From cave dwellers to Neuromancers

A brief history of interface design

by

Toby Chance
Managing Director
eCompany Interactive Solutions

Conference on Human-Computer Interaction in South Africa
University of Pretoria

9th May 2000

"*Nothing is in the mind that did not first pass through the senses*" - Aristotle

"The next care to be taken, in respect of the Senses, is a supplying of their infirmities with Instruments, and as it were, the adding of artificial Organs to the natural... and as Glasses have highly promoted our seeing, so 'tis not improbable, but that there may be found many mechanical inventions to improve our other senses of hearing, smelling, tasting, and touching." - Robert Hooke, 1665[1]

"Cyberspace: The tablet become a page become a screen become a world, a virtual world. Everywhere and nowhere, a place where nothing is forgotten and yet everything changes." - Michael Benedikt, 1991[2]

Life's rich tapestry

The Museum of the Middle Ages in Paris is, perhaps, an unusual starting point for a paper on the history of interface design. But step into the most celebrated room in the museum, (as I did in December 1996) displaying the set of six tapestries known as The Lady and the Unicorn, dating to between 1484 and 1500, and you will soon see the connection. As Dutch writer Cees Nooteboom noted of his own pilgrimage there recently, "I was alone in the fifteenth century, in the middle of a pictorial allegory, consisting of five senses and a riddle." [3]

In formalistic terms, the tapestries display a unique content and style within the Milles Fleur - thousand flowers - weaving tradition. They are large (in some cases life size) and pulsate with rich colour, symbolism and allegory. But it is their names that struck me most. Five of them are named after our senses - *The Sense of Sight, The Sense of Hearing, The Sense of Taste, The Sense of Smell, The Sense of Touch* - while the enigmatic sixth is named *A mon seul désir*, abbreviated by some as simply, love. For one such as me steeped in a 'multimedia' frame of mind, the sight of these artefacts from another era, named in contemporary terminology, was too much for my imagination to resist.

We know nothing about the artist or the weavers - this was an age before the artist had an individual persona, but we can imagine that had they been living today they would have

---

[1] Robert Hooke, Micrographia, 1665
[2] Benedikt, Michael, *Cyberspace, First Steps*, (MIT Press, Cambridge, Massachusetts, 1991), p 1
[3] Nooteboom, Cees, *Life's Rich Tapestry: The Lady and the Unicorn* (Qu**art**erly, Winter 1999, The National Art Collections Fund, UK), p 39

been designing an elegant 3-dimensional navigation system, or animating Toy Story 2 or directing a new historical drama for pay TV. Our skills as artists and technicians, and more recently scientists, are perennial and appear in different guises from generation to generation using new technologies which our unceasing creativity give birth to at an ever accelerating pace.

This paper is an eclectic look at how we made this journey and created the rich tapestry that is the modern human-computer interface, expressed through the languages and traditions of art, science and technology.

The focus of our lens settles on two radically different timescales. A wide lens reveals the evolutionary path taken by man from a primitive animal to a sentient being, specifically focusing on how we have invented means of communication that make highly refined use of the different sense organs.

A narrow lens, by contrast, provides a more detailed focus on developments of the human-computer interface since around 1960. Here, we shall encounter the discovery of the world of computers by artists and the world of art by computer scientists. Specifically, the evolution of the multi-disciplinary field of human-computer interaction has been the quest for a new body of knowledge to make sense of the technology that defines the Information Age.

One constant theme running through both accounts is the concept of the user. In the early days of computing, users were the unwashed masses, and computer scientists were the appointed priesthood. Throughout history, enlightened revolutionaries have invented technologies that have liberated the user - the individual - from the tyranny of the expert and clandestine knowledge. We could happily equate Douglas Engelbart's invention of direct manipulation of an information space with Giotto's early experimentation with perspective or Gutenberg's invention of movable type: in all cases, the role individual is enhanced.

The interface and perception

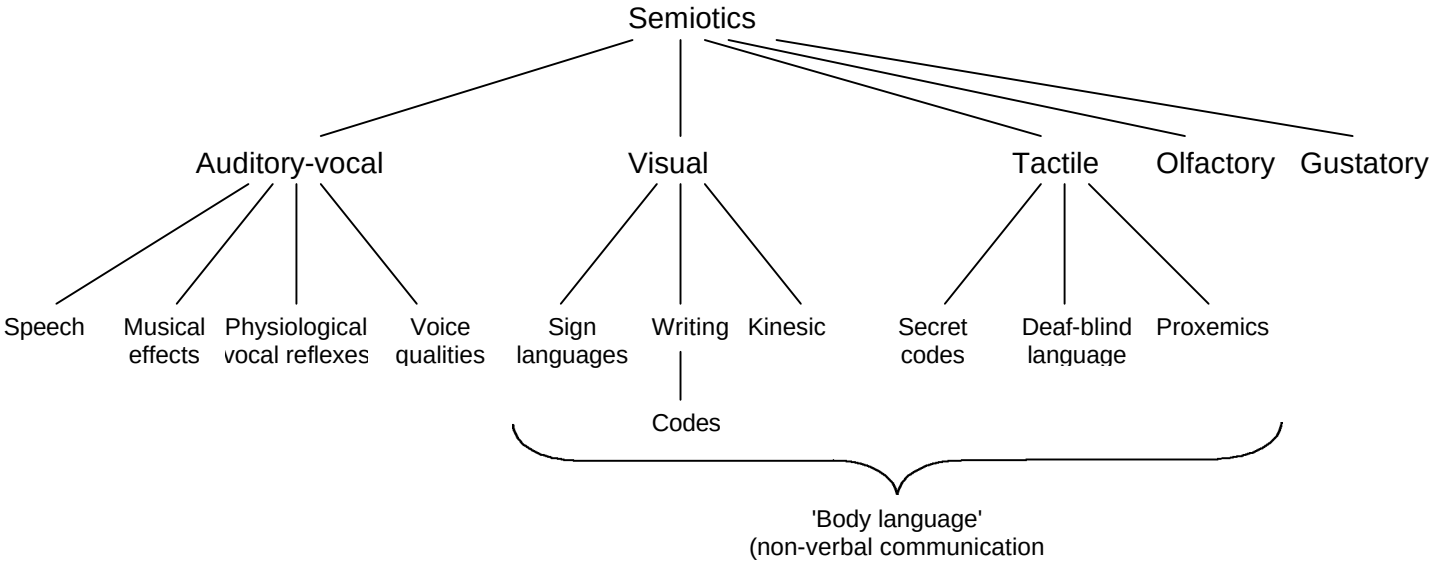
The "interface" describes a place 'between', somewhere in the middle, a space mediating adjacent objects through a subjective haze. The very need for an interface is premised on difference. The primitive human hand interfaced to tools, and the earliest hand weapons and tools attest to man's innate ability to shape them around his own form, to augment his physical ability to perform a task. This example indicates two fundamental elements of the interface: it is about form and about function. At its most basic, for an interface to work, it has to 'fit' and, where language is the interface, it has to mean something to the parties concerned.

This leads to a third element of the interface, and that is the central role it plays in perception. Our five senses generate stimuli in the form of nerve signals that are interpreted by the brain, which responds in turn by generating thoughts, physical actions or sounds. To use systems terminology, all of the five senses receive input. It is left to our speech facility, body language (including eye movements) and our creation of media and artefacts to communicate output

Using one organ, the eye, we can perceive colour, shape, form, space, light, movement, dynamics and balance, (after Arnheim) while using the ear we can perceive pitch, volume, resonance. Touch has several facets as well - including the senses of feeling pressure and temperature on the skin, and the proprioceptive sense of knowing where one is in a three-dimensional space. These capacities of perception are what enable us to differentiate

between image and text, between words and music, and between the feel of a keyboard and the touch of a mouse.

The domain of semiotics, which "investigates the structure of all possible sign systems, and the role these play in the way we create and perceive patterns (or 'meanings') in socio-cultural behaviour"[4] presents us with the diagrammatic relationships between the senses as they pertain to language:



This is a useful guide but does not show the whole picture, for it omits the non-symbolic visual communication media that play a vital role in our story - painting, architecture, drama, film, TV and so on.

The history of the interface is really the history of man's biological development (of sensory perceptions) interacting with his technological development (of tools and mediums of communication). Man has applied his imagination and communication organs to the invention of language expressed in painting, sculpture, architecture, poetry, music, theatre, film, computer games and so on. As man has invented new mediums of communication, that stimulate the senses with varying degrees of intensity, so our responses and view of the world have varied. As McLuhan said, "Electronic media alter the ratios between the senses". This underlies the meaning of another of his famous epigrams, "The medium is the message."

For millennia, the dominant medium was the painted or drawn image. Around 10 000 years ago, man began to develop an awareness of three dimensions and constructed dwellings and later grand structures to express his place in the world and to communicate symbolically. Around 5 000 years ago, language was recorded as writing, but only among the tiny minority of literate individuals. It was not until the adoption of movable type in the 1450s that knowledge through text became widely available, and this led directly to the Renaissance and the Scientific Revolution. The 20th century is the age of electronic media - radio, telephone, television - that has accelerated the pace of economic and cultural change even faster. This was primarily an analogue world, in which information was created and

[4] Crystal, David, *The Cambridge Encyclopedia of Language*, (Cambridge University Press, 1997) p 399

distributed in a linear fashion - the electronic equivalent of linear text. Control of the message was very much in the hands of the author, and power was wielded by the distributors of information via the mass media.

Now we are in the early years of what might be called the digital era, characterised by randomness, interactivity, and increasing control of the message by the individual. The development of interfaces for mediating between man and this world began around 1960 and is now manifest in the convergence of all our previous technologies into what Nicholas Negroponte in 1970 called "media technology", known today as multimedia. The dominant expression of this technology is the Internet, but there are offshoots such as interactive television, mobile telecommunications and immersive computer games that are all part of the proliferation of digital media that characterise our age.

The sense of sight

The magical power of the image

The history of art is clearly beyond the scope of this paper, so I will highlight just some of the developments that are relevant to our retrospective look at interface design. I have chosen, fairly arbitrarily, the role of images in communicating and exercising control over our environment, and the use of metaphors and icons for organising and presenting knowledge visually.

For early man the depiction of images was the dominant means of recording information and communicating meaning to the group. This was a pre-literate age, but many of the features of contemporary interface design were present even then. The images created by the cave painters at Lascaux in France and the Drakensberg are examples of man's first attempt to express his imagination as art. But this was not art as we understand it today. "Palaeolithic artists were not making works of art in order to express their personal view of the world, they were attempting works of magic".[5] Cave paintings were communicating with the higher powers through the medium of painting in order to understand and influence the physical world. At a higher level of abstraction, Rheingold believes cave paintings were "primitive but effective cyberspaces" because they were used to create a "three -dimensional sound and light show" that was possibly used to "imprint information on the minds of the first technologists."'[6] Both views suggest the use of images as the first truly mind-altering medium.

The cave painters used metaphor as a device to communicate meaning. "The experience of trance is so overwhelming and so difficult to describe that people have to use metaphors - that is, they have to compare their experience with some more familiar experiences that others can understand."[7] The painters were often the shamans of the community and, for example, used images of dying eland to express the experience of trance: perhaps the earliest use of visual metaphor. In this case, the painters were communicating with the other members of the family group, using a visual language that could be understood by their uninitiated brethren.

The use of metaphor to simplify and communicate meaning is at the heart of the desktop interface. When attempting to take the mysticism out of computing, the designers at Xerox

[5] Storr, Anthony, *Solitude*, (Harper Collins, London, 1999) p 75
[6] Wooley, Benjamin, *Virtual Worlds*, (Blackwell, Oxford, 1992) p 243
[7] Lewis-Williams, David and Dowson, Thomas, *Images of power: understanding Bushman rock art* (Southern Book Publishers, Johannesburg, 1989) p 50

PARC took a physical space containing well-known objects with clearly understood functional characteristics - an office containing filing cabinets, folders, trash cans, documents, drawing, writing and erasing tools - and represented them on a screen. 'In Poetics, Aristotle defined metaphor as the act of "giving the thing a name that belongs to something else". The crucial element in this formula is the difference that exists between "the thing" and the "something else"'.[8] The power of the desktop metaphor is its superficial resemblance to the real thing without its attempting to be a simulation. It still leaves a role for the imagination to make sense of the missing bits to create meaning for the individual in the accomplishment of a task.

Perceptual experiences are 'seen' primarily in images and then coded into words. The interplay between images and words in communicating meaning is illustrated in the spatial pneumonic device known as the Memory Palace. The Memory Palace was first exploited by the Greek Simonides in the 6th century BC as a way of remembering the positions occupied by people at a grand banquet. It was a technique fully exploited by the rhetoricians in their professional need to remember long passages of narrative. The Memory Palace was perhaps "the original information-space, stories turned into architecture, abstract concepts transformed into expansive - and meticulously decorated - imaginary houses."[9] The Memory Palace exemplifies that old saying "in the mind's eye". The mind creates pictures because it is this sense that best facilitates memory. We can remember the characteristics of a face for longer than we can remember the name of the person.

The converse of this is the creation of an image that has some formal relationship to what it represents. This is the primary function of icons - they are visual signifiers of meaning. Thus, a photograph or painting of a tree are iconic. In the hierarchy of mentalities described by Kay, in human development the iconic is preceded by the enactive (know where you are, manipulate) and is superseded by the symbolic (tie together long chains of reasoning, abstract).

A feature of icons and metaphor is that they compartmentalise thinking into pre-determined units of meaning defined by the artist or creator. They make assumptions about what today would be called the user's mental model. Interface designers have been grappling with this since the invention of the graphical user interface.

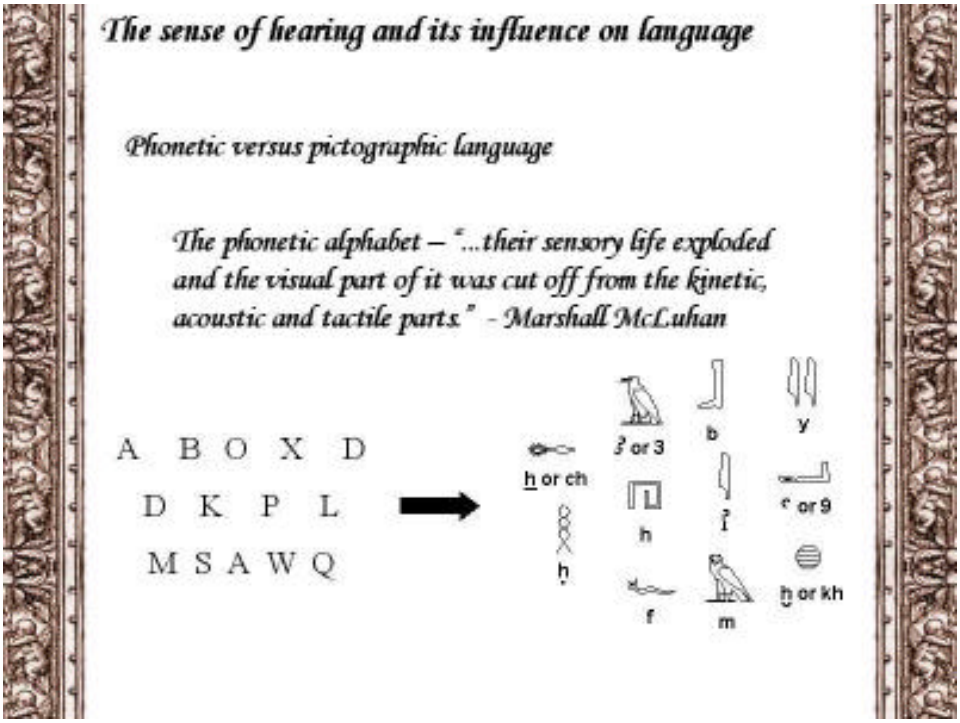
It was the inadequacy of images that led to the development of the symbolic mentality manifest in pictographic and alphabetic text. A similar historical inflection point has been reached in interface design: "One of the problems is how to get concrete signs to be more abstract without simply evoking the kinds of symbols used by the symbolic mentality." [10]

The appearance of text

Writing took hold as the dominant means of communication, expressed in visual form as pictograms, in the Fertile Crescent around 8 thousand years ago. The Rosetta Stone was the key to unlocking the meaning of these symbols, as they evolved in Egypt, because it was inscribed in two other languages understood by scholars, in particular ancient Greek. Socrates believed writing had disadvantages as a communication medium when compared to speaking: "The fact is, Phaedrus, that writing involves similar disadvantages to painting.

[8] Johnson, Steven, *Interface Culture*, (Basic Books, New York, 1997) pp 58-9
[9] Ibid, p 12
[10] Kay, Alan, *Interface: a personal view*, in Laurel, *The art of human-computer interface design*, (Addison Wesley, 1990) p 202

The productions of painting look like living beings, but if you ask them a question they maintain a solemn silence. The same holds true of written words:…"[11]



Text, especially in the form of the stone tablets and manuscripts that were used in Socrates's time, was indeed an inert medium, and the preserve of the privileged few. The invention of movable type by Gutenberg in the mid 15th century and the spread of printed text to a wide audience helped drive the attempt to recapture the culture and learning of those classical times. The surge of creativity in the arts and sciences in the period 1450 to 1750 was powered by a medium of communication that was highly adaptable, reproducible and accessible.

When we say something is 'written in stone' we imply it is permanent, cannot be changed either by the writer or his audience. The printing press brought writing to a larger public and the book overtook the manuscript as the principal format for recording text. It is extraordinary how we have hung on to the traditions of the earliest forms of writing. Today we have the digital 'tablet', the 'notebook' computer, 'scrolling' text, 'leading' (derived from the use of lead to position text on a printing block), and many other ancient conventions brought into the digital world. We have adapted the best (and some of the not so good) conventions to the modern user interface because they have become ingrained in our culture.

The Victorian novel was a particular form of textual expression that played a similar role in 19th Century society as the cave paintings did in the Neolithic Age. "The novel was the answer to the question: 'what connects all these bewildering social realities?' ".[12] The novel perfected the intricate narrative form of textual story-telling and became a way for authors to uncover the mechanisms of culture and society and explain it to their audience, just as cave paintings elucidated the shaman's interpretation of the magical powers that they believed governed their society.

[11] Mengham, Rod, *The descent of language* (Bloomsbury Publishing, London, 1993) p 31
[12] Johnson, ibid, pp 32-3

The Memex is the bridge that linked the old, mechanical way of recording and communicating information to the modern digital way. The Memex was conceived in 1945 by Vannevar Bush in a classic paper *As we may think*, in which he described a machine that was a "kind of mechanised file and library." "Bush's vision was remarkable. He not only foresaw the application of the computer to information storage and retrieval, and the value of associative indexing in that activity, but he also correctly anticipated the multimedia nature of computer use in the future."[13] The Memex threw out the old notion of text as fixed strings of information. Instead, text - indeed other media, including pictures and voice clips - could be "gathered together from widely separated sources and bound together to form a new book. It is more than this, for any item can be joined into numerous trails…."[14] It is no wonder that the Memex was the touchstone for the new generation of thinkers - Douglas Engelbart, Theodore Nelson, Alan Kay - that applied its core principles to the digital medium of the computer.

The sense of hearing and it influence on language

For most of human history, the oral and aural world has predominated over the visual world. That changed in the West with the arrival of Greek civilisation. Around 500 BC, "something happened which flipped them out of the old Homeric world of the bards into this new rational, philosophically logical, connected, private, individualistic, civilized world. And that thing is called the phonetic alphabet." [15]

The Western philosophy of language assumes that speech came before writing, that writing is the pictorial means of representing speech, that alphabetic writing established an effective means of representing sounds, and that the superceding of pictographic by alphabetic writing should be judged as evidence of cultural progress (after Mengham). Though this way of thinking is now being questioned, there is no doubt that the breaking down of language into phonemes represented by alphabetic letters has lead to a distinctive Western culture when compared to the older and now largely Eastern practise of written language comprising glyphs and pictograms that have no direct relationships with patterns of sounds.

"The Chinese ideogram is a wonderful instrument of unified sensations".[16] McLuhan, though he did not know it because he was fifteen years before his time, foresaw the unification of the senses in form of the graphical user interface. The command-line interface common to computing before the mid 1980's came out of the engineering and computer science community, who were schooled in the Socratic, Greek tradition of thinking and knowledge representation. The graphical user interface was created largely by mavericks who eschewed these conventions and aimed for an earlier form of knowledge representation. They achieved re-unification of the senses in the interface through, for example, the representation of whole thoughts as icons with meaning embedded in them, rather than sounds as letters and numbers with no individual meaning.

According to McLuhan, we are re-entering an aural, or what he calls an acoustic world, driven by the electric media. "In the electric world, the simultaneity of information is acoustic in the form that it comes from all directions at once."[17] For McLuhan, the electric acoustic world has "no continuity, no homogeneity, no connections, and no stasis." (He uses the word

[13] Preece, Jenny, and Keller, Laurie, *Human-computer interaction* (Prentice Hall and The Open University, Hemel Hempstead, 1990) p 5
[14] Ibid, p5
[15] McLuhan, Marshall, lecture at Florida State University, 1970
[16] McLuhan, ibid
[17] McLuhan, ibid

connections not in the sense of hyperlinks but in the sense of family and stable patterns of society). He was writing at a time when television had become the modern bard. Today, we can confidently say that this reality has been accentuated by the arrival of the Internet, mobile communications and interactive TV. Our interfaces are having to cope with a simultaneity of information bombarding our senses from all directions.

Another sign that McLuhan may be right in his analysis is the growing interest being shown by researchers in the marrying of narrative and interaction in modern media and interfaces. The soap opera, sitcom and period drama are genres that hark back to the Homeric tradition of story-telling. Greek theatre was the first incarnation of this media type. Computer games, MUDs, chat rooms and interactive novels are today's digital equivalents, and are harbingers of greater things to come.

One of the Holy Grails in interface design is a natural language or speech recognition capability. One could argue that a speech interface will forever change the relationship we have with computers. Through its immediacy of interaction, it will accentuate conversational interaction and take the computer further towards Engelbart and Licklider's goal of acting as a device for augmenting the intellect. "A linear medium cannot represent the simultaneity of processing that goes on in the brain - the mixture of language and image, the intimation of diverging possibilities that we experience as free will." [18]

The conversational interface is itself soon to be obsolete, according to John Walker and others who have predicted the exploratory interface as the next boundary for human-computer interaction. The sense that makes this possible is the sense of touch.

The sense of touch

"Touch is the oldest sense, and the most urgent. If a sabre-tooth tiger is touching a paw to your shoulder, you need to know right away." [19]As I said in the introduction to this paper, the sense of touch has to do with tactile sensation but also with gesture and spatial awareness of the body's positioning in a physical environment - the proprioceptive sense. Science fiction and horror writers use the "touchability" of an experience to accentuate its realism. 'Sending shivers down my spine', 'it seemed so real I could feel it', 'it makes my hairs stand on end' - these are expressions of thoughts having involuntary physical manifestation, and show the interplay between physical sensations and the psyche. Aldous Huxley, in Brave New World, written in 1932, wrote:

"Sunk in their pneumatic stalls, Lenina and the Savage sniffed and listened…
The house lights went down; fiery letters stood out solid and as though self-supported in the darkness. THREE WEEKS IN A HELICOPTER. AN ALL-SUPER-SINGING, SYNTHETIC-TALKING, COLOURED, STEREOSCOPIC FEELY. WITH SYNCHRONISED SCENT-ORGAN ACCOMPANIMENT.
'Take hold of those metal knobs on the arms of your chair,' whispered Lenina. 'Otherwise you won't get the feely effects.' "[20]

The image he creates leaves one feeling faintly repulsed and un-nerved. A less dystopic version of this same idea is Morton Helig's Sensorama and is at the heart of virtual reality systems developed from the mid 1980s.

[18] Murray, Janet H, Hamlet on the Holodeck: the future of narrative in Cyberspace, (The Free Press, New York, 1997) p 281
[19] Ackerman, Diane, *A natural history of the senses* (Random House, New York, 1990) pp 80-81
[20] Huxley, Aldous*, Brave New World* (Bantam Press, New York, reprint 1953) p 134

More mundanely, our ability to physically identify objects and materials requires the interaction of highly complex receptors and brain functions that technology has not come close to replicating. The science of robotics is mostly concerned with simulating touch, and in marrying that with cognitive psychology we are beginning to invent machines that create tactile pictures that serve the same purpose for the sense of touch as visual pictures do for the sense of sight. This form of artificial reality is particularly important for the creation of future human-computer interfaces because, as mentioned at the beginning of this paper, speech and body movement are the only two ways in which humans communicate outwardly with the world. (Direct conversion of brain signals into actions or synthetic speech is still some way off).

An important aspect of touch is resistance[21] and force-feedback devices are still mainly confined to the research laboratory. An early exception was Atari's video game *HardDriving,* which entailed the driver receiving force-feedback via the steering wheel. A contemporary technology is "Phantom", a versatile 6-degree-of-freedom force-feedback input device that has been used in medical training. It is easier to create input devices for touch than to create output devices, and perhaps the most commonly used is the touch screen. The touch screen and its cousin the light pen are appropriate to certain categories of application, but the fact that they have not been widely adopted suggests that the communication they facilitate can be better achieved using other interfaces.

The senses of taste and smell

The senses of taste and smell are our most primal, the least studied by artists, cognitive psychologists and computer scientists, and are thus marginal to the history of the human-computer interface. That does not mean that that they are not complex. To identify the smell of a rose, the brain analyses over 300 odour molecules. The average person can discriminate between 4,000 to 10,000 different odour molecules. Researchers have discovered that an odour can only be detected in liquid form. Hence, taste and smell work in tandem.

Evolution of homo sapiens has rendered these senses the poor cousins of the human sensorium. Morton Helig's Sensorama attempted to simulate the smell of a New York street, but the dominant media of our time ignore it. Devices that can create simulated smells are only now emerging from the laboratory into commercial products. Poetry and cooking might pay them the some respect, but somehow they do not lend themselves to the world of the computer. I shall allow them a graceful exit from our story.

Space, time and movement

The perception of space and time is essential to our concept of reality. I group the two together for the sake of brevity, but also because they are linked closely in the theoretical and the physical worlds. Mathematicians from Euclid to Newton to Einstein to Hawking have grappled with their interrelations. So have artists, architects and designers in their attempts to represent them in 2-dimensional and 3-dimensional media and structures.

As Rudolf Arnheim notes, "Movement is the strongest visual appeal to attention."[22] Movement entails change both in space and time, and the important role its depiction plays

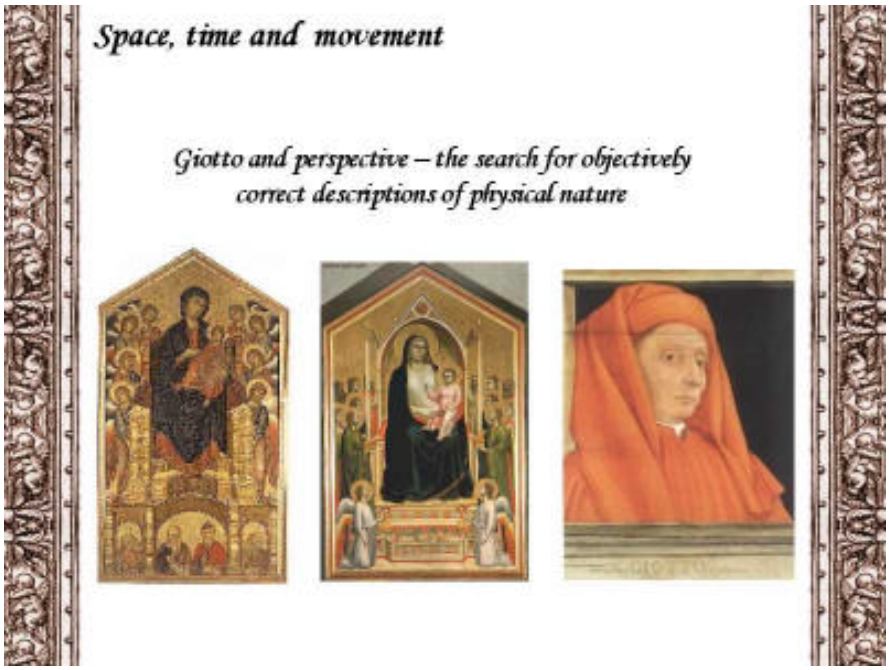
[21] Krueger, Myron, *Artificial Reality II* (Addison-Wesley, 1991) p 143

[22] Arnheim, Rudolf, *Art and visual perception: a psychology of the creative eye, the New Version* (University of California Press, Berkeley, 1974) p 373

in the modern human-computer interface can be seen by the impact animation and video have had on popularising the use of computers.

For primitive man, moving images attracted the eye because they signified danger. Today, they are more likely to signify excitement, opportunity, playfulness. The first pictorial expressions of movement did not require the third dimension of depth, nor even the fourth dimension of time. The sense of depth was conveyed by the layering of paint and markings on protuberances inside caves, which when seen in the rays of light at a certain time of day gave the impression of three-dimensional figures. Movement was conveyed through the positioning of human and animal limbs at various stages of articulation, such as during the pursuit of animals or during a trance dance.

In the 1400s a Florentine artist, Giotto di Bondone, invented a way of painting a three-dimensional space onto a 2-dimensional canvas. Perspective had been born, and was to have a fundamental effect on art and civilisation. (Actually, the Greeks invented a version of it in their design of tapering columns in the Acropolis to provide the illusion of parallelism). For the first time, paintings were represented from the individual's point of view, not from "some disembodied or divine locus"[23] Artists, architects and scientists of the Renaissance, like Brunelleschi, Albert and Leonardo, picked up on perspective in their "search for objectively correct descriptions of physical nature" that "led to the great voyages of exploration as well as to the development of experimental research and the scientific standards of exactitude and truth." [24]



Perspective is important also because it began to focus the attention of the artist on the individual. Though artists generally do not intentionally create their works for consumption by users, the fact that 500 years ago they began to place the individual human experience at the centre of their painting showed that the concept of "I" was taking hold. 'If I can see it from my perspective, you can see it from yours.' When applied by commercial artists and designers, the viewer's perspective had important consequences. It resulted in design for

[23] Johnson, ibid, p 214
[24] Arnheim, ibid, page 283

mass consumption based on an appreciation of the ergonomic and cognitive limitations of the user. One could draw an analogy between an abstract artist who paints imagery that is inaccessible to the unschooled viewer except through an expert's interpretation, and a programmer who writes an application that is unusable without interpretation by an interface. In both cases, the needs of the viewer are catered for by an interpretation, or interface.

A central feature of the current age is the acceleration of events through time with no regard for geography, leading to a growing cognitive overhead for users of devices that communicate these events - whether they be newspapers, TV or computers. The challenge for the interface designer is developing interfaces that can cope with this 'electric' time.

Television has become the dominant medium of the last 50 years. Television, and its earlier cousin cinematography, has a mesmerising effect on people, illustrated by the apocryphal tale of the cinema audience who evacuated the building when they saw a film of a train coming towards them. "The motion-picture medium has an extraordinary range of expression. It has in common with the plastic arts the fact that it is a visual composition projected on a two-dimensional surface; with dance, that it can deal in the arrangement of movement; with theatre, that it can create a dramatic intensity of events; with music, that it can compose in the rhythms and phrases of time and can be attended by song and instrument; with poetry, that it can juxtapose images; with literature generally, that it can encompass in its sound track the abstraction available to language." [25]

The narrative power of television and its stimulation of the two most highly developed senses - sight and hearing - are what make it a powerful medium. We have now reached the point in our discussion - roughly 1960 - when a new medium of protean capability began to emerge that has the potential of becoming even more significant for human communication and culture.

The origins of the modern computer interface

"The networked computer acts like a telephone in offering one-to-one real-time communications, like a television in broadcasting moving pictures, like an auditorium in bringing groups together for lectures and discussion, like a library in offering vast amounts of textual information for reference, like a museum in its ordered presentation of visual information, like a billboard, a radio, a gameboard, and even like a manuscript in its revival of scrolling text. All the major representational formats of the previous five thousand years of human history have now been translated into digital form." [26]

Students of the history of interface design will be familiar with some of the icons of invention in this field. Different strands of thought and research are drawn into today's discourse on the discipline. The media and communications strand has McLuhan as its patron saint, while the technology strand has Doug Engelbart as its hero. Alan Kay was perhaps the first person to weave them into a coherent body of knowledge, and it is not a co-incidence that Kay was the principal designer of the graphical user interface - seeing the computer not as a tool (as the technologists had) but as a medium. We also have the psychologists chipping in, beginning with Donald Norman and the cognitive ergonomics community.

Some of the characters are perhaps less well known, for their inventions or ideas gathered dust through a lack of finance - and perhaps imagination - necessary to turn them into commercial reality. But the common thread tying them together was a focus on the way

[25] Dern, Meyer, *Understanding Movies*, 2nd edition (Prentice Hall, New Jersey, 1976) p XI
[26] Murray, ibid, p 27

computers are designed to by used by people, rather than on the way computers worked. The mainstream in computing wanted to build ever bigger and more powerful machines, perpetuating the need for a select group of skilled professionals to use them. But technological advances such as transistors and integrated circuits reduced the size, and hence the cost, of computers by packing in the basic switching mechanisms into miniature packages. This made it possible to conceive of a computer designed for an individual user. The pioneers and mavericks on the fringe, as Howard Rheingold points out, built the first personal computers because they wanted mind amplifiers for their own use. This was the beginning, to borrow Brenda Laurel's phrase, of the art of human-computer interface design.

John Walker, founder of Autodesk, wrote a ground-breaking internal memo called Through the Looking Glass, in 1988 in which he identified six user interface generations.[27]

Generation	**Description**	**Barrier**
First	Plug boards, dedicated set-up	Front panel
Second	Punch card batch	Countertop
Third	Teletype timesharing	Terminal
Fourth	Menu systems	Menu hierarchy
Fifth	Graphical controls, windows	Screen
Sixth	Cyberspace	?

The 1960s saw the beginnings of the shift from the second to the third generation of interface. There was a sudden spawning of experimentation with human-computer interface technology in research laboratories and government institutions, mainly in the USA.

Probably the most celebrated innovator of this period was Douglas Engelbart. Engelbart's vision was to "use audio-visual media to match human perceptual and cognitive capabilities with computers' representational and computational capabilities….to increase the power of our most important innate tools for dealing with the world - our ability to perceive, think, analyze, reason, communicate."[28] In 1957 Engelbart to took a job at the Science Research Institute at Stanford University where he worked in isolation on his ideas, and eventually published them in 1963 under the title "A conceptual framework for augmenting man's intellect". With funding from NASA he went on to found the Augmentation Research Centre at SRI, which was responsible for all the major inventions of what came to be known as the 'direct manipulation' of computers. These inventions included word processing, the mouse, hypertext, windows, icons, bitmapped video graphics, online processing and multimedia documents. In 1968, Engelbart demonstrated his radical view of operating computers to the Fall Joint Computer Conference and the world of computing would never be the same again.

A fellow traveller on this lonely quest was JCR Licklider. In 1960, Licklider was a professor at MIT and wrote a book entitled Man-Computer Symbiosis, in which he conceptualised a machine that would help man's thinking in real-time, to take over the mechanical tasks and leave the thinking to human beings. His vision closely matched Engelbart's. Licklider advocated time-sharing computing that increased the community of users of any one machine, as opposed to the conventional batch computing of the time. This enabled users to think while sitting at the terminal, rather than just type in pre-digested procedures. This was the beginnings of the idea of time-sharing and interactive computing. Licklider went on to

[27] Walker, John, *Through the looking glass*, in Laurel, op cit
[28] Rheingold, Howard Rheingold, *Virtual Reality* (Secker and Warburg, London, 1991) page 49

become director of ARPA, which is where Alan Kay and his team cut their teeth before moving on to PARC.

Two years later, Ivan Sutherland, working at the Lincoln Laboratory, demonstrated another ground-breaking concept technology which emerged from work done in computer graphics. Sketchpad, his programme that used a light pen to draw images on a computer screen, fundamentally changed the way people used a computer, by enabling them to use their hand and eyes in real-time, and was the forerunner to Computer Aided Design. "Sketchpad was Lascaux, thirty thousand yeas later: instead of pigments on limestone cave walls, Sutherland used electrons and glowing phosphors on the surface of a glass bottle." [29]

A witness to both the Engelbart demonstration and an early demonstration of Sketchpad was Alan Kay. Kay was a researcher at ARPA with an interest in media and psychology, heavily influenced by Marshall McLuhan and Seymour Papert. Kay's conceptual breakthrough was to see the computer as a medium rather than a tool, and while at ARPA and later at PARC pioneered the graphical user interface and the first personal computer, the Alto. The Alto, and its descendants the Star, Lisa and the Macintosh, finally brought together the innovations of Licklider, Engelbart and Sutherland into one integrated computer system and led to the eventual dominance of the Windows interface in personal computing.

Kay and his PARC researchers were the first to apply the concepts of metaphor and icon to interface design. In 1972 he demonstrated a painter's palette, which enabled the user to 'paint' pictures on the computer screen. "This was just the start. They did not want simply to create metaphorical tools, they wanted to create a whole metaphorical world, one in which the user would explore to discover what the computer could do." [30]

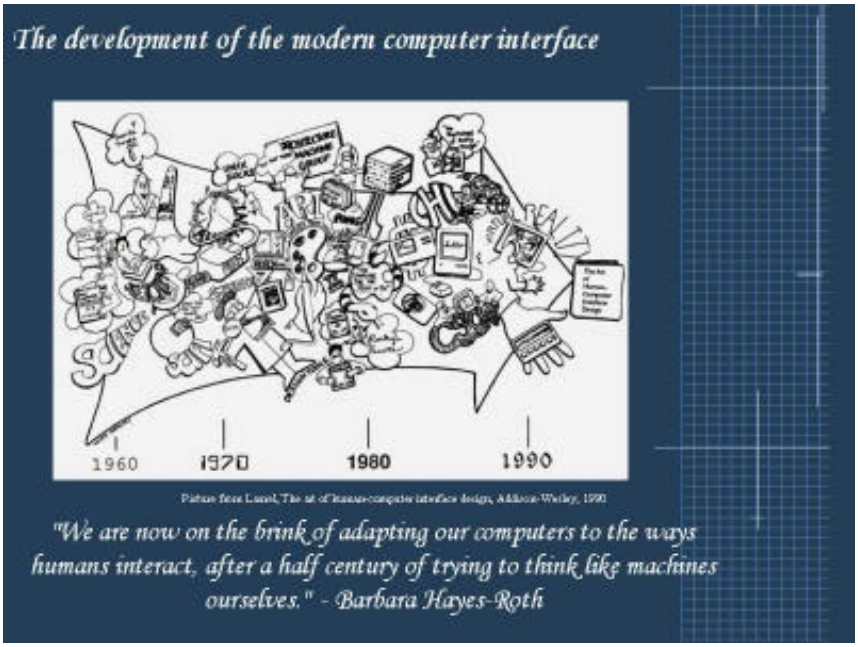
Kay remains one of the leading theorists in this field. In his paper "User interface: a personal view" - now admittedly ten years old by still pertinent - he points out some of the challenges for user interface designers. He divides them into two categories, manipulation and management, based on man's tool-building and social-oriented nature. The key challenge for designers of better manipulation tools is encapsulating complex symbolic language into readily understandable iconic representations, with the goal of enabling ordinary people to use computers to both read *and* to write in the computer medium, in other words, be computer literate. Within the management category, he advocates the ability to indirectly manage agents rather than directly manipulate objects.

In Kay's terms, we communicate with ourselves, our tools, our colleagues and others, and our agents. Until the mid 1990s, the computer was used to augment our communication with the first two. In 1995, the advent of the World Wide Web and Web browser, and the widespread adoption of Internet Protocol as a communications standard, has seen explosive growth in the third form of communication. Though not often recognised as contributing to the human-computer interface, the network and a common user interface to access it are, in fact, vital technologies for enabling the cyber equivalent of the community, an essential component of human society. It is Kay's contention that the agent will become the means by which humans will eventually mediate between themselves and the network.

In parallel with applications of pictorial representation to interface design was the development of a new way of handling text and, ultimately, all screen-based objects. This was the idea of hypertext, that originated with Vannevar Bush's Memex, was articulated clearly in 1960 by Theodore Nelson, but is today most closely associated the World Wide

[29] Rheingold, ibid, p 89
[30] Woolley, op cit, p 147

Web. As we have seen, narrative had been the dominant means of communicating for thousands of years, represented by the cultural artefacts of books, paintings, plays, films, V, buildings. The conceptual premise behind hypertext is the arrangement of information in non-linear form. Most importantly, it changed the relationship between the author and his audience. From the non-negotiable pact between author-producer and reader-consumer implicit in narrative, interactive discourse gives power to the user to influence the outcome of the experience.



Virtual Reality, gaming and the Internet

We have traced the development of the human-computer interface as is manifested today in the graphical user interface (GUI), the keyboard and the mouse. But concurrent with these developments was a view of human-computer interaction that wanted to expand Licklider's notion of 'intimate contact' between human and computer to the point where a person would actually reside within the computer program. It also drew fundamentally on the idea of information space as being non-linear. Ivan Sutherland can be credited with the first VR prototype in 1968, but the real work went on at the University of Utah in Salt Lake City where in 1970 the first working stereoscopic head-mounted display was demonstrated.

The early 1990s saw a feverish interest in VR and the associated concept of Cyberspace, a term first coined in 1994 by novelist William Gibson in *Neuromancer*. Looking back over developments of the last ten years, one can see that Gibson's fictional vision of cyberspace, now manifest in the Internet, arrived as the common way for people to interact with information before the full-body immersion experience of VR. It might not have been so, had Hollywood seen the promise of a still-earlier technology involving full-body immersion.

Morton Helig's Sensorama was perhaps the first device that attempted to simulate the human sensorium, in this case for the purpose of entertainment. It was The Lady and the Unicorn in techno-drag: "It is an object of the present invention to provide an apparatus to simulate a desired experience in the senses of an individual."[31] Howard Rheingold describes the experience of a simulated motorcycle ride in Brooklyn in the 1950s: "I heard the engine

[31] Helig, Morton, *Sensorama Simulator*, US Patent # 3,050,870, 1962, Rheingold, op cit, p 50

start. I felt a growing vibration through the handlebar, and the 3D view photo that filled much of my view came alive, animating into a yellowed, scratchy, but still effective 3D motion picture."[32] The idea of a new entertainment medium involving hyper-reality was before its time, and its failure can be attributed to many of the same factors that caused the stalling of VR as a new medium.

Though VR is applied in some real-world applications, such as architectural, scientific and business visualisation and computer games, certain characteristics of the VR interface have made it unsuitable to mass adoption as a human-computer interface. Principal among these is the encumbering nature of the technology itself, which makes it difficult to use and expensive. Many practitioners of new media technologies, such as Nicholas Negroponte at the MIT Media Lab, are sceptical of the value of VR for everyday use in the office, home or factory, for these very reasons. Interface technologies that will lessen the physical burden of experiencing VR, such as miniature screens that can be worn as contact lenses, and biosofts, or electronic implants, are still some way from commercial application.

Far more successful than immersive VR have been their cut-down siblings, computer and video games. As media types, their success derives not only from the visual realism of the imagery (Pacman, one of the first killer games, uses blocky characters and one colour) but from the illusory sensation provided by immediate interaction and the "suspension of disbelief". Gaming is a form of human behaviour that is as old as humans themselves, and the popularity of computer games was arguably the most important driver behind the development of multimedia and 3-D graphics which form the cutting edge of digital imaging technology today.

Interactive television and mobile interfaces

The current obsession with the Internet is soon to be substituted by the new kids on the block, interactive television and mobile telecommunications. Between them, these three technologies will provide interface designers with enough challenges to keep them occupied for at least the next five years.

Interactive TV (ITV) has been in gestation for many years. An early, non-broadcast, version of it was interactive video, comprising a PC connected to a VCR or laser disc player. Various forms of interactive video have come and gone, perhaps the highest profile being CD-I. CD-I failed for a number of reasons, but its main problem was that it was neither a TV nor a computer - it fell between two stools. From an HCI point of view, its interface was designed for use at-a-distance (remote control, as in TV), but the content and low resolution display were not sufficiently compelling to motivate users to use its interactive features: there was not sufficient value added over and above the narrative programming contained on the discs.

Truly interactive TV systems work when they facilitate the networked, interactive capabilities of the Internet to combine seamlessly with the narrative broadcast features of television. Anyone familiar with the DSTV or Open TV interfaces will immediately revel in their simplicity and value-added features. The rich content available to browsers, such as home shopping, on-line games, weather updates and news, make the investment and short learning curve worthwhile. I was initially sceptical of sitting back in my couch and using the remote where I had become used to sitting at my desk to perform the same functions. But I found it easy to use and intuitive, as have millions of others.

[32] Rheingold, op cit, p 50

Combining the Internet with cellular telephony promises to unleash yet another wave of new ways for people to interact with information. Early renditions of WAP interfaces are clumsy, and the cell phone screen was not designed to display more than a few lines of text and low resolution icons. These features will force application and interface designers to select the services they offer very carefully. Those that require instant access to information and require little content depth or analysis will be most successful. One can imagine lastminute.com adapting much of its Web content to WAP, as well as financial institutions, on-line gaming companies and ticketing companies.

Beyond the desktop

The dominant metaphor for interacting with computers for fifteen years has been the desktop. It has its virulent critics and its passionate supporters. Most sit somewhere in the middle, accepting its limitations but bewailing the absence of better alternatives. One such critic is Donald Norman: "The desktop worked very well on the old 128K Mac with no hard disk. You could scatter everything around and still see it on the screen. But today we have thousands and thousands of items - more than can possibly fit on the screen."[33]

Steve Steinberg categorises alternatives to the desktop into four: spatial, semantic, networked and temporal.[34] The desktop and the idea of cyberspace are spatial metaphors. Where the desktop deals with space in two dimensions, cyberspace adds the third dimension of depth. One attempt to expand on the spatial theme is Negroponte's Spatial Data Management System, conceived in 1976, but it never left the drawing board. The main objective of spatial interfaces is to represent the huge amount of information available to users in a "usable" format. Stuart Card and others at Xerox PARC have developed a set of new interactive visualizations for hierarchical information (cone tree, Hyperbolic Tree, Spiral Calendar, Disk Tree), linear information (Perspective Wall), matrix information (Table Lens), information space x time (Time Tube), and document visualization (Document Lens). Their most recent research focuses on Information Workspaces (Rooms, Information Visualizer). One group from PARC has co-started a new company, Inxight, to market the results of this research, but its products have so far hardly made a dent commercially. (From XEROX PARC's User Interface Research Group website)

Semantic interfaces are broadly based on the functionality of search engines such as Yahoo and AltaVista, an example being MIT's Semantic File System. Networked interfaces are best illustrated by Theodore Nelson's Xanadu project, but even with the powerful support of Autodesk funding it never saw commercial success.

Little attention has been paid to the development of a temporal metaphor, where time rather than space is the conceptual framework for designing the interface. One attempt to do so is Lifestreams, developed by computer scientist David Gelernter. Lifestreams works on the belief that people create and use information over time, and can best develop, store and manipulate it when thinking chronologically. It is more than just a file system with time and date information, "it's an entire architecture that defines the structure of our lifestreams, how they are stored, and the operations that can be performed on them." [35]

Perhaps the final word on the subject is appropriately left to Hollywood, heirs to the cave painter magician tradition. The latest attempt to redefine human-computer interaction evolved from the film The Matrix, and CNN.com reports that the US military has enlisted the

[33] Donald Norman, in Wired, 5.02 Feb 1997
[34] Wired, ibid
[35] Wired, op cit

support of the film's technical team as advisors.[36] The idea is to recreate the Holodeck from the film that allowed Keanu Reeves to dodge bullets, and build a simulator for training solders in virtually any environment. In 1999 the Defense Modeling and Simulation Office and Paramount Digital Entertainment began working on using Hollywood multimedia technology and movie storytelling skills for training soldiers to make better decisions on the battlefield. Also involved in the venture is the University of Southern California which has teamed up with the Army to establish the Institute of Creative Technologies "The holodeck is the Holy Grail of the institute," Paul Debevec, a member of the team said. "It will be a next-generation virtual reality simulation technique that will make it possible for a person to go into a room or put on a headset and really feel like they are in a different place. They will be able to see, hear, touch and even smell everything. The terrain or environment will be realistic, and eventually there will even be other characters to interact with and teach and learn from." [37]

"If you work from photos of real environments, you can get 95 percent realism, but that doesn't include people or dynamics," George Borshukov, a colleague of Debevec said . "Capturing people doing real things on film and stringing that together with real environments [will be done]," he said. But "the ultimate goal to simulate all the physics, as opposed to simply image-based rendering, is 20 to 25 years away."

It seems that man's quest for representing, augmenting and interacting with his world by artificial means will carry on in many new guises for as long as we remain in control of our senses.

[36] CNN.com, May 2nd, 2000
[37] CNN.com, ibid