Designing a New Medium

from
Brenda Laurel (ed.):
The art of human.computer interface design.
Reading, MA.: Assison-Wesley 1990

Tim Oren
Advanced Technology Group
Apple Computer, Inc.

THOUGH CURRENT COMPUTER-ASSISTED "multimedia" systems are clumsy and expensive, progress at the hardware level is accelerating. What has not kept pace is our understanding of software and content design for the multimedia environment. Particularly, there has been little examination of the implications for computer-human interaction to be found in the emerging use of the computer as a medium.

Computers are usually viewed as tools, instruments for storing and manipulating data that ultimately will be printed out. But now the computer is beginning to be used as a medium in itself, a means of communication in which the content is never reduced to print. Examples include multimedia education systems, collaborative computing environments, and desktop presentations. The computing medium has both real-time and stored forms and is multimodal, incorporating text, voice, music, graphics, video, and animation.

We should now begin thinking of the computer-human interface as a media process. This means enlarging our study to include issues such as the psychology of media, evolution of genre and form, and societal implications of media biases—issues heretofore peripheral to computing. The shift of focus to the media level may give us a new language in which to describe such choices and to analyze the prior evolution of computer systems. It suggests that attention to artists employing the computing medium,

and the incipient genres they will originate, is a source of inspiration and understanding for user interface and computer systems designers.

Taking Control of Media

The forms of existing media, such as television, have been largely determined by opportunism. The first technically feasible configuration is rushed to market and exploited and then sets an *ad hoc* standard for form and content in that medium. Television shows are certainly designed, but they are set in a framework of hardware and marketplace over which the writer has no control. Altering a medium based on obsolete hardware is extremely difficult, as the current high-definition television (HDTV) controversy shows.

In contrast, the computer medium is malleable and tunable—"soft." As hardware improves, the face that the computing medium presents to the user and author is increasingly determined by the design and limits of software. Because software is more easily manipulated than hardware, the computing medium is craftable, a metamedium. Its existence implies, for the first time, a design space that may be *deliberately* explored.

The wrong way to deal with this situation is to muddle about until we hit gold, and then let the marketers take over. The current status of interactive media is a prime counterexample, with its proliferation of specifications: CD-ROM, CD-I, DVI, CD-V, LV-ROM and so on and on. All of these were motivated by market positions that might be obtained by leveraging current hardware products. None of them involved prior attention to what content and forms might make interactivity an attractive option, or to the biasing characteristics of the medium that would result.

The choices designers make, or the decision to leave the matter to the vagaries of the marketplace, will have far-reaching consequences. Our society increasingly communicates its knowledge and values and conducts its affairs through media, rather than face to face. Each medium has its own set of biases and proclivities that alter and govern the form of information it transmits. The growth of institutions around a medium is likewise influenced by its biases. For instance, the effects of network television on public debate and education are widely perceived [Mander, 1978].

The computing media described in this book and elsewhere [Ambron, 1988] are proposed as vehicles for collaborative work, learning, and play and particularly for the education of the young. Media biases in such systems will have a powerful effect on our culture, as they govern what may be discussed, created, and transmitted now and from generation to generation.

If we as designers want to take control of the process, there must be a vocabulary in which to articulate choices. At present, much of the computer-human interface field is committed to the detailed study and modeling

of low-level cognitive phenomena [Newell, 1985], an inappropriate level at which to address media issues. We should instead consider whether the language of media and criticism is a useful next stage to our understanding of systems of humans and machines.

The Evolution of Media

in-cu-nab-u-lum [fr. L *incunabula*, pl., swaddling clothes, cradle, fr. *in-* + *cunae* cradle] **1**: a book printed before 1501 **2**: a work of art or of industry of an early period

What can we learn about the new computing medium by looking at the past? Every medium goes through an early incunabular stage in which old forms persist into the new medium before being modified and finally replaced with new, better adapted, forms. For example, the text shown in figure 1 replicates in early movable type the interlineation and marginalia typical of a hand-written manuscript. With its detailed cross-references among six parallel streams of text, it resembles the notion of hypertext more than it does a modern book.

What happened to this form? It was driven out in part by the difficulty of typesetting such pages. With early printing technology, each character had to be placed individually on the page and replaced in the type fount after use. New forms of content, such as outlines, tables of contents, alphabetical indexes, and footnotes, arose to fill similar functions while simplifying the layout problem and reducing the number of typefaces needed. These forms are now so stamped in our consciousness as to be thought essential for scholarly literature.

In this early day of computer media we are also using incunabular forms. We are busily replicating hierarchically organized books on CD-ROM and distributing libraries of still photos on videodisc. To get beyond this stage and exploit the computing medium, we will have to invent forms and genres that are more than borrowings. Thomas Edison created the movie camera, but D. W. Griffith and Sergei Eisenstein invented filmmaking. Likewise, the advances of the new medium will be defined by seminal visions of those who are expressing themselves in ways heretofore impossible. This is the key to success of the computer medium, not digital video, broadband fiber networks, or some other hardware advance.

The problem is harder now, however. The limits of the film and cameras of the time set essential parameters for Griffith and Eisenstein. In the computer, above the essential hardware substrate, both form (user interface) and content (data) may be implemented in software. For instance, a HyperCard stack stores both interface and content in the same way. Not only is the form defined in software and therefore malleable, but the boundary between form and content is itself plastic. Without some guiding

Figure 1 : Image of incunabular Biblical commentary. Text and two parallel commentaries, interlinked with numeric references.

notion of conventional form, the viewer may have trouble distinguishing the two. Indeed, the traditional separation of form and content is called into question by the computing medium.

Defining the Design Space

What are the conceptual dimensions of the new design space? Our experience is too limited for any conclusive answer, but the following will be

considered below: What is the human role in the range from active to passive—is it user, viewer, or participant? To what extent do the author or editor and final viewer of content share the same environment? Are the means of creation widely distributed or centralized, costly or inexpensive? Is the medium suited for a mass or an elite audience?

This design space is so large that a perception that all computer-based media (CD-I and DVI, for instance) must compete with one another is superficial. Rather, the *forms* built of these technologies will compete amongst themselves and with older forms for the time and attention of potential authors and audiences. The expressive range of the forms, the types of messages that can be sent, and their potential to engage the human participant, will be the grounds of comparison. The psychology of the media experience must be investigated if we are to have a notion of *a priori* design for engagement [Laurel, 1986a].

It should also be obvious that in such a space there is unlikely to be "one true user interface." A stubborn adherence to known methods would cripple the potential of new media by restricting expressive range. The knowledge won in the Desktop setting on the Macintosh now has to be tested against a wider set of choices, and we must look for overarching principles that can explain both the Desktop and the multimedia experience.

Genre

Genres are conventional, familiar ways of setting expectations of the experience to come. All works in a given genre have certain underlying ideas and themes in common. One does not expect to pick up a romance novel and have it suddenly turn into a police procedural, nor to have a soap opera change into a variety show. Genre recognition invokes our memories of conventional stories, characters, and handling of form and leaves us free to enjoy the nuances of the story of the moment, rather than relearning basics. In order to gain the advantages of a particular genre, an artist must work within its boundaries. Artists may choose to break the boundaries of a recognized genre for several reasons. They may wish to comment on the conventions of a form by satirizing or burlesquing it, for instance, or they may deliberately violate the established conventions of a genre for the sake of innovation. Whenever artists depart from a known genre, they risk losing the benefits of genre recognition. But successful innovation in a particular work can feed the genre and cause it to evolve.

As older, fixed media have found a necessity for genre, so we may suspect that the new flexible computing medium will have a greater need. With the range of design available, there must be a means of setting expectations and transferring knowledge from past exposures, if the content itself is to be appreciated.

Because of the role of genre in setting expectation, borrowings into the computing medium from existing media must be approached with caution.

For instance, a value of computers is interaction. Use of an established noninteractive genre for the sake of familiarity may sabotage this value by lulling the user into old, established habits, which do not include an active user role (see the chapter by Oren et al. in this volume). One also risks comparison with the highly evolved and intense production values employed by established media, such as print and video.

The notion of genre has already arrived on the computer, under a different name: *metaphor*. *Metaphor* as applied to the interface is a different sense of the word than its literary use. In the literary definition of metaphor, our interfaces are hopelessly mixed: We place *windows* on our *desktop*, then put *folders* within the *windows*, thus forming a *tree*. What an odd collection of natural, architectural, and office imagery! And yet it doesn't seem to matter. If we instead think of the so-called Desktop metaphor as a *genre*, a developed set of expectations for content and form, then it seems reasonable, because a genre can embrace many metaphors. And indeed, the Desktop's appeal to skill transfer from program to program is solving the same problem as the evolution of genre. It at once sets user expectations and self-enforced limits of expression.

Because of the familiarity and strength of its underlying ideas, the Desktop has burgeoned into a genre with many individual expressions. Other genres of the computing medium are discussed in the trade periodicals daily: spreadsheets, WYSIWYG text editors, draw and paint programs. Each has one breakthrough ancestor that established it in the public's eye, and a succession of new, innovating programs that influence and refine the underlying genre. Try thinking of Bill Atkinson as Edgar Allen Poe, and MacPaint as "The Murders in the Rue Morgue."

Convention and Phrasing

Existing media have established conventions that punctuate the viewer's experience. For instance, an establishing shot is conventionally used in film before a close-up is done. A cut from one head shot to another facing the opposite way indicates conversation. If creepy music starts playing just before the hero opens a door, you can be sure he's in for a nasty surprise. Like genre, media conventions set expectations and allow a focus on content, but they do it on a different scale. Media conventions cut across genre lines; for instance, the filmic conventions mentioned above are employed in all genres of movies, from horror to comedy to adventure.

The punctuation of experience is even more important in computer media systems. For instance, the perceptual jump between a color video and black-and-white computer text can be abrupt. To ease such jumps, we will need segue conventions that set up expectations for the transition. We will also need standard ways for setting context in nonlinear media. The common occurrence of interface features such as maps, backtrails, paths,

and tours in current multimedia prototypes may indicate that they are becoming recognized conventional uses that address this problem.

For conventions in new media to have value, much of the potential audience must be "media literate," conscious of the media conventions. This process has already begun. A videotape control panel would not have been recognized by most people ten years ago but is now the prime means that most consumers have of dealing with "interactive" sound and video.

Most computer interface ideas can be recognized as directed at this simple, media-wide level of convention. Heretofore, simplicity and even "monotony" have been recognized as important characteristics of interface [Raskin, 1986]. Experience from more mature media suggests that a progressive unfolding of depths of sophistication is a more likely final outcome. The surface story of film or novel is accessible to all, but a lifetime may be spent studying their forms and gaining a richer understanding and appreciation of the works. Future computer systems may have simple surface conventions and many layers of sophistication for those willing to learn.

Conversational Systems

The ruling paradigm of the Macintosh interface is the combination of a passive, tool-like computer with an active human. The machine is to be nonmodal and reactive, and should intrude as little as possible on the task at hand. A traditional medium like television is in sharp contrast. The human is the passive observer of the content, and the machine is the active element.

Between these poles may be options for new media that are *participatory*; that is, where all actors, both machine(s) and human(s) may take active or passive roles at various times. We will need conventions for the exchange of initiative between the actors in the system. Such conventions need to be modeless; in other words, it should be possible for any actor to ask to "take the floor" at any time. Let us call a participatory system with conventions for such exchange of initiative a *conversational system*, taking human conversation as a metaphor (though not implying the necessary use of human language).

Conversational systems are needed because the knowledge of the human is incomplete. For instance, the novice cannot begin to use a system unless there is some volunteering of information on operation. A student cannot use an educational multimedia work effectively if there are no starting points, no examples or concepts that are offered spontaneously. Also, the sheer quantity of information available in present and predicted systems overwhelms attempts to keep it mentally organized, and the computer, acting as agent, must volunteer the information and organization as needed, in response to the human's explicit or inferred needs (see Laurel's chapter

in this volume). In computer games, the entire experience is the interaction between the player and the behavior of the program.

A comparison of written and spoken language shows the enormous impact that conversational systems will have on form. In published media, the author produces a finished work, which relies on culture and knowledge that the eventual reader is assumed to share in order to ground the discussion. Face-to-face communication is an iterative process in which meaning is negotiated, with the ability to digress to explain a cryptic referent if understanding breaks down (see Brennan's chapter in this volume). Anyone who has tried to turn a recording of a conversation into a published article has experienced the difference in forms firsthand.

This use of spoken and written language as an example does not imply that a conversational medium must use human language. In a system that includes computers as well as humans, the methods evolved for human-human communication may or may not be appropriate for any particular situation. The "utterances" of computer-human conversations may take many nonverbal forms (see Buxton's and Schmandt's chapters in this volume). But whatever the type of utterance, a conversational medium should be capable of flexible exchange of initiative, two-way transfer of information, and maintenance of context through the analysis of discourse.

What are the current prospects for building conversational abilities into computer media? The ability to mock a conversational entity is the so-called Turing Test, and this task falls within the bounds of artificial intelligence (AI). As yet, AI has failed to produce a system for discourse understanding that is at once general and robust. Formal symbolic methods suffer from brittleness when they venture outside constrained domains. Attempts to improve them with "common sense" suffer from the so-called knowledge acquisition bottleneck: the need for human intelligence to encode knowledge into a prescribed format and to ensure consistency with the remainder of the system.

There are other, nonsymbolic, approaches to giving the computer the ability to act independently. Connectionist methods show some promise for spontaneous generalization, but their use in large, unconstrained domains is unproven and their workings are often unexplainable to the user.

The methods of information retrieval in unstructured text deserve notice. Information retrieval attempts to break the knowledge acquisition bottleneck by automatic indexing and searching of text. But there are no "reliable" retrieval methods; all retrieval is inherently probabilistic as a result of the imprecise use of words by authors and readers and the computer's inability to fully understand context.

The prospect, then, is that the intelligence of the human participants in a conversational medium will dominate that of the computer for the foreseeable future. The advantages of the machine are an eidetic memory,

infinite patience, and its role as the mediator of all interactions. It is able to watch as you work and communicate. But a robust portrayal of a conversational entity that can operate across contexts is beyond us at this time.

Storytelling

> A story is a little knot or complex of that species of connectedness called relevance.—Gregory Bateson [Bateson, 1980]

One way around the current lack of generalized conversational systems is to consider subclasses of conversation. For instance, some collaborative work systems use the formalized conversation of business [Winograd, 1985]. Here I will consider the possibility that storytelling, viewed as a constrained conversation, may be a tractable approach, with the potential to engage the human participant at deep symbolic levels as well as explicit cognitive levels.

Like conversation, a storytelling metaphor departs from a responsive tool-like model for the computer. But rather than free interchange of initiative, a storyteller holds the content and structures its presentation to achieve a dramatic effect, partially in response to the spoken and unspoken reactions of the audience at hand. Thus, it has some conversational aspects. This can be contrasted with pure drama, where the form is supplied by the artist, or again with real or simulated worlds, where we follow a self-determined path through space and time and form and meaning are self-supplied.

The so-called "navigation" of databases is similar to this wandering of a real world [Oren, 1987], because the items are viewed in a linear sequence. Interpretation is up to the user: if we are successful in casting the items in terms of our own experience, we may assimilate the information contained [Mandler, 1984]. The role of a storyteller in such a system could be to arrange the narratization of content into useful patterns, by constraining and guiding the choice of what to read or view. This is one origin of the "what's next?" question that led to the formulation of guides (see the chapter by Oren et al. in this volume).

Results from the guides experiments suggest that people find the addition of storytelling and personification to the interface to be intriguing and engaging, particularly in comparison to third-person, omniscient voice text and narration. There seems to be a desire for more carefully defined characters and associated points of view within the database. We are led to suspect that the further extension of storytelling method into new media is worth examining.

Storytelling has the convention of describing only what deviates from the expected and is significant to the advancement of the plot. This dovetails

with the finding that humans do not encode, recognize, or recall something that is *expected* in a context [Mandler, 1984]. Mentioning any low-level detail raises its perceived importance to the story (see Don's chapter in this volume). The convention is so strong that the mention of a seemingly unrelated item in the course of a story may instigate a search for its relevance. This may be the psychological basis of foreshadowing (and the entire mystery genre), and is precisely the user behavior desired with guides: to produce a suspension of belief such that plausible connections will be investigated before the narrative explanation is discarded.

The goal of the story is to evoke the user's engagement at both conscious and unconscious levels. Given the limits of intelligence in the computer, we must create a storyteller that can "tell more than it says." The medium becomes as much a mirror to the user as a window on the data. Myron Krueger suggests that: "Ambiguity is an instrument of efficient communication, for while you may not have succeeded in saying one thing clearly, you have suggested several ideas at the same time" [Krueger, 1983].

The imprecision enforced by our poor tools may work in our favor in the storytelling mode, if it is not so dissonant as to break suspension of disbelief.

Our present understanding of creating and portraying story elements in computing media is crude. But there are intriguing questions here: can we heighten engagement by moving from generic characters to figures who are very individualized, even flawed, or which develop with time? How can we represent elements such as plot, tension, and catharsis in the computer, and is their usage in the new media appropriate and helpful [Laurel, 1986b]? How can these elements be maintained while using the human participant as a source of unpredictability? How can interaction best be introduced to the storytelling process: branching, point-of-view shift, some form of interest feedback, or some extension of next-move generation? The best guidance available may be from studies of previous changes in media [Ong, 1982] and criticism of innovative literary forms [Lem, 1984; LeGuin, 1989].

In introducing story to the computer under the guise of new media, we move from the symbolic logic of AI to the class of subconscious symbol described by Jung and Campbell [Jung, 1968; Campbell, 1988]. We begin to explore whether the computer can transmit the "affective image" that speaks in archetypal terms. Examination of the computer human interface in these terms is unprecedented. The best parallels are in the psychological and critical analysis of myth and film. This is not unfitting, for the computer is one of the greatest artifacts of power of our time, an embodiment of the creation archetype for those who can wield it. The public fascination with the computer as Frankenstein and its power to attract young hackers have

a mythic quality. We have spent a great deal of effort on the Appolonian side of the computer; perhaps the advent of new media will usefully engage the Dionysian aspects as well.

Media and Markets

> Convivial tools are those which give each person who uses them the greatest opportunity to enrich the environment with the fruits of his or her vision.—Ivan Illich [Illich, 1973]

Existing commercial media, such as television, radio, and most print and computer software forms, are mainly mass media. They are characterized by a marketplace with centralized production and distributed consumption. The ability to have a voice in a mass medium is regulated by one's access to the costly, centralized means of production.

Let us borrow *convivial* from Illich to describe other media in which all participants have the possibility to be authors or readers. Examples include the telephone and many electronic messaging and bulletin board systems. If we believed conviviality to be a desirable property of new media, how would we go about designing for it?

An essential principle is symmetry—that the same authoring tools be available to all. In the mass video medium, the networks have megadollar studios, whereas the consumer may have a single camera. In electronic messaging systems, a symmetric medium, everyone writes using the same tool. In a computing medium, the question is whether a potential artist needs a quarter-million-dollar budget to even begin. Symmetric media produce a talent-limited creative process; in mass media, the process is economically limited. We see that tools make the medium as much as its delivery vehicle.

We should also prefer "low thresholds" of accessibility to "high ceilings" of bandwidth and modality. Low thresholds let in authors with minimal resources. It is possible to "sketch" without a great commitment of time and expense. High ceilings yield a medium that is capable of absorbing a great deal of production value, which correlates to production costs. High production costs lead to the need for a mass market to amortize the expense, or for an audience able to subsidize production costs. We end up with "Three's Company"—the subjection of content to titillation, or with Broadway shows—content for the affluent few.

A medium may trade off production value against *salience* in gaining the attention of a user. *Salience* is the pertinence of a piece of information to a particular person's needs at the time it is presented. For instance, targeted direct mail advertising is an attempt to increase the salience of the message, as opposed to the intense production values that characterize

television commercials. If, through devices such as agents, we are able to create personalized adaptive media that optimize salience, we may be able to compete with the production values of mass or elite media.

It may be useful for a convivial medium to have the property of collage-ability, that is, the ability to create new works by combining old ones from the same or other media. This admits the possibility that a legitimate expression may consist of the meaningful juxtaposition of existing works. It opens up the medium to those who are not primary creators in some of its forms. Collage-ability requires gateways, the technical means to bring information into the medium. Video digitizers, scanners, and optical character recognizers are examples of gateways. Once in, the medium should not erect artificial boundaries to the joint use of information from various sources. Collage-ability is desirable for convivial media, but raises many legal and ethical issues, such as the status of copyright and royalty in collage works, and the ability of artists to control the context of use of their work.

To promote understanding, propagation, and improvement of new forms, a convivial medium should be inspectable. It should be possible to "look under the hood" and see how an effect was achieved. This alone is not enough—though film and video are mostly in the open and literal, the tools are not yet accessible. But inspectable software is increasingly a thing of the past, with the replacement of hacker sensibilities by commercial interests. This trend should be reexamined if we are to move to new media where the value of content outweighs any transient advantage of programming.

One final aspect at which computing has been notoriously poor is the survivability of content. By *survivability* I mean the possibility to move the bulk of content and structure from one system to another. Music moves easily from vinyl to tape to CD, video from tape to disc to digital format. But software seldom goes anywhere at all, and the problems of moving a file system from an obsolete computer to a new one are notorious. If artists and commercial organizations are to make investments in new media, this must stop.

It is interesting to note that Apple's HyperCard has some characteristics of a convivial medium. It is not convivial in a true sense, because access requires purchase of a rather expensive computer system. However, for the community of Macintosh owners, it is symmetric, because everyone gets the same version. It emphasizes easy access over high production items such as color and digital video. It displays collage-ability of sound, digitized images, and scanned or typed text. Scripts are inspectable, and the sale and swapping of "buttons" is a recognizable phenomenon. The various updates have all been able to accept content from previous versions. Insofar as HyperCard has succeeded

as a "Volksmedium," it is because it shares these aspects of convivial media.

Technologists Should Be Listening to Artists

Building a new medium is a talent-limited process because we are early in the evolution of genre, searching for seminal visions and rallying points. As with Edison and film, it is unlikely that the inventors of the technology will be the creators of such visions. This is more likely to fall to the authors, educators, and artists who have something to say and who find that the new medium offers them a way to say it for the first time. Early forms of the computer medium, such as HyperCard, are significant in providing this expressive opportunity to many.

Because these original uses and forms were not envisioned during the design of existing computing systems and interfaces, they are likely to produce points of strain as these tools are applied. Designers of such systems should get feedback by paying close attention to the problems experienced by artists who are pushing the edge. These observations can guide the construction of the next generation of computing media so that they are better fitted to evolving forms rather than replicating old ones.