

4.0 Human-Computer Interaction

Introduction and Definition

As soon as the users began to be other than the designers of computing machinery, the so-called human-computer interface became an object of design itself. That boundary has been loosely referred to as an “interface”, much as any other computer peripheral. But the term “interface” is losing its hold, mainly because it suggests a sterile, inflexible specification that doesn’t match the steady introduction of more abstract and natural means of communications we will call modalities. The more descriptive and anticipatory term is “interaction”.

The next twenty years will see two important ways computers will come to serve us. The first and undoubtedly most compelling vision is to see their migration from a hands-intensive tool toward a delegatable assistant. This concept is not an insensitive attempt to gloss over difficult concepts with anthropomorphism but to point to an inevitable direction whose ultimate measure will never quite be realized. Simply put, the evolution of human-computer interaction should be thought of as a direction and not an end. Straining to grant the human user an interaction for which training is not a prerequisite, the ever-affordable surplus of computing cycles will be more and more dedicated to making the computer not just user-friendly but user-like.

Thus, human-computer interaction (HCI) encompasses a very wide range of functionality and work. In general, it is all the functions, mechanisms, and conventions that provide the means for users to interact with computers. Some interface specifications or definitional languages may exist until natural language dominates. In a research context, it is all the enablements whereby humans and computers interact and is NOT an interface in the sense of the specification of a standard.

4.1 Principal Motivations

As with most machines man has built, the computer requires its user to learn what functionality the machine has as well as how to interact with it. Though the potential for vastly easing these burdens has been there in the case of the computer, it has given little quarter to its operator....so far. But, as stated above, the computer has the potential to offer its user a more abstract way to convey orders and the next decade or two will see substantial changes and improvements in this regard. Table I lists some of the reasons why this trend will occur.

The motivations listed in Table 1 can be aggregated into just two: the computer as an increasingly delegatable assistant to an information-centered user and the computer as an augmentation to the physical, not just intellectual human.

The more straightforward use of computers today is as a framework within which application programs are run. The processors, the connectivity, the operating systems are becoming invisible...subordinated to the applications programs at hand. Within a decade those too will begin to sink below the conscious surface, yielding to task-oriented computing at the next level of work abstraction. To make this possible, the way people interact with computers will take on some of the attributes used when they interact with other people. Within ten years, very acceptable speech recognition and understanding software will permit a wide range of computer tasking now confined to the keyboard and mouse. This will be important wherever keyboards are

Table I Motivations for More Natural Human-Computer Interaction

- The present wide range of inconsistent interaction environments requiring retraining as a user moves between systems
- The need for easier, more intuitive interactions to achieve objectives making human language the dominant, but not only, interaction language
- Tasks of greater complexity can be accomplished using higher-level abstractions with attributes such as completeness, veracity, timeliness....
- The notion of delegation - the computer or computer-based machine accepts increasingly abstract instructions and tasks
- The computer as a sensory and action mediator, matching human dexterity and senses to situations in which humans cannot function or easily adapt.

impossible to accommodate. This change does not mean over that time frame the computer will necessarily display very much reasoning power, but it should be capable of both understanding and synthesizing natural language.

An important conceptual aspect of HCI is just what the human is interacting with. Today's platforms suggest that the user interact with a local client or host; whatever is personal to the user is located within easy, direct reach. That notion may broaden considerably toward a personal space that is quite virtual.... distributed in ways the user may be unaware of. The virtues of such a convention may be that the user's location may one day be totally irrelevant. Present systems hint of this now but all access is cross-net to the user's personal physical machine.

As mentioned, the second compelling use of computers will be to augment the physical rather than just the intellectual human. While the range of such use may not be as broad as more mind-centered work, there are many situations where our dimensionality, our sensory package, our reaction times, our motor sensitivities, or simply our remoteness are not adequate to the task. Through relatively new concepts such as telepresence, computers will effectively transport us into other dimensions and places to let us view or interact with what we find. (See Personal Computing, Chapter 3.)

4.2 Important Counter-Pressures

Undoubtedly the most important resistance to a common and consistent interaction environment is the progress of technology. That movement is apparently in the direction of more natural, more abstract, more consistent levels of interaction but its form is of a few companies vying for the user's devotion to their operating environment. One important aspect of that competition is that the progress in functionality within each camp is both small and cloistered enough to keep users captive, all the while forcing a lifelong user adaptation. Today's user is a paragon of tolerance to these parallel and lucrative strategies of technology advancement and payment for upgrades. The concomitant advancement of the technology and this tolerance for small, incremental change are impediments to a global and consistent standard.

Another obstacle against an HCI standard is the continuing change in the nature of computers. In this case we refer mainly to their shrinking size and, to a lesser degree, their ubiquity. As computers get smaller, the keyboard is no longer an option and, if one is to retain some fluency in the interaction, speech understanding becomes important. Total ubiquity of computers is the preferred future for some, with the interface becoming either unobtrusive or a constant set of simple functions, the first implying a more natural interaction. But ubiquity may also arrive in the form of the long-discussed information utility with its countless uses. Managing the access and those uses would benefit from an intuitive and natural HCI. Of course, ubiquity also means that many users will only see the consequences of computers with no direct accessibility to them at all.

Environmental factors can also compromise or even invalidate a common and consistent interaction environment. In the wireless world, high noise or multipath will force more narrow bandwidths that may cause abbreviated frameworks for interaction, particularly if the HCI is centered in some remote server.

4.3 Models and Modalities

In this HCI discussion, “model” refers to the range of computer types and, to some extent, how they are used. “Modality” refers to the channel or medium that is used, such as pointing, speech, or keyboard. Table II shows a taxonomy of different models for HCI.

Table II Important Types of Human-Computer Interaction

- Desktop Model - Portability is not primary. Wired media is expected. Large displays, rich input output modalities.
- Mobility Model - Portability and power saving are paramount. Radio or wire. Much more constrained input output modalities.
- Distributed Model - May have some of each of the above. User sees constant environment regardless of location.
- Virtual Reality Model - Immersive interaction within a totally simulated environment.
- Augmented (Enhanced) Reality Model - Simulation overlay to real world environment. Generally a real-time process.
- Telepresence Model - Computer-mediated interaction with a real environment. Generally a real-time process.

Apart from the *models* in Table II is a collection of underlying *methods* used as a basis for building the interaction environment. Examples of these are: situation theory, tasking models, agents, learning-metaphors, requirements/tasks/usability, tool-based, common ground, and blackboards.

The spectrum of interaction modalities is defined by our communications-oriented senses: displays, speech, handwriting and gesturing, text, haptic feedback, and their integration. In most tasks we undertake, we benefit from the use of multiple modalities. Computers to date have been dominated by the keyboard, display, and mouse (KDM) convention. But the recognition accuracy of speech and handwriting are opening the door to other modalities...albeit slowly. Multiple-agent systems are now being demonstrated that permit very natural integration of the different modalities common to HCI.

4.4 Realizations Within 10 Years

In the context of the models mentioned above, here are some predictions about HCI over the next decade:

- The Desktop Model - As the best example of how the need for standardization defeats rational design, the omnipresent QWERTY keyboard seems eternal. Because of the need for authoring text, the keyboard will remain. But because speech is the most convenient method for addressing or seeking objects not displayed, it will see increased use and in circumscribed environments will yield almost 100 percent accuracy. Modality integration will be greatest under this model and expect a few interaction environments to illustrate modality independence. The infrastructure will be present to reach 100 Mbps peak data rates at the desktop terminal for about 25 percent of the market.
- The Mobility Model - Computing power in laptops and PDAs will be roughly equivalent to desktops but the input/output methods will lag. Speech and handwriting recognition will improve and some will be integrated together. The mobility model will be dominated by lap-tops of ever increasing power. Some smaller ultra-portable wireless systems will use speech in addition to a stylus and buttons.
- The Virtual Reality Model - A broad assortment of virtual reality systems and their components will become available through the entertainment industry. The present major impediment to VR, the combination of display resolution and rendering speed, will improve greatly, offering much more realistic, totally synthesized environments. But matching the eye's capability using head-mounted displays is a very daunting task. Extremely high resolution is needed and the costs to achieve even a million-pixel fields is now prohibitive. Today's displays also have a fixed focus for all objects, leaving depth perception and occlusion unreconciled. This and several other factors may lead to VR sickness after long use for some users.
- The Augmented Reality Model - Augmented reality systems require more modest computing power but a great deal more registration accuracy, probably at least an order of magnitude. This need is in the head tracking system, with the amount depending upon the size and distance of the object being interacted with. This is essentially the spatial synchronization problem between the real and synthesized worlds. The applications such as training, maintenance, and medical care are important enough, however, that the AF should help leverage the development of

this technology. Typical will be the projection onto real objects (or on special eyewear) of graphical images that reflect some training aids to the user. Some systems using this will be ready for AF employment within 10 years but it is not clear whether the non-DoD market will be big enough to do this without military investment.

- Telepresence Model - Systems will emerge for medicine, hazardous waste handling, and other tasks with accessibility problems. Telepresence surgical systems will be in clinical trials or beyond. Telepresence is probably the most relevant information technology for RPV systems, including the possible extraction of the fighter pilot from the cockpit. Scale transformations are also a use of telepresence. Satellite repair or space station construction are examples as are microsurgery and microsystem assembly. Telepresence systems will permit distance-spanning meetings with far greater realism than the “head in the box” video or computer conferencing of today. Again, AF R&D investment will likely be needed to speed telepresence technology along.

Perhaps the most intriguing question for a 10-year projection is just how content users will be to keep using the present dominant keyboard-display-mouse (KDM) configuration for HCI. It is clear it will not vanish and may have considerable longevity. The information technology world has other good examples of inefficient or awkward conventions that are maintained simply for the sake of standards and often because they are implemented using products that have very low marginal cost. KDM will likely persist.

KDM dominance will just slow but not prevent the inevitable movement toward interactions that are more natural for humans. It may mean that new HCI functionality may have to emerge from niche markets. For example, speech recognition, arguably necessary for input in powerful but small computers, has found its first utility in the labor cost saving replacement of operators. Continuous, speaker-independent, large vocabulary, real-time, domain- or context-constrained speech recognition is beginning to find utility. Speaker independent accuracies have gone from about 21 percent word error rate in 1987 to 3.6 percent in 1991 for a 1000 word Navy battle management vocabulary. More recently, error rates on Wall Street Journal text with much greater complexity were about 7 percent. Couple such recognition accuracy with the added utility of circumscribed, context-aware, natural language understanding and very useful applications arise. Thus, the increasingly capable ability to understand speech will find more and more applications in HCI over the decade.

Another evidence of progress in HCI is the integration of the important interaction modalities into more natural and intuitive combinations. Speech understanding along with the pen-based inputs of text, symbols, and gestures are being integrated through the use of agents that not only enable better interaction but are adaptable to both task and user preferences. An example is the Open Agent Architecture, a schematic of which is shown in Figure 4.

Our access to broad information stores has been much simplified by the search and relational addressing conventions of, say, the WWW. But the fusion of data taken from disparate data bases ultimately needs knowledge mediation and this decade will see research demonstrations of that integration. (See Information Access Technology, Chapter 6.)

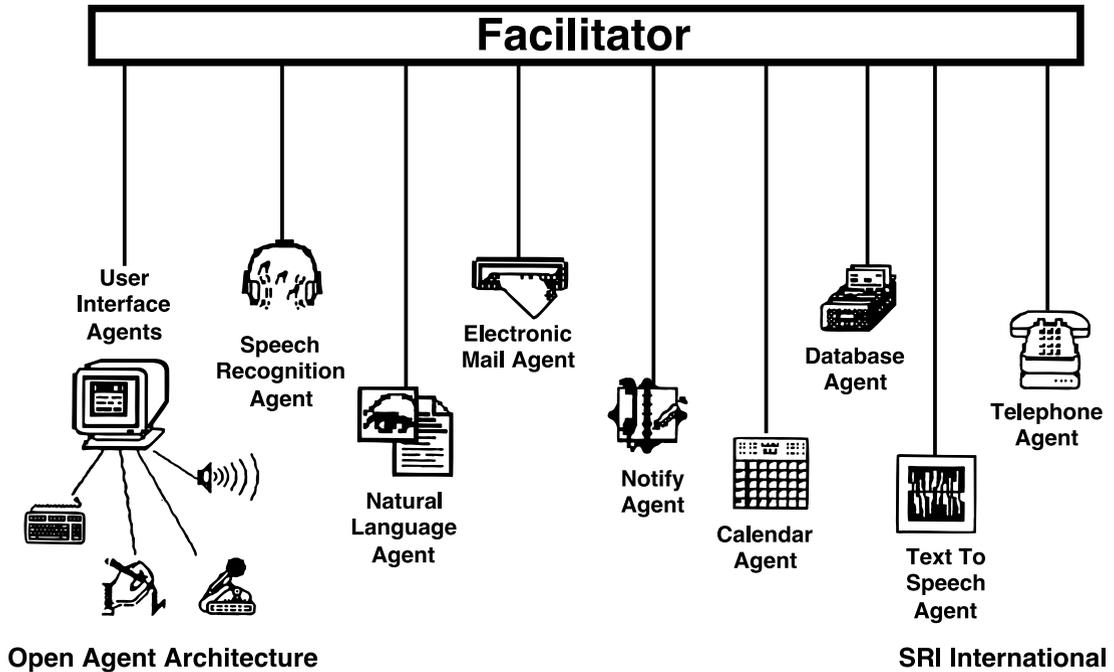


Figure 4 SRI's Open Agent Architecture

4.5 Realizations Within 20 Years

Again repeating the model metaphor from above, here are some longer term projections:

- The Desktop Model - There will be a very wide variety of task-oriented versions of HCI. Interactions between user and machine will be in natural-language, semantically flexible, and redundant. Displays and scanners will operate consistent with the limits of human vision. From a "home" machine full networking will enable direct bilaterally safe access to other machines, most with a common "look and feel."
- The Mobility Model - This will be embodied in three versions: ultra-portable or wearable, palm-tops of pocket size, and notebooks. The first is for speech input/output only or, in some cases speech/pen-based input. Some network connections and high resolution, eye-worn displays that can also double as projecting devices. The second is now called PDAs. They will become communications-rich terminals with, in some instances, enough local storage to make the limiting propagation and noise environment adequate for communication via differentials; i.e., only information different from that locally stored. Notebooks will have computing power equivalent to most desktops with perhaps less storage and I/O capabilities. (See Personal Computing, Chapter 3)

- Virtual and Augmented Reality Models - Improvements in these areas will be incremental over the second decade but that will ensure, at least technically, the availability of high resolution synthetic environments from virtually any location. Collaborative virtual and augmented reality environments will also be distributed and available. Adequate resolution displays and tracking system accuracy should be available.
- Telepresence Models - This type of computing will grant increasing accessibility for humans to otherwise inaccessible space. In medicine computers will mediate delay-bound remote surgery, microsurgery, all types of minimally invasive surgery and will do so giving the surgeon consistent, natural, sensory-rich accessibility. Manipulation of micron-sized objects with synthetic haptic feedback will enable the assemblage of nanosystems. Imaging and other human-compatible sensors will be compressed over communication links to permit remote experiences of high realism. Teleoperation will also benefit from computer-mediation that can, for example, guarantee the avoidance of undesirable actions.

Other supporting technologies will also be present:

- Totally integrated modalities will offer the flexible means to express a given semantic notion.
- Intelligent agents will learn users' intentions and either advise or react accordingly.
- User-crafted rules, conventions, appearance will yield a "behavior" for a computer keyed to its present user. That tailoring might then appear at any access point.

Displays will be both high resolution and very large. Mobile systems will have projectable displays.

4.6 Untethered Realizations

- Human interaction will be defined only by the task and user preferences - Until now the user has been more adaptation-prone than the machine. That will reverse itself. All tools should be selected or constructed by their user(s) according to the task. The computer as a tool is so malleable its utility seems boundless.
- The transparent computer - The computer is the first machine that can take tasking and delegation in abstract, human-compatible terms. How it accomplishes the task can become totally transparent (not necessarily irrelevant) to the one who is using it. One may wish to give it form, but the input and output devices may also be hidden and ubiquitous. If the interaction involves physical input/output objects, then the user will obviously face associated location constraints. (See Artificial Intelligence, Chapter 7.)