needed to advance HCI research is an open software infrastructure to support capture of increasingly rich activity data, speed and improve analysis, and facilitate sharing with a wider research community.

A Final Thought

The current state of HCI owes a deep intellectual debt to research at Xerox Parc and the development of the Alto. Researchers in the early days at Parc designed a new computational environment with crucial attention to the importance of the interface. Is it not time to do this again? Might not a major limitation in building the next generation of HCI be that we are building additional layers on top of operating systems themselves designed with virtually no conception of the kinds of services needed today?

References

[1] BAUER, D., FASTREZ, P., AND HOLLAN, J. Computationally-enriched piles for managing digital photo collections. In *Proceedings IEEE Symposium on Visual Languages and Human-Centric Computing* (2004).

[2] BEDERSON, B. B., AND HOLLAN, J. D. Pad++: A zooming graphical interface for exploring alternate interface physics. In *Proceedings of the ACM Symposium on User Interface Software and Technology* (1994), pp. 17–26.

[3] CLARK, A. *Being There: Putting Brain, Body and World Together Again*. MIT Press, 1997.

[4] DOURISH, P. *Where the Action Is: The Foundations of Embodied Interaction*. MIT Press, 2001.

[5] GOODWIN, C. Gesture, aphasia, and interaction. In *Language and Gesture* (2000), D. McNeill, Ed., Cambridge University Press.

[6] GUIMBRETIÈRE, F. Paper augmented digital documents. In *Proceedings of the ACM Symposium on Suer Interface Software Technology* (2003), pp. 51–60.

[7] HOLLAN, J., HUTCHINS, E., AND KIRSH, D. Distributed cognition: Toward a new theoretical foundation for human-computer interaction research. *ACM Transactions on Human-Computer Interaction* (2000), 174–196.

[8] HUTCHINS, E. *Cognition in the Wild*. MIT Press, Cambridge, 1995.

[9] HUTCHINS, E. L., HOLLAN, J. D., AND NORMAN, D. A. Direct manipulation interfaces. *Human-Computer Interaction 1* (1985), 311–338.

[10] NUNEZ, R. Could the future taste purple? reclaiming mind, body and cognition. *Journal of Consciousness Studies 6* (1999), 41–60.

[11] REKIMOTO, J., AND SAITOH, M. Augmented surfaces: a spatially continuous work space for hybrid computing environments. In *CHI '99: Proceedings of the SIGCHI conference on Human factors in computing systems* (New York, NY, USA, 1999), ACM Press, pp. 378–385.

[12] SHNEIDERMAN, B. Direct manipulation: A step beyond programming languages. *IEEE Computer 16*, 8 (1983), 57–69.