What is the Next Generation of Human-Computer Interaction?
CHI 2006 WORKSHOP

George Fitzmaurice
Autodesk
210 King Street East, Toronto, Ontario, Canada
fitzmaurice@acm.org

To answer what is the next generation of HCI, we need to look back a bit and determine what generation we are on now and what constitutes a generational shift. Everyone can agree that command line and GUI are two distinct generations. So, how do we know the boundaries of one generation from the next? For example, what about the cell phone and PDA interfaces? What about the internet revolution? One could argue that these systems still use the WIMP (windows, icons, menus, pointer device) elements and so belong to the GUI generation, yet I think they introduced some significant changes and usage models.

GUIs are composed with WIMP (Windows, Icons, Menus, Pointing device) elements. There should be a corresponding set of elementals for this next generation UI. However, do we have to start from scratch? Are we trying to reinvent the wheel? It's like saying the caveman invented the wheel so "what's the next generation of the wheel?" Yet, the wheel will always be valuable. It just continues to get refined and augmented.

In fact, perhaps we have recently reverted back to command line interfaces since we type in text for specifying web addresses and conducting web searches. But let's look forward.

Future Trends

In general I believe the domain of human computer interaction has expanded from focusing primarily on "office worker" activities to consumer activities (e.g., entertainment, content delivery and commerce). Here are some thoughts for the future HCI generational changes:

Less representational and more actual. We will expand our interactions to real world situations and I think "Augmented Reality" is a good encompassing phrase. Currently, GUIs deal with representations of things (e.g., a file icon represents a document). The next generation will deal with the manipulation of real world artifacts. Within the GUI desktop, interface elements will continue to look more realistic (e.g., like VCR transport controls for playing DVDs)

Provide transparent transactions. Our transactions with the computer are very explicit and everyone around us knows when we are interacting with a computer. In the future, we should have the capability to provide transparent transactions. That is, other people observing us are unaware that we are performing transactions (e.g., looking up a person's

name while having a conversation with the person). In terms of augmented reality, it is also possible to have transparent transactions with a computer while manipulating physical artifacts (e.g., an RFID tag embedded in our clothes signals when we are in a store).

No more storage management. Why do I have to keep managing my data (moving it between computers and devices, conducting backups, getting more storage, moving my data when I get a new system)? Perhaps there can be a central repository on the net (say iron mountain) that manages, stores and caches my data with an unlimited storage capacity. While this would be handy for a GUI based system, with an augmented reality system, I doubt we want to have to deal with storage issues. Instead the reality-based UI concept will focus on issues of *ownership* of an object and if I currently have an object *present* with me.

More pleasurable rather than efficient articulation. For explicit, human-computer dialogs, next generation interfaces should focus on defining pleasurable instead of solely efficient designs. For example, much of the research effort has been in designing efficient interaction techniques (e.g., reduce pointing time, reduce number of button clicks). This may be akin to the progression of design and overall experience for the automobile. The original automobile was purely functional. Slowly it evolved into offering luxury finishes, smooth manual gear shifting with high performance suspension systems.

These last three are more targeted towards existing GUI based systems:

GUI: Move from ugly aesthetics towards cinematic visual experiences. On traditional desktop systems, most of the visual experience is quite poor. Shifting towards a more cinematic visual experience is important when dealing with people who are more casual users. The Apple Mac OS X operating system is a good example here of shifting towards more cinematically rich visual experiences. For the augmented reality systems, the analogy does not hold up as strongly – it means shifting from real world objects and physical environments to highly stylized environments (e.g., theatrical stages or Disney world) and desirable artifacts (e.g., iPod nano).

GUI: Tackle the huge learnability of complex software systems. Many applications are very difficult to learn from an initial "walk-up and use" situation as well as providing continual learning and improvement for the user. How can we provide better ways of finding and presenting this knowledge? One way may be to use video-based examples.

GUI: video friendly interactions. For GUI based systems, our future designs must handle video as a core element. We need to be able to go beyond the idea of creating, editing, or just watching a video. I think the time has come where video can be used as an input stream such that the content is analyzed, objects are recognized, tracked and selectable with embedded links to augmented content.

Target a task or activity

In all of this analysis, we first must be guided by determine what people want to do. Next we need to determine if and how we can improve upon their desired activity by introducing computational artifacts. Only then can we begin to design and determine the proper input/output devices, computer infrastructure and preferred interaction model.

Here are my favorite samples of tasks or activity for the next generation:
- As we may think ("memex") by Vannevar Bush. His vision was to augment intelligence by providing a memory extender.
- Retrieve and display a person's name when you meet them.
- Commerce. The ultimate killer application. Shopping occurs in the real world with real artifacts and is ubiquitous.

Real World Interaction (RWI)

With digitally augmented real world interactions, there are a number of determinations users will conduct when interacting with the digitally augmented objects. These properties include the :

(1) *Scope of operability*. What environments does this object operate in? For example, it only works when it is on a sensing table or above ground.

(2) *Location sensing.* The object has an awareness of it's physical location (such as GPS) and reacts to events based on locations.

(3) *Sharability.* Will the information or state of the digital object be transmitted to other elements or users in the interactive system for collaborative use?

(4) *Online or Offline states*. Does the digital object have access to the internet, cell phone network or a personal PC network? Secondly, if the digital object has network capabilities, is it currently connected or disconnected.

(5) *Data flow*. Does the digital object represent a "push" technology or a "pull"? A push technology is one that if an event happens with the object, the results are manifested at a different location. A pull technology is one that if an event happens with the object, the object reacts and has added information or value immediately available (often manifested with visual display or feedback).

(6) *Anatomy of a transaction*. What signals the start and end of a transaction? What is the duration of a transaction? Some transaction could last days or months, etc.

(7) *Persistence of transaction*. Users need to determine how lasting the effect of the interaction will be – how long with this information be stored and relevant for future interactions.

(8) *Privacy of transaction*. Who can access information retained by the digital object or the transaction log of interactions the digital object has participated in and possibly transmitted to other devices on a network?

All of these descriptions have an implicit assumption that a physical artifact serves as a container or conduit for some digital artifacts. Consequently, the physical characteristics of the physical artifact will determine the "chunking" and phrasing of the combined physical and digital interaction.

Sample technology for discussion - Electronic Shelf Labels

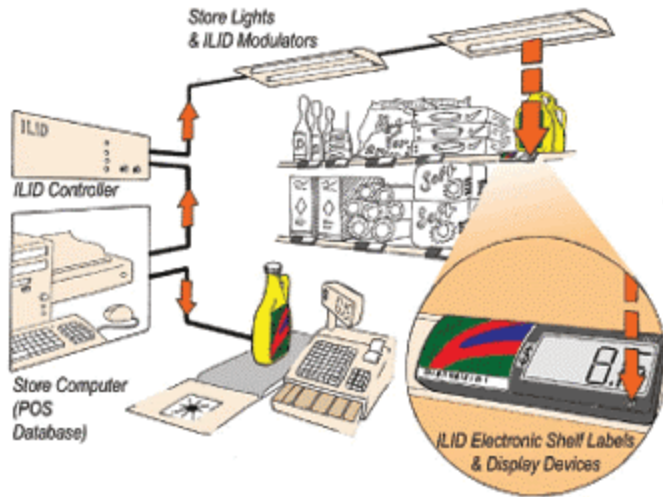1) NCR RealPrice (http://www.ncr.com/en/products/hardware/esl.htm)

These electronic shelf labels allow prices to be changed dynamically and almost instantly throughout the store by communicating with a wireless base station. They are a good example of augmenting the real world with digital computation and displays.

2) ILID Fujitsu[1]

"ILID is a unique technology that uses a building's existing lighting system to carry information to electronic display devices such as electronic shelf price labels. The technology provides a simple low-cost means of transmitting data from a central computer to a large population of receiving devices.

---

[1] Excerpt taken from ILID Fujitsu web page:
(http://www.fujitsu.com/au/services/industries/retail/solutions/esl/)

ILID Electronic Shelf Labels are fully portable, self-powered and operate in adverse environments such as freezers and fresh produce areas. Because light is used to transmit price changes and other information, there is no need for radio frequency or infrared transmission hardware. The devices themselves can be located almost anywhere – they do not need to be in a direct line of sight as they will work with reflected light or even light transmitted through a translucent object.

The ILID system revolutionizes the way both retailers and customers view product pricing. For retailers it becomes an effective pricing tool by eliminating time-consuming price-marking and by simplifying product promotion. For customers it guarantees that the price on the shelf will be identical to that paid at the checkout. Also, other information such as savings on discounted prices can be quickly and easily displayed."