

From UNIX to the desktop : a semiotic perspective on digital (r)evolution.

Jean-Marc Orliaguet
jmo@medialab.chalmers.se
01-09-1999

Abstract

From a semiotic point of view this paper discusses why indexical and iconic forms of interaction with computers are, in the perspective of digital evolution, reactionary positions, and why the only revolutionary approaches are to be found in symbolic modes communication.

INTRODUCTION

Computer-aided publishing started in the 80's as a replacement of large and expensive printing machines — from ink to electronic code. The metaphor was rather straightforward : what you see is what you get (WYSIWYG). Computer-aided design (CAD), computer-aided music (CAM) ... all share the common feature of being computerized variants of artistic or technical activities with pre-established codes and conventions : musical notation, technical drawings ... These are examples of iconic transformations, from the paper to the screen.

Digitality has gone a step further than simple computer-aided activities through a systematic digitalization of all media and knowledge, the transformation of analog material into bits, using a most extreme symbolic notation : binary code. Films, photographs, books, sounds are being converted into long series of 1's and 0's breaking apart the continuity of the original model into a complex binary system.

In physics, the principle of entropy¹ states that the system will tend towards chaos unless effort is expended to organize its information units, in other words unless new rules and conventions are defined.

However, on the biological level, natural evolution shows that the opposite happens : as systems become more complex, they also become more sophisticated, more organized, less chaotic, apparently contradicting the principle of entropy. To explain the contradiction, physico-mathematical models developed in cybernetics in the 50's led to consider information science and communication as a negentropy² phenomenon (“negative entropy”) that has for effect to reduce the degree of uncertainty³ in a system, increasing its energy instead of dissipating it. [Joly, 1994].

In either cases, because binary data is not directly understandable by humans, digital information has to be *re-presented* using a sign system and a representation mode — either symbolic (ex : Unix) or non-symbolic (ex : the desktop ”metaphor”).

In which direction does the nature of a sign system designed for human-computer interactions influence digital evolution?

¹“when a system containing a large number of particules is left to itself, it spontaneously assumes a state of maximum entropy - that is, it becomes as disorderly as possible.” [Young, 1987]

²A meaningful interpretation of negentropy is that it measures the complexity of a physical structure in which quantities of energy are invested, e.g., buildings, technical devices, organisms but also atomic reactor fuel, the infrastructure of a society. In this sense organisms may be said to become more complex by feeding not on energy but on negentropy (Schroedinger). [Krippendorff, 1986]

³*information* : that which reduces uncertainty. [Shannon, 1948]

⁴PEIRCE, Charles Sanders (1839-1914), American philosopher and physicist. Founder of semiotics.

⁵N.B. not to be confused with religious iconography.

SEMIOTIC TERMS

Communication between humans and computers, like any type of communication, implies the presence of sign systems organized according to codes. Understanding a sign involves applying the appropriate rules of a code.

In Peirce's ⁴semiotic model, a sign is a triadic relation between a *signifier* (or *representamen* : the sign vehicle, the form which the sign takes ex: a command, a menu, an icon ...), a *referent* (or immediate *object* : the concept, the action performed, what the sign stands for ...) and the *signified* (or *interpretant* : how the sign is interpreted in a given context, the sense made of the sign).

The process of making sense of a sign, also called *semiosis*, is done through a series of hypotheses performed on the signifier (in peircean terms : *abduction*) to determine if the signifier is an instance of a familiar rule. Signification is the result of applying the rule.

Signs can be classified into three main categories, derived from three modes of relationship between the sign vehicle and its referent: *indices*, *icons* and *symbols*.

The *index* is a mode where the signifier is directly connected to the referent, either causally or physically. The connection is always observable or easily inferred, for instance a flashing red light (the signifier) indicates that something is wrong (the referent), a progress bar indicates a level of progress in an action being performed, or a volume slider represents the output volume of a soundcard ... Indexical signs are the expression of a contiguity between a signifier and its referent since they are physically and temporally dependent on each other.

The *icon*⁵ is a mode where the signifier is indirectly linked to the referent by a resemblance, a similarity in appearance or quality. For example a folder (the signifier) is an iconic representation of a file directory (the referent) based on similar functional qualities. Icons reveal a sense of continuity between the signifier and its referent, however because by contrast with indices, the connection requires a leap of the imagination, there is no contiguity in iconic representations.

The *symbol*⁶ is a mode where the signifier is connected to the referent by arbitrary rules, chosen by pure convention and usage. As a result, symbols exist for themselves, free from any sort of motivation and necessity as present in icons and indices. For instance in UNIX, the command *ls* (the signifier) outputs a list of files in a directory (the referent), and even if the English verb “to list” is at the origin of the symbol “*ls*”, it neither excludes other spellings nor makes this very interpretation a necessity.

It should be noted that the classification in distinct categories is artificial; signs are not exclusively either indices, icons or symbols but often a combination of all three.⁷

IDENTIFYING SIGNS

Icons and indices are attractive because of their necessity to signify; in other words, unlike symbols, their presence is never gratuitous (for example the thumbnail of a photograph — an iconic representation by excellence — implies the near presence of another pixel image, similar in all aspects to the thumbnail but the size).

However, to be interpreted correctly, icons and indexes need first to be identified as such: an icon is not an icon until it has been interpreted as an icon. Therefore there is always a part of convention, a rule of interpretation, in identifying iconic and indexical signs.

As an example of this : a hyperlink on the World Wide Web is represented by a word of active colour underlined (a *symbolic* representation), when the mouse pointer moves onto the hyperlink, it takes the shape of an index finger (an *iconic* representation), the change of shape itself being an *indexical* sign that the word underlined is actually a link and not just a typographic effect [Codognet, 1996]. There is no such thing as pure iconicity:

For a sign to be truly iconic, it would have to be transparent to someone who had never seen it before — and it seems unlikely that this is as much the case as is sometimes supposed. We see the resemblance when we already know the meaning. This is especially true with onomatopoeic words which supposedly imitate the sound of their referent. The Russian words *puknut*’ and *pyornut*’ for example are regarded as onomatopoeic by Russian speakers, but it is not possible for someone who does not speak Russian to work out their meaning from the sound alone. [Cook, 1992]

Therefore, in the process of transposing a concept from a symbolic to an iconic or indexical representation, for signs to be discovered and identified, a whole set of explicit arbitrary rules become implicit. They generally are regarded as “natural” or “taken for granted”, because they are not in contradiction with our physical reality. For instance, the desktop “metaphor” uses most of the conventions and habits that people acquired in an office environment.

Convention is necessary to the understanding of any sign, however iconic or indexical it is. We need to learn how to understand a photograph... Convention is the social dimension of signs... it is the agreement amongst the users about the appropriate uses of and responses to a sign. [Fiske, 1982]

⁶N.B. not to be confused with Jungian or Freudian terminology.

⁷*There are no pure signs.* No sign perhaps, can perfectly realize any one of these types [namely, icon, index, symbol]. They are like chemical elements, which the very laws of chemical reaction prohibit us from obtaining in absolute purity, but to the purification of which we can so far approximate as to get tolerably accurate ideas of their nature, and which present themselves habitually in such a degree of purity, that we have no hesitation in saying, This is gold, that silver, and the other copper; or this is iron, that nickel, and the third cobalt; although all are strictly mixtures of the three. [Peirce, 1902]

Consequently, the use of non-symbolic representations is justified as long as the level of conventionality is smaller than in symbolic representations. The identification of indexical signs being usually easier since the signifier is directly linked to the signified, the question is especially pertinent to icons.

THE ICONIC TEMPTATION

Even when a symbolic approach would be preferred over an iconic, there is a natural propension amongst viewers to try and recognize familiar objects in visual messages, at any price — an “iconic temptation” that makes us feel uneasy when the iconic link with the real world is missing.

By drawing an analogy with how violently people reacted against abstract art in the beginning of the 20th century and how users today react against non-iconic user interfaces, one understands why the amount of realism involved in iconic representations is a sensitive issue. Too little realism in icons will make sign interpretation ambiguous and if the icon fails to communicate its iconicity, it is interpreted as an arbitrary sign : a symbol. As a corollary, symbols can be seen as icons or indices that have failed to communicate their indexicality or iconicity.

This has led interface designers to push the level of realism to the extreme through extensive use of analogies. Unfortunately, often excessive realism is synonymous with lack of flexibility:

Cast-iron realism prevents the user from getting ideas for modifying her tool and changing her working conditions because the technical workings of the system are so to speak sealed up. The Macintosh system I myself use is a good example of this. If a user wants to go beyond its friendly user interface, he enters a completely new world consisting of files, forks, resources, and similar strange creatures he has never encountered in the use situation. [Andersen, 1990]

In cases where flexibility is paramount, realism must be sacrificed. To illustrate the dilemma, let us consider the object *chair*. In order to represent a chair as iconic, one can choose a particular model of chair as signifier to keep a certain level of continuity with the object, but in order to be functionally as general as possible one should remove all peculiarities specific to a certain style of chairs and to a certain epoch. Unfortunately, by dint of simplifying the representation to the extreme of only drawing a few skeletal lines, moving from representational to abstract graphic, from a lawful to a proper sign, one eventually comes closer and closer to a symbolic representation.

The symbolic approach, very unsentimental in a way, is neither concerned with the spatial or temporal contiguity of the index nor the continuity of iconic representations. Practically it means that symbolic representations ignore realism, and can therefore distance themselves from the original material through several layers of abstraction and meta-information. (ex: in text processing: the *editorial* approach vs the typographic approach). That offers the possibility to contemplate the system from several standpoints instead of a having a unique representation of it.

Indexical and iconic approaches on the other hand will try to reshape the material after its original image, bringing back the

continuity that was broken, for instance by creating the illusion of the desktop according to the concept of “What You See is What You Get” or rather “What You See is *All* You Get” (ex: in text processing, the *typographic* approach), thereby denying any original aspect in digitality by tying it to the analog world. Of course, the whole process is in a sense very emotional, reminding more of the quest for paradise lost than a creation *per se*, as if to make us believe that all our belongings in the analog world survived the violence of digitalization.

FROM ICONS TO SYMBOLS

If in the space of our representations, indices are the first dimension and icons the second dimension, moving from icons to symbols is an act of adding a third dimension to our representations. Working directly on the symbolic level offers a deeper understanding of the media by abstraction of the levels underneath.

For instance, the what-you-see-is-what-you-get approach to text processing is both indexical and iconic. It is iconic because the screen is used in place of the printed media by virtue of the resemblance between pixels and ink dots, and it is indexical because if the typography is modified on the screen, by causality it is also modified on paper. Indexicality helps understand iconicity and vice-versa.

What do we gain by breaking iconicity and introducing symbols?

We gain by moving away from the ink, and as a result the media is no longer the *paper* but instead the *text* is the media. And because the symbolic approach⁸ focuses more on the structural organization of a text than its appearance on paper, there are no longer paper sizes, page orientations, page breaks, carriage returns, big or small typefaces ... there are instead introductions, abstracts, descriptions, sections and subsections, quotes, citations, references, conclusions ... The new dimension introduced by symbolic representations in text processing is the meta-information that forces us to reason on the editorial rather than the printing level.

HTML was at the origin purely structural, i.e. most tags had a semantic function (for EMPHASIZE, <H1> for HEADER 1...), but it turned over the years into a descriptive language (<I> for ITALICS is now used instead of because both tags achieve the same visual effect on screen) — again, the iconic temptation.

By writing <I> instead of , we apply a semantic rule that says that emphasis is represented in italics. Then, when encountering words in italics, we perform a series of hypotheses on which rule was applied. By abductive reasoning based on our personal experiences with texts, we may interpret italics as emphasis or simply as a typographic effect.

In the case of photographs as another example of iconic representations, the missing third dimension is also recreated mentally, from our experience with the three-dimensional world.

Unfortunately because we have no cognitive awareness of digitality, the interpretation of digital data from iconic information leads nowhere but back to the analog world. Therefore the key to digitality is the symbol.

SYMBOLIC LINKS

Furthermore, there are important concepts that are impossible to represent otherwise than symbolically.

When duplicating files, Unix makes a distinction between *hard links* and *symbolic links* (symlinks). A hard link from file A to file B presupposes that file A exists⁹, while a symlink from file A to file B does not presuppose that file A already exists.¹⁰ . (Accessing file B is then equivalent to accessing file A).

Giving an iconic representation of the concept of a hard link is straightforward because the target (the file being duplicated) already exists. It is done by clicking on the icon representing the file and by creating an exact visual copy of it, a *double*. This visual procedure is of indexical nature: by inference it implies that the file that has just been physically created is interchangeable with the original one.¹¹ B is a clone of A.

However, in the case of a symlink, the target does not necessarily exist, in the sense that it has not been defined. Technically speaking, the class of symlinks exists but it has not been given an instance. In this case, there is no physical data to select and duplicate. Although the operation of linking is easy to perform mentally — it is one of mapping between sets, which can be written down as $B \mapsto A$ because A and B are generic terms — representing it iconically is a failure.

For that reason, symbolic links have found their nearest equivalent on the desktop as “shortcuts”, or “aliases” with the restriction that they mix the notion of a *link* with its *target*.

FROM SYMBOLIC LINKS TO THE WWW

One may object that symbolic links are not so important after all. Yet the concept of symbolic linking can be considered at the very origin of the World Wide Web, much more than hypertext.

Even if the Web is based on the idea of hypertext that first appeared in the 60’s, and later found an incarnation in HyperCard (1987) and various online manuals (ex: Insight, Microsoft online help), one should not forget that the direct ancestors of the Web are the *Gopher*, *Archie*, *Veronica*, FTP servers, the e-mail — all based on the concept of symbolic links, the idea that destinations outside a given set may disappear while links to them continue to exist.

Is that to say that the Web owes its existence to UNIX’s symbolic links? Certainly not, but it is remarkable that the Xanadu Project, as described by Ted Nelson in 1974 in *Dream Machines* [Nelson, 1974] as a global hypermedia environment, despite the richness of the hypertext conceptions if offered, never materialised in its proposed form; maybe because it defined a link as

⁸cf. L^AT_EX, SGML, XML

⁹The concept of making a copy of A and calling it B. Can be written “*ln A B*”. Both files share the same inode.

¹⁰The concept of creating B as a reference to A, written “*ln -s A B*”, represented as : B -> A. Operations done on B are actually executed on A: the link is “dereferenced” by the kernel.

¹¹Umberto Eco in “the Limits of Interpretation” [Eco, 1990] : “Two objects are doubles of one another when for two objects O_a and O_b their material support displays the same physical characteristics (...) and their shape is the same (in the mathematical sense of “congruence”)”

“the *traversable* connection between two nodes” where “documents must remain accessible indefinitely”.

In that sense, the Web takes the concept of hypertext a step further, accepting the idea of “dead” links, it gives documents a temporal existence besides their spatial presence, i.e. documents have a past, a present and a future : they become alive, unlike icons that ignore temporal fluctuations.

FUTURE THOUGHTS AND CONCLUSION

To put the argument in relation with a long-term evolution : moving from indices to icons, and from icons to symbols is a refining process where representations become more articulate, breaking away from the continuity of icons and the contiguity of indices. It is also the evolution of mankind to move from primitive to abstract representations, that which demands most effort. Man has evolved from primitive state to civilization and culture, acquiring knowledge through the mastering of symbolic ways of thinking and communicating, going beyond the simple understanding of indexical and iconic signs overly represented

in the vegetal and animal world.

Art, imagination and dreams, on the other hand, take the opposite direction of our technological evolution, they strive to de-symbolize our representations, trying to find the icon beyond the symbol, and the index under the icon... The dreaming act for example transforms our verbal thoughts into images (iconic representations), all full of hints (indexical representations). [Joly, 1994]

Of course, de-symbolizing activities allow us to look back on our epoch, they offer us a necessary *reflection* on our time. As digital evolution has only begun, it is no wonder that indexical and iconic representations take an overwhelming place in our computer environments, they are after all natural reactions against an excessive amount of symbolic abstraction in digital systems. Nevertheless, one must understand that purely indexical and iconic approaches strive backwards.

In summary, the danger is to create a digital world that eventually has nothing more and nothing less to offer than the analog world we already know, because we tried so hard to imitate it.

REFERENCES

- [Andersen, 1990] Andersen, P. (1990). *A Theory of Computer Semiotics*. Cambridge University Press.
- [Codognet, 1996] Codognet, P. (1996). The semiotics of the web. <http://pauillac.inria.fr/codognet/web.html>. unpublished.
- [Cook, 1992] Cook, G. (1992). *The Discourse of Advertising*. Routledge, London.
- [Eco, 1990] Eco, U. (1990). *The Limits of Interpretation*. Indiana University Press.
- [Fiske, 1982] Fiske, J. (1982). *Introduction to Communication Studies*. Routledge, London.
- [Joly, 1994] Joly, M. (1994). *L'image et les signes. Approche sémiologique de l'image fixe*. Nathan Université.
- [Krippendorff, 1986] Krippendorff, K. (1986). A dictionary of cybernetics. <http://pespmc1.vub.ac.be/asc/indexASC.html>. unpublished.
- [Nelson, 1974] Nelson, T. (1974). *Computer Lib / Dream Machines*. Microsoft Press.
- [Peirce, 1902] Peirce, C. S. (1902). Reason's rules. unpublished.
- [Shannon, 1948] Shannon, C. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, 27:379–423, 623–656.
- [Young, 1987] Young, P. (1987). *The Nature of Information*. Praeger, New York.