Modeling Links in Hypertext/Hypermedia

Oreste Signore
CNUCE - Institute of CNR - via S. Maria, 36 - 56126 Pisa (Italy)
Phone: +39 (50) 593201 - FAX: +39 (50) 589354 - email: O.Signore@cnuce.cnr.it

Abstract

The links between the nodes constitute one of the most significant aspects of the hypertext, and very much contribute to the support of their dynamic interaction paradigm, so relevant to the implementation of effective applications.

In this paper we firstly recall the characteristics of the hypertext, pointing out some of the main problems the hypertext applications' designer must face. Afterwards, we outline a hypertext design methodology that leads to a consistent design of the nodes' structure, and allows modelling the different kind of links that can exist among the nodes. In particular, introducing the "concept space" allows an effective modelling of the intensional links. Finally, we discuss some general implementation aspects, and point out the need of offering to the user multiple and interchangeable interaction paradigms.

Introduction

During the last years, the interest in hypertext has accelerated sharply ([Conklin87]). Several factors explain it: more powerful workstations, high resolution graphic displays, decreased costs for large on-line storage, availability of low cost and easy of use systems, and so on.

This led to an explosion of the hypertext applications, and, perhaps, to the tendency of considering it a "magic" solution to some old and known problems. According to [Nielsen90a]:

> *"hypertext is non-sequential writing: a directed graph, where each node contains some amount of text or other information. The nodes are connected by directed links. In most hypertext systems, a node may have several out-going links, each of them is then associated with some smaller part of the node, called an anchor. When users activate an anchor, they follow the associated link to its destination node, thus navigating the hypertext network."*

© Archives & Museums Informatics, 1995

When the content of a node is made not only of textual information, but also of other kind of media, like graphics, sound, and so on, we may use the term hypermedia.

Two major areas of application are education and documentation.

In the following, we will briefly illustrate the importance of the links, that are a fundamental component of the hypertext, and will discuss some design issues that can lead to the implementation of effective hypertext applications.

Hypertext/Hypermedia: General Issues

The Original Idea

Hypertext was originally conceived essentially as a tool for the management of the personal information, that is, the information needed by the single researcher for his/her own purposes. As a matter of fact, the description of the mythic "Memex" by Vannevar Bush ([Bush45]) sounds:

> *"A device in which an individual stores his books, records, communications, and which is mechanised so that it may be consulted with exceeding speed and flexibility. It is an enlarged intimate supplement to his memory."*

As a consequence, the user of a hypertext was thought as having a complete awareness of the content and the organisation of the information, i.e. the user is supposed to have a complete knowledge of the semantic domain the hypertext is concerned with. This implies that he/she should find quite evident and natural all the links, either they represent a "structural" organisation, either they just accomplish the task of representing some association among different information nodes.

However, we must note that even a user with a thorough knowledge of the possible associations among the different information nodes can meet many difficulties in finding the relevant chunk on information, exactly as it happens when we are looking for a book or a document that, we remember, "*has been stored in the bookcase, near to other documents related to the same topics.*"

The Relevance of the Links

As we have pointed out before, hypertext is constituted by nodes and links, which are of greatest importance. In effect, we can say that they are, in a certain sense, the 'essence' of the hypertext.

© Archives & Museums Informatics, 1995

Really, the links are the basic component that contributes to the *enrichment of the knowledge*, as they have a fundamental role in *stimulating the user's interest*, implementing *various types of connections* among the nodes.

According to some proposals in the literature ([DeRose89]) we can classify the links in two main classes: extensional and intensional links, each one of them is subdivided in more subclasses, and so on. Without going in more detail, we simply remind that the extensional links are explicitly stored in the hypertext, while the intensional links can be deduced from the context.

We have also to remember that each node can be linked to many other nodes due to many reasons, and this is especially true for the intensional links, that can be seen as the "added value" to the hypertext.

It is customary that the designer analyses the content of the nodes and identifies the "*anchor point*" ([Nielsen90a], [Nielsen90b]), that constitutes the smallest element in a node that can give rise to a link towards other nodes. The physical implementation of a link makes use of buttons, icons, hot-words, and similar.

This is the phase where a great number of links can arise. As a matter of fact, many designers realise that the links are the power of the hypertext, and try to emphasise this concept introducing as many links as possible.

Pros and Cons of the Hypertext

The advantages. Hypertext offers some relevant advantages: the first one is the possibility of immediately accessing the nodes containing the relevant information. Secondly, it is possible to have a non-linear organisation of the data, that very much resembles that of the human mind. Finally, it is possible to reduce the amount of information stored in each node. Therefore, instead of overloading the single node with plenty information, that can disturb the user, it is sufficient to suggest possible items that can give supplemental or deeper information. The user will activate the links at his/her will, so enriching his/her knowledge according the specific needs or interests

There are several ways for accessing the information: *navigation* from one node to another, *browsing*, i.e. visualisation of a set of nodes and choice of the most relevant or "promising", *querying*, i.e. the selection of a set of nodes that fulfil the information requirements contained in a single query. Quite obviously, we must expect that the user will intend to use these access modalities as many times as he/she wants, and in arbitrary order.

Some "structural" problems. At least two problems, that constitute a real challenge, are common to all implementations: the *disorientation* and the *cognitive overhead* ([Utting89] [Parsaye89] [Nielsen90b]).

© Archives & Museums Informatics, 1995

The first one is originated by the structure of the hypertext itself. In fact, this model should not only organise the information, but also give us advice about our location in the network and the feasible path to follow in order to reach a specific node, whose existence we are well aware. Similar difficulties can arise in linear texts, too, but in this case the user can simply move back and forth. The hypertext offers a much higher degree of freedom, as we have more than one dimension. Therefore the possibility of being "lost in the hyperspace" is much higher.

A possible way out to this difficulty is given by the availability of browsers, so that the user could easily realise the context he/she is in, and be aware of the structure he/she is navigating upon, according to his/her peculiar perspective. However, the possibility of browsing the whole database has some counterparts: it may increase the possibility of becoming lost, or at least to be confused, while performing the search. In many cases, the browser simply represent the nodes' schema, without taking into account that a global view can be of reduced utility to the user, who can become confused looking at a complex and intricate nodes' network.

The second problem arising in the usage of the hypertext is related to the difficulty of being adapted to the *cognitive overhead*. This difficulty comes up in the design phase as well as in the navigation phase. In the first case, the designer must decide about the structure to give to the texts, which links, and how many links to create, where anchor the links, etc. During the navigation, the user must operate a choice in front of a wide number of possible links.

Both the problems can have a relevant effect if the user will not have a suitable help.

Implementation problems. Independently of the difficulties due to the peculiar tool adopted for the implementation of the hypertext, some considerations have a more general relevance, and can heavily impact on the design phase.

The first aspect is related to the *extensional* or *explicit links*: as they are defined by the designer, they necessarily are dictated by his/her knowledge on the specific domain and mental organisation, that therefore are "hard coded" in the hypertext. As a consequence, we can have a large variety of choices, comprised between the two extreme cases: the first is a completely "flat" hypertext, that in fact, in spite of some technological effects (sound, animation, special effects, etc.) reproduces a poor quality book or a traditional application (menu cascade). On the other extreme, we have the case when the designer realises the power of the links and put on every node an enormous amount of links, referring to the most disparate destinations, according to a tumultuous process of association. In both cases, the user in fact looses his/her intellectual independence, as he/she is either forced to follow a non stimulating path, or receives an enormous quantity of stimuli, and therefore cannot follow his/her logical and conceptual associations.

© Archives & Museums Informatics, 1995

The second problem arises from the need of representing the *intensional* or *implicit links*, implementing interaction mechanisms that will emulate as far as possible the human mind's associative mechanism.

The third problem, that in some way has some overlap with the others, is well known in the context of the cognitive psychology, and is much more common that we can suppose at first glance. In fact, we face this problem every time we try to make use of a common use object that is not "well designed"[1]. In all these cases, we are in presence of a problem of interface design. As shown in figure 1, the system's interface accomplish the task of communicating to the user the designer's mental model. As a consequence, the user will build a mental model according to the "messages" communicated by the interface. In practice, the two models only rarely are consistent, in many cases they can be totally different. [Norman88] reports many examples where the poor interface design was the cause of disasters.

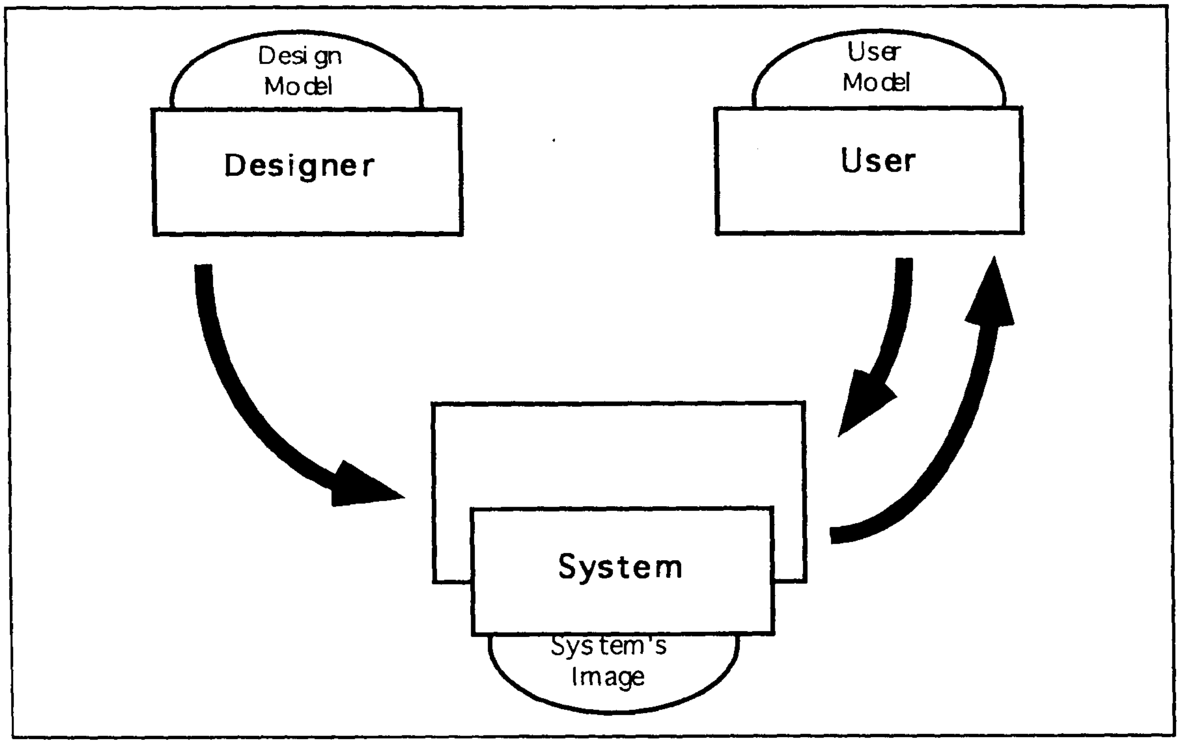

Fig. 1 - The role of the interface as a communication medium to the establishment of a mental model (from [Norman88]).

1 Any of us can remember cases when he/she was in trouble in using some shower controls, or car doors openers, or tools with the on/off switch in secret or unthinkable positions.

© Archives & Museums Informatics, 1995

Hypertext Design Issues

Generalities

Hypertext is related to other areas, either of computer science as well as pertinent to other disciplines (figure 2). We have to note that some of them are "mature" and can give a valuable help in the definition of a design methodology.

More precisely:

- *Database Technology* can be of invaluable help in the process of *Data Modelling*, i.e. in the phase where we define the structure of the nodes and their content;
- *Information Retrieval* is a sound theoretical basis for the identification of the most suitable techniques to be adopted for the *Indexing* of the free text information that is so relevant in hypertext applications;
- *Artificial Intelligence* can contribute to the problem of *Connecting data items*, i.e. capturing the relations that can be established among the nodes on the basis of the choices taken by the user;
- *Cognitive Psychology* is essential to implement an effective *User Interface*, capable of conveying to the user the designer's mental model.

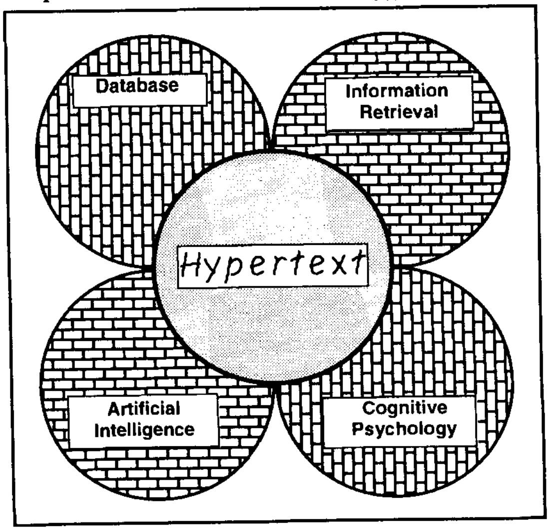

Fig. 2 - Hypertext and other disciplines

Some additional considerations must be done to identify the characteristics of an effective hypertext application:

- The *purely sequential visit* of the hypertext nodes is a too simple interaction, can help in the very initial stage the less experienced user, but will soon appear less appealing that a conventional book.
- Obliged connections should be avoided as far as possible, as they force the user to follow associative paths that are evident to the designer, but sometimes obscure to the user. On the contrary, the user should have the possibility of *associating the concepts* that can be tied to the

© Archives & Museums Informatics, 1995

information nodes, so that he/she will be able to exploit the navigation capabilities at the highest level.

- The user should have the possibility of choosing *several interaction paradigms*, as the association among the nodes can be dictated by several reasons (contiguity in space or time, relationship among the associated concepts, etc.). In addition the user must have the possibility of *switching* from one paradigm to another, so emulating the way the human mind works.

- The implemented *links* must be of different *types* (or, at least, their meaning must be quite evident to the user). Even more important is to *visualise* the links, so that the user can have a clear idea of how much a link is "promising" in terms of the number of nodes he/she can access. Finally, the *weight* of the links (i.e. the importance given to the specific association) should vary accordingly to the user's interests.

The nodes

The implementation of an effective hypertext has as a prerequisite the accurate analysis of the information to be managed. Making an analogy with other application environments, especially the database environment, the definition of a conceptual schema, that is a semi-formal description of the world of interest, takes a fundamental importance. The conceptual schema allows to represent the relevant objects and the relationships among them. This approach appears to be very useful as it is based upon consolidated methodologies, that allow to represent the world of interest independently from the peculiar technological instruments that will be adopted in the subsequent implementation phase.

The architecture of a hypertext design methodology ([Signore94]) is reported in figure 3.

Once the conceptual schema has been designed, it is possible to proceed to the definition of the hypertext's structure by a straightforward process. Roughly speaking, entities are transformed into hypertext nodes, while the relationships can be mapped into extensional links.

Going down to the physical level, the identification of nodes' types and links' types allows the definition of nodes' structure that make them clearly distinguishable, while maintaining a "family feeling". By this, the structure of the node can automatically transfer information about the kind of information it contains.

In defining the nodes' structure, we may introduce the additional concepts of *components* and *perspectives*. This means (figure 4) that every node is made by several components, that can be seen as an enrichment of the concept of "field" of the card or the "attribute" of an entity. In fact, the components of a node can be of several types, e.g. sound, animation, image. This assumption leads to the extendibility of the proposed methodology to the design of hypermedia as well as hypertext applications, by means

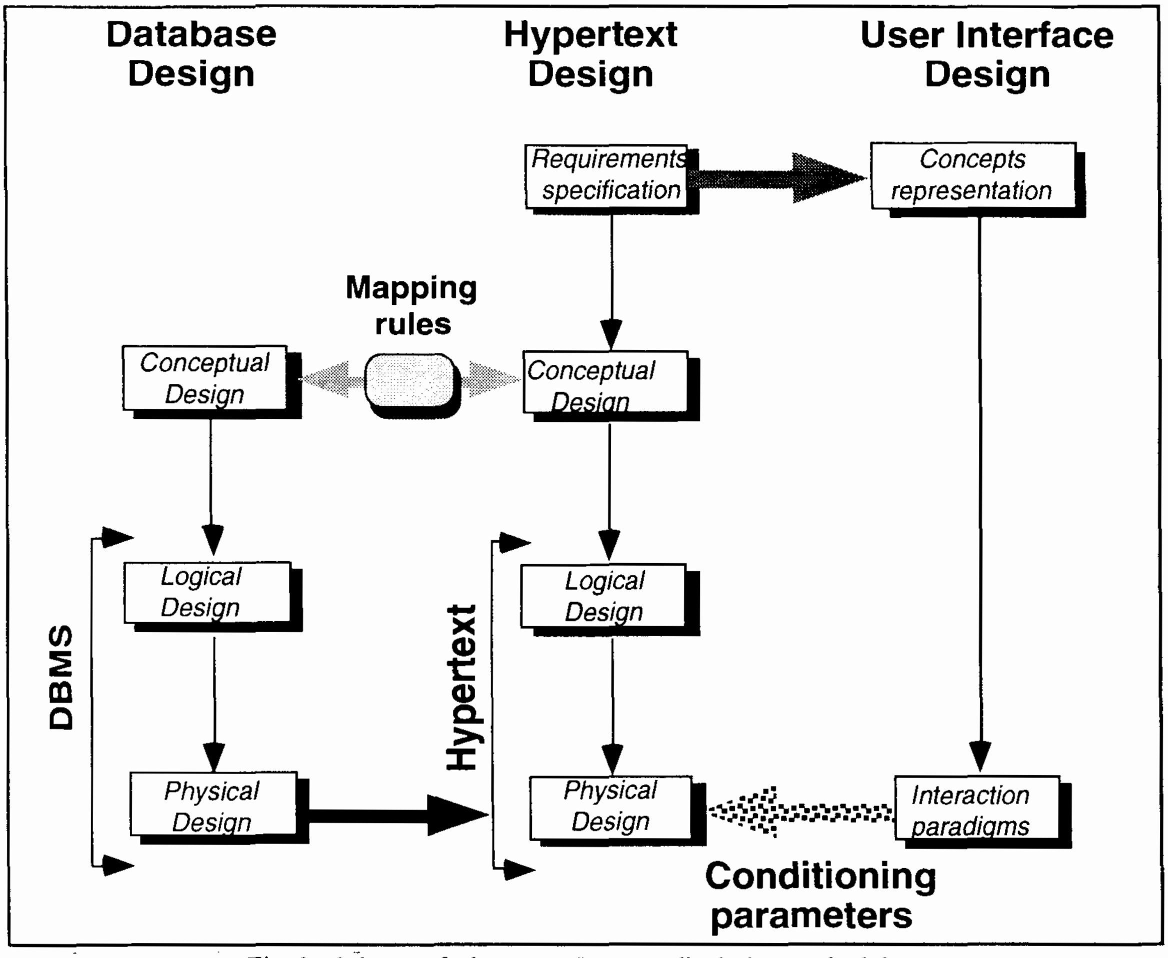

Fig. 3 - Schema of a hypertext/hypermedia design methodology

of a unified approach.

In addition, if we consider that every component can be seen from different perspectives, we can easily manage, at the design level, many problems, like the different resolutions of the monitors, user preferences, multilingual support, etc.

Finally, every node will contain several links towards other nodes, dictated by the associations modelled during the conceptual design phase.

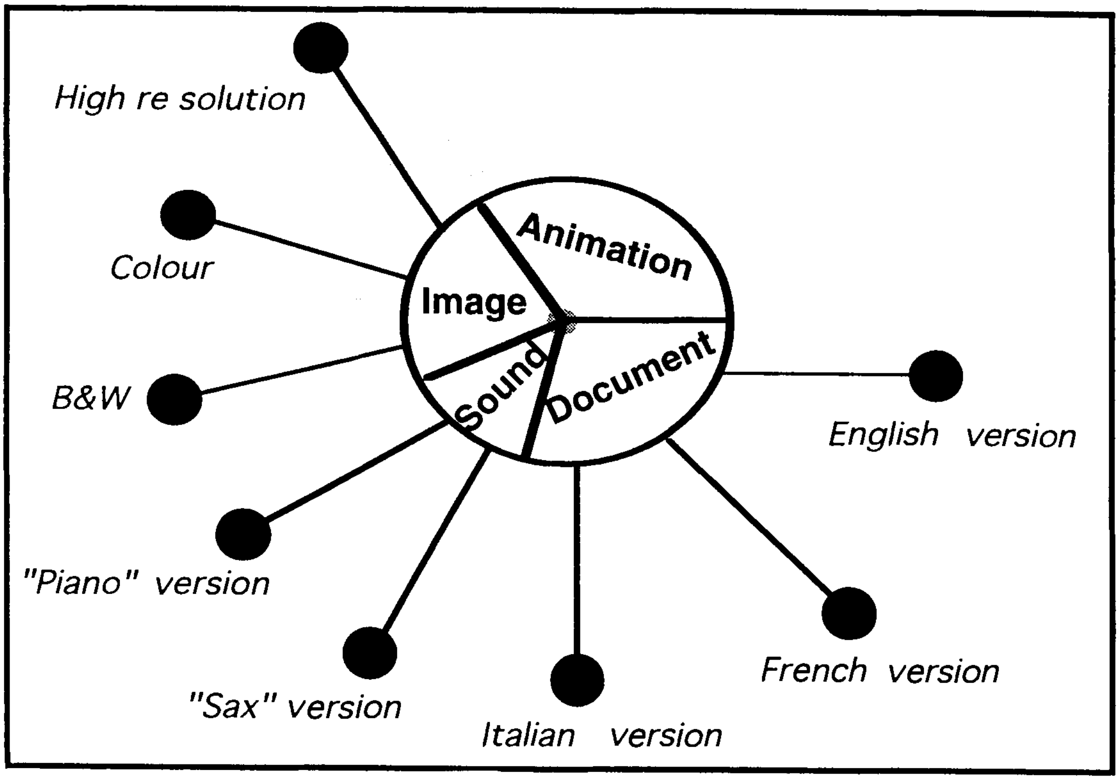

Fig. 4 - The structure of a node: components and perspectives.

Associations among the Nodes

As we have pointed before, the links are the hypertext component in charge of representing the associative aspect, extensional as well intensional.

In fact, the extensional links are essentially structural links. In figure 5, the user made use of a relationship between the descriptive card of a scientific instrument, its picture and bibliographic references.

However, the intensional links constitute a very different case, as they are modelled on the associative process typical of the human mind.

We can emulate this process implementing a "*concept space*", that makes explicit the relationships existing among the concept that can be attached to the single information node (figures 6-7). By this approach, we can implement a representation of the knowledge on the specific domain, and therefore we are no more forced to make explicit all the possible links existing among the various nodes. As a consequence, we can reduce the cognitive overhead given by the presence of an excessive number of links on the single node. It is evident, however, that this solution will be effective only for really "active"

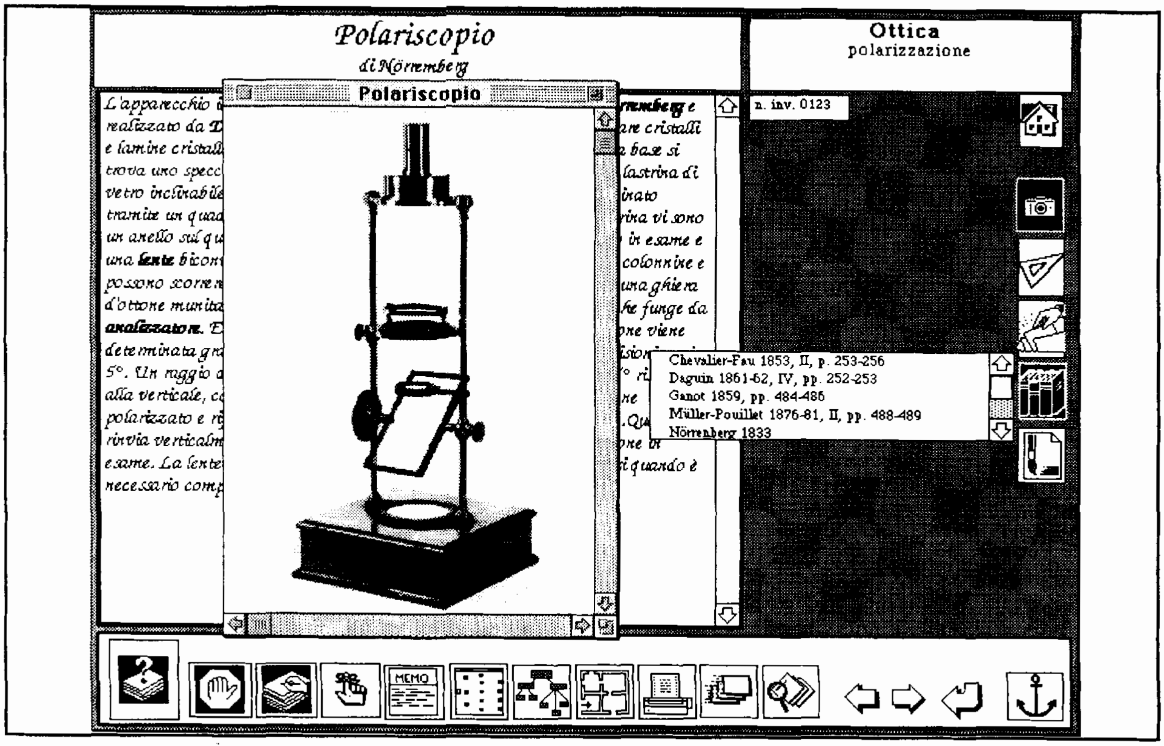

Fig. 5 - An example of activation of extensional links to visualise the information related to a node.

users.

Such a mechanism hypothesises that the user has the ability of abstracting the significant concepts from a single node, and subsequently associate them in producing a personalised cognitive path. Therefore it can be considered especially valuable in the educational applications. In fact, the implementation of an interaction paradigm based on the navigation through the concept space very much supersedes the widely diffused concept of "electronic encyclopaedia". These last, really, very often are simply a (possibly large) set of nodes, connected by explicit references ("hot words"), mapping predefined relationships.

The navigation through the concept space, on the contrary, enables the user to operate an abstraction process, then following the associations among the concepts, finally descending again to the information space. This process appears much more similar to the natural process of the human mind. The relevance of the proposed approach consists in the e introduction of a large number of potential links among the nodes, but avoiding that their proliferation will overload the node or disturb the user[2].

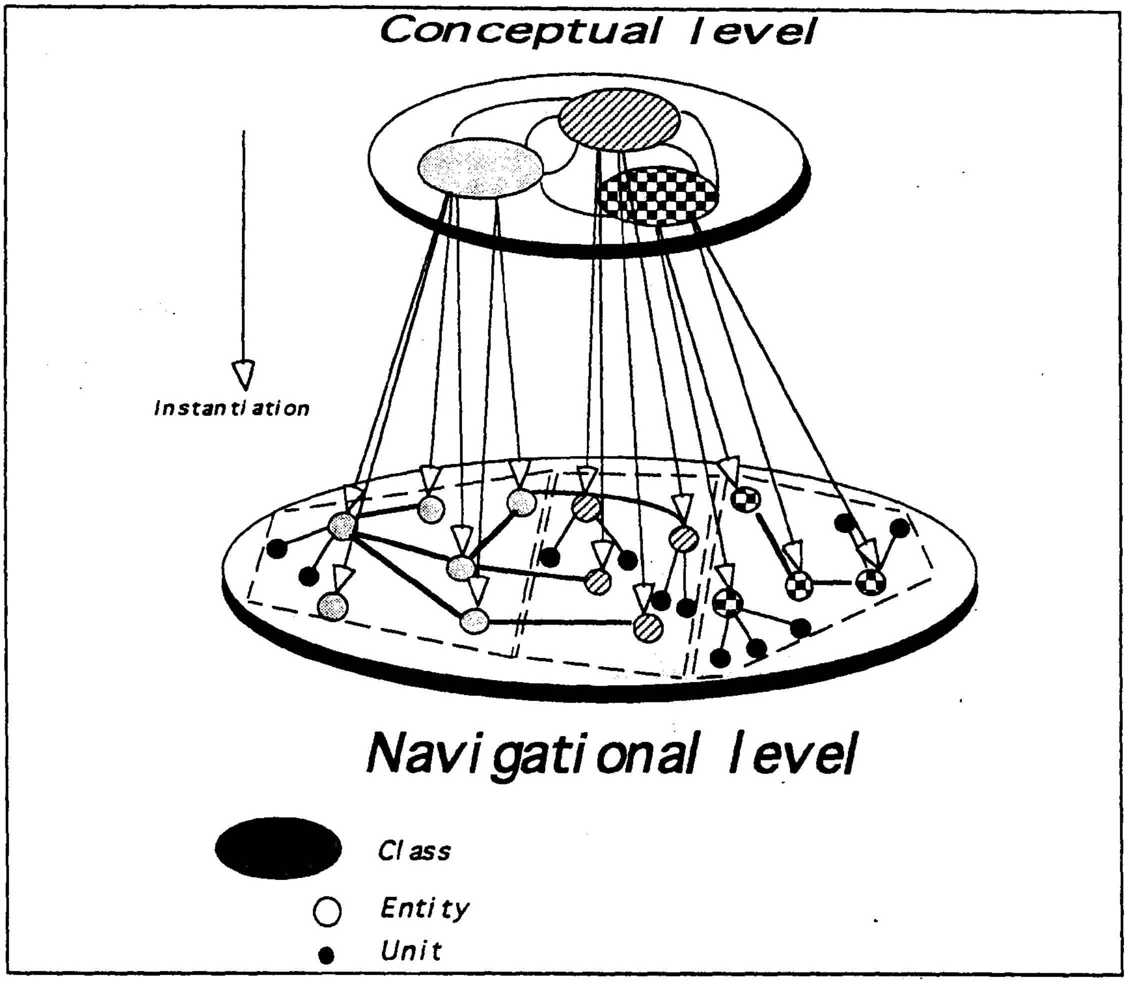

Fig. 6 - The two abstraction levels in the proposed hypertext model

Implementing the Links

As we stressed the importance of the links, clearly we must pay much attention to their implementation as well as to the way their are made explicit to the user.

In the literature, the authors assume contrasting positions about the opportunity of having "typed" or "non-typed" links. Without going into technical details, we must stress that the user must be made well aware of the semantic meaning of the links, and therefore of its destination node type. As an

2 In the 18th century, the authors of the famous Encyclopédie had very similar findings and elaborated a geographic metaphor.

© Archives & Museums Informatics, 1995

example, when managing a collection of objects, there is an intrinsic difference among the link connecting two objects, the link connecting the object to a person, the link addressing a dictionary item.

Another arising difficulty comes from the necessity of making evident these differences to the user, especially when a single "*anchor point*" is the starting point towards many other nodes (figure 8).

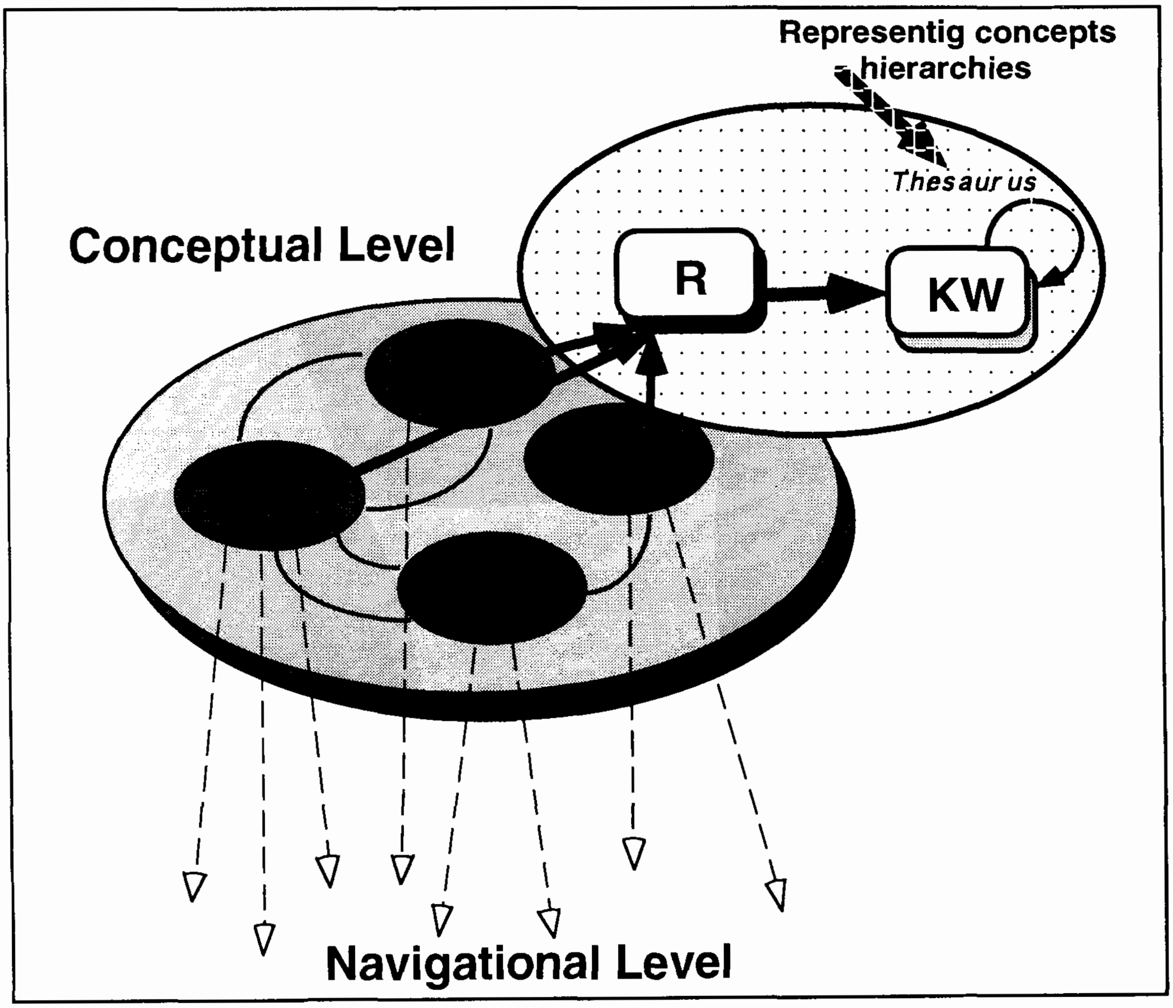

Fig. 7- The concept space at a higher detail level

We can reduce the difficulties originated by the cognitive overhead and disorientation by making evident the links that exist among the various nodes, so that the user could immediately realise how much the activation of a link can be "promising." To emphasise this concept, it is worthwhile to introduce the concept of the link's weight, that will measure the affinity degree between two nodes. This concept allows to distinguish among the links representing a strong interrelation between two nodes from the

© Archives & Museums Informatics, 1995

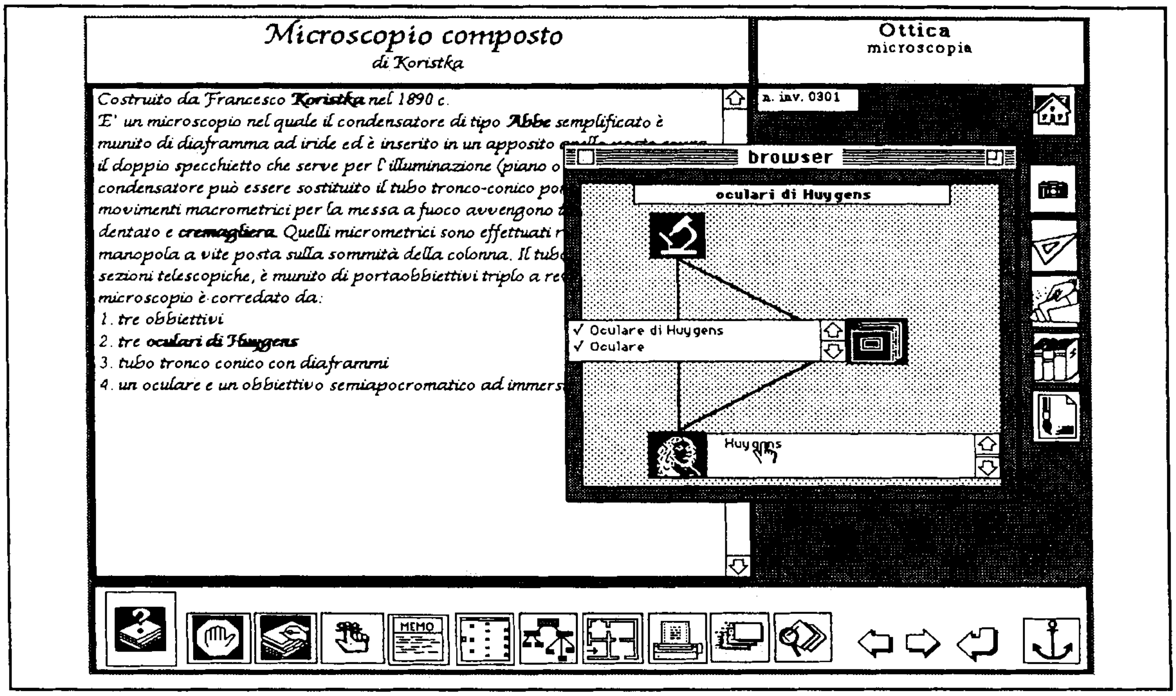

Fig. 8 - An example of visualisation of a "multiple link"

loosely connected nodes. It is quite obvious to give a "spatial" representation of the affinity degree, where the distances between the nodes decrease when the affinity increase (figure 9).

Looking to this problem at a higher detail level, we can realise that the weight of the link between two nodes cannot be considered an "objective" quality, but it is related to the specific user's interests. To take an example, would we implement a hypertext on plants, it will be possible to imagine several links among the various exemplars on the basis of their botanical affinity, or from a single exemplar to cultural specificity, the historical interest, usage and properties. However, the importance of these links depends on the peculiar aim the user has in his/her mind when navigating the hypertext. In particular, some of these links could be of null importance, and the user would be happy do not have them showed. As a consequence, we must consider the usefulness of defining a "user profile" so that the user should be able to modify at his/her will the importance given to the different links (figure 10).

The Interaction Paradigms

Traditionally, the designer's knowledge is "hard coded" in the hypertext, so forcing the user to follow undesirable cognitive paths. We can reduce this disturbing effect by implementing many and interchangeable interaction paradigms, so that the user will be able to follow the mental paths that he/she will find the most natural and significant time to time.

© Archives & Museums Informatics, 1995

The most obvious paradigm is based on the physical contiguity of the nodes. In this case, we will

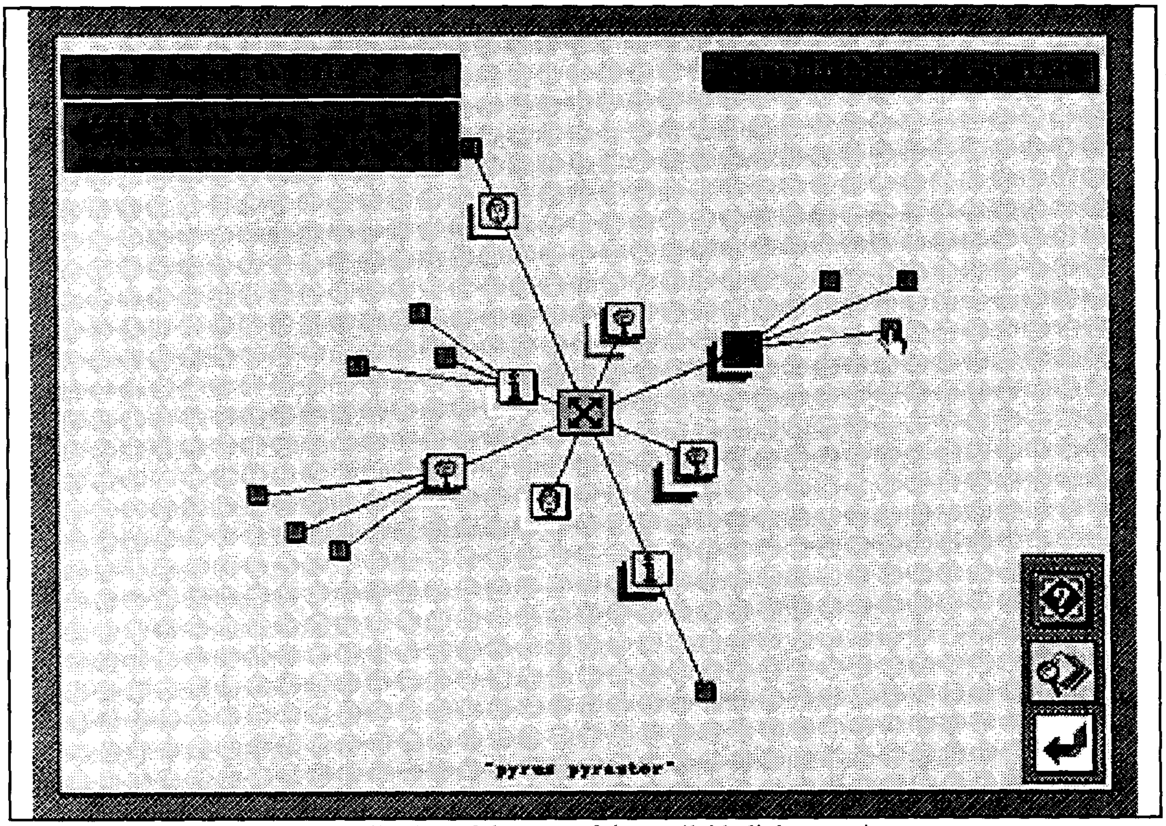

Fig. 9 - An example of the visualisation of the map of the available link to navigate.

navigate the hypertext in the same way we would browse the pages of a book. Unfortunately, in many cases this is really the only available interaction paradigm. In fact, links towards other nodes, or the possibility of tracing back the followed path, do not substantially modify the characteristics of this kind of interaction, which very little stimulates the user, and do not much differ from the conventional cascade of menus. In addition, some undesirable effects can arise, like the navigation to a totally out of context node, reached simply because of its physical contiguity with the previous one. In conclusion, this kind of interaction paradigm forces the user to follow the cognitive paths defined by the hypertext's designer. Only predetermined and foreseen associations, explicitly implemented by extensional links, can be followed.

It appears much more relevant to implement more flexible and stimulating interaction paradigms, that will allow the user to follow intensional links ([Signore93a], [Signore93b]).

On the basis of previous experiences, it seems that we can reconduct the most useful paradigms to three basic classes: classification, map, time.

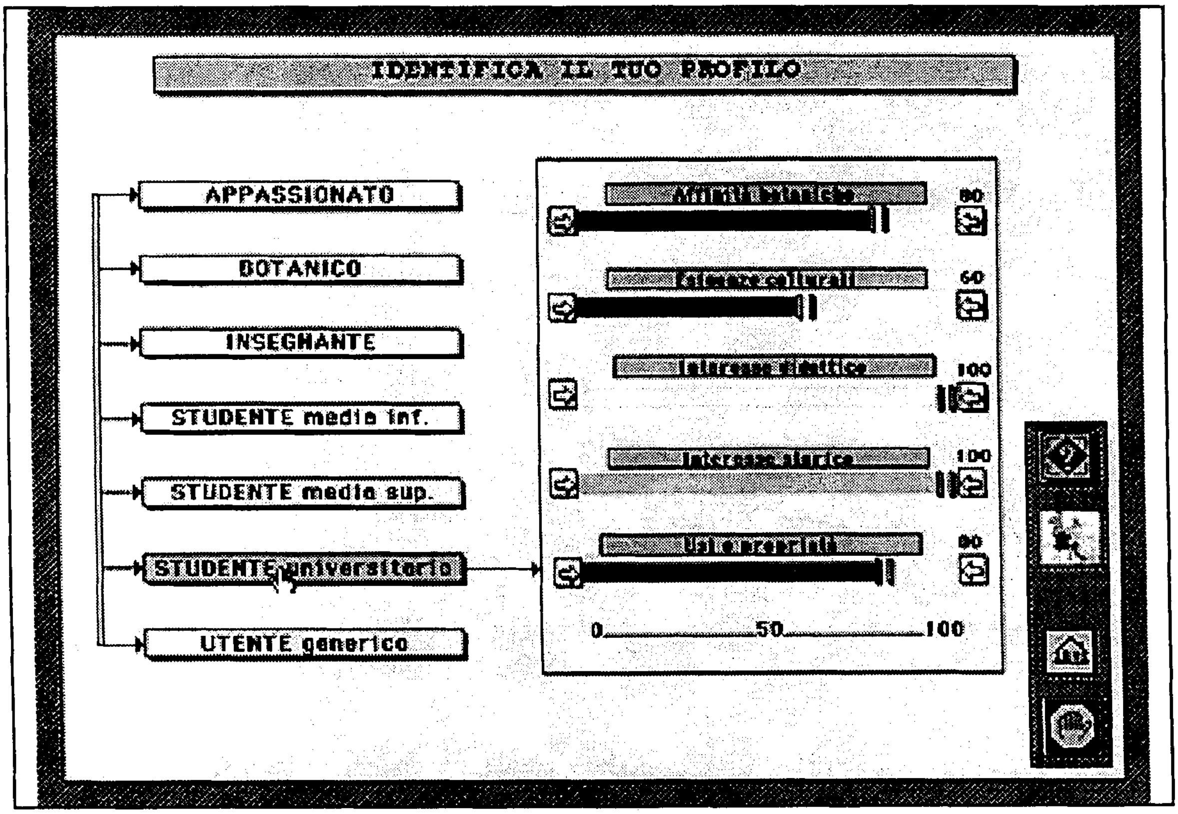

Fig. 10 - The definition of a "user profile"

The *classification paradigm* allows the navigation through the concept space: from a single node the user will rise up to one of the concepts associated to it. Afterwards, moving across the relationships that map the domain knowledge, he/she can identify other concepts. From these, it is possible to go down again the nodes' space. Figure 11 shows an example where from a classification of some topics in physics it is possible to directly access the related instruments.

Widely used and useful is the *map paradigm*. In this case, the user can interact with a topographic or geographical map, selecting the interesting zones and choosing the nodes to be reached on the basis of their physical location (figure 12).

In many applications it can be very useful the implementation of a *temporal paradigm*, that will allow to link the nodes taking into account the contemporaneity (restricted or extended), the temporal sequence, the overlapping of temporal intervals, and so on.

Many applications can get advantage from the availability of the map and time paradigms at the same time. This allows to put the information in the right space-time context, according to a mental model widely used in many areas, especially in the management of the cultural heritage ([Signore90]).

© Archives & Museums Informatics, 1995

The implementation of all these paradigms requires the existence of retrieval intensional links, and

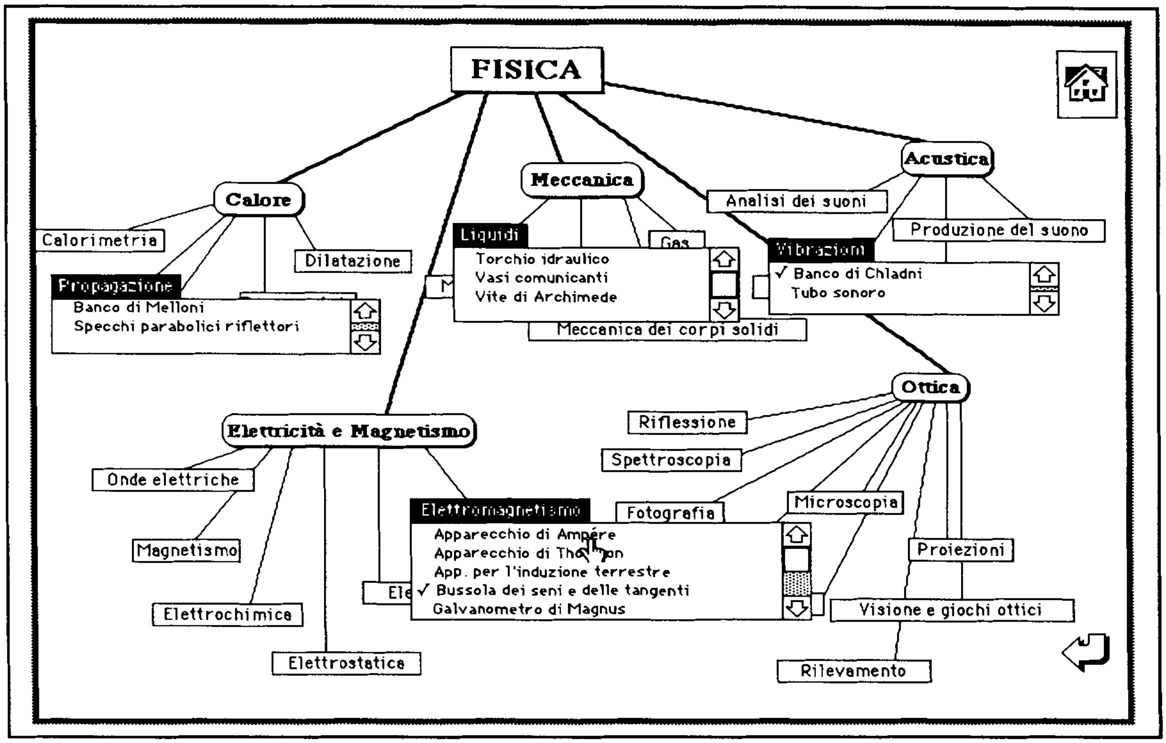

Fig. 11 - An example of interaction paradigm based on a classification scheme.

therefore of a mechanism for retrieving the information associated with every node ([Signore91]).

Other aspects

When implementing a hypertext, we must consider that the target user is supposed to be particularly "active" and "curious". As a consequence, we must foresee the implementation of *annotational links*, that will allow to capture the user's knowledge, that can be added to the hypertext. In a subsequent phase, the annotations can be made accessible to the whole users' community, and will become part of the common patrimony of knowledge.

In many cases, especially in implementing educational applications, we must provide for, and implement, *vocative implicit links*, that will reference dictionary entries.

Finally, the designer must provide guided tours, that will satisfy some peculiar needs. However, the user must be free of leaving the tour at his/her will, would he/she desire to perform a deeper investigation of some aspects, or follow associations dictated by his/her experience in the specific field. It is

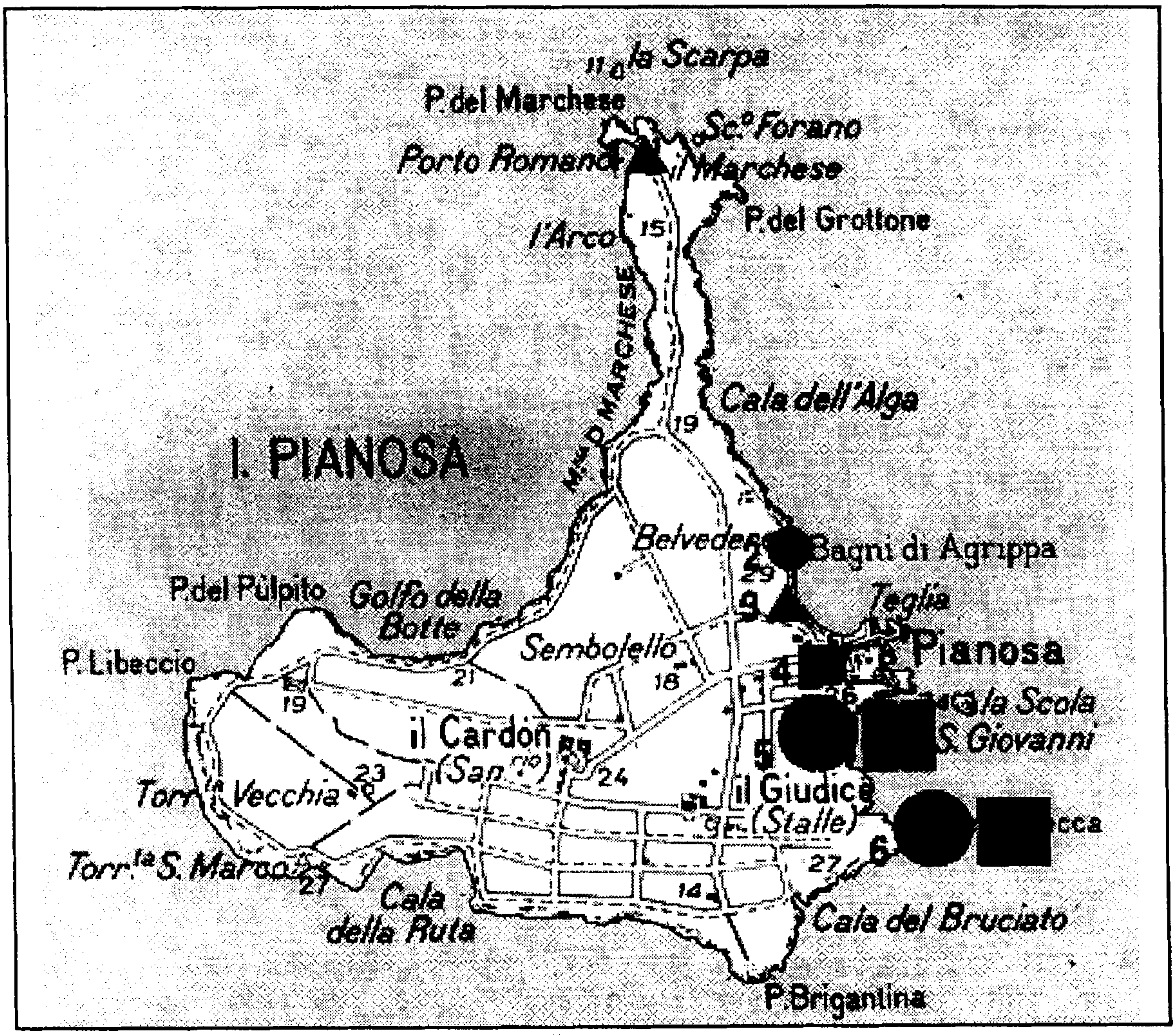

Fig. 12 - An example of map/classification paradigm

worthwhile to note that guided tours can be useful to the user to become familiar with the hypertext's structure and content.

Conclusion

Many hypertext applications pay much attention to mere technological aspects, while disregarding other important issues, like a consistent and expressive user interface, or the usefulness of an accurate data analysis.

© Archives & Museums Informatics, 1995

In this paper we showed as the links among the nodes in a hypertext give the possibility of representing the associative mechanisms that undoubtedly constitute a point of strength of this kind of systems. Distinguishing between extensional and intensional links leads to the definition of a hypertext design methodology that makes use of consolidated approaches in the database and information retrieval areas and allows a clear identification of the two different types of links.

The need for an effective and expressive user interface makes evident the importance of the availability of multiple interaction paradigms, that the user will select at his/her will, switching from one to the other according to the associative mechanism that will appear to be the most adequate.

References

[Bush45] Bush V.:*As We May Think,* in Atlantic Monthly (July 1945), 101-108. Republished in S. Lambert, S. Ropiequet (Eds), "CD-Rom: the news Papyrus", *Proceedings of the First MS CD-ROM Conference*, Microsoft Press, Redmond,WA, 1986.

[Conklin87] Conklin J.: *Hypertext: a survey and introduction,* IEEE Computer, September 1987.

[DeRose89] De Rose S. J. (1989), Expanding the Notion of Links, Hypertext ' 89 Proceedings, (November 1989), 249 - 257.

[Nielsen90a] Nielsen J.: *The Art of navigating through Hypertext*, Communications of the ACM, Vol. 33, N. 3 (March 1990)

[Nielsen90b] Nielsen J.: *Hypertext and Hypermedia.,* Academic Press inc., Harcourt Brace Jovanovich Publishers, USA 1990.

[Norman88] Norman D. A.: *The psychology of everyday things,* Basic Books, Inc., Publishers, New York (1988)

[Parsaye89] Parsaye K., Chignell M., Khoshafian S., Wong H.: *Intelligent databases.,* Ed. J. Wiley and Sons, Inc., 1989.

[Signore90] Signore O., Bartoli R.:*Implementation of a historical/geographical database with support of imprecise dates*, DEXA'90: Database and Expert Systems Application, Proceedings of the International Conference in Vienna, Austria, 1990, (Tjoa A.M., Wagner R., Eds.) Springer Verlag Wien New York, ISBN 3-211-82234-8, pp. 271-274

[Signore91] Signore O., Aulisi R., Ceccanti V.: *Hypertext for Hypertext: a Figured Thesaurus*, DEXA 91: Database and Expert Systems Application, Proceedings of the International

© Archives & Museums Informatics, 1995

Conference in Berlin, Germany, 21-23 August 1991, (D. Karagiannis., Ed.), Springer Verlag, Wien-New York, ISBN 3-211-82301-8, pp. 514-519.

[Signore93a] Signore O., Malasoma S., Tarchi R., Tunno L., Fresta G.: *A Hypertext with Multiple, Interchangeable Paradigms of Interaction and Dynamic Links*, The Art of Communication, VIII International AHC-Conference, Graz, Austria, August 24-27, 1993, Abstracts, pp. 75-77

[Signore93b] Signore O., Malasoma S., Tarchi R., Tunno L., Fresta G.: *A hypertext for an interactive visit to a science and technology museum*, Proceedings of the Sixth International Conference of the MDA and the 2nd International Conference on Hypermedia and Interactivity in Museums (ICHIM'93) Cambridge, England, 20-24 September 1993, ISBN 0-905963-89-X, pp. 204-209

[Signore94] Signore O., Fresta G., Cimbalo A., Di Lecce S.: *A Methodology for Hypertext Design: a Case Study*, Proceedings of Structures and Contingencies in Computerized Historical Research, IX International Conference of the Association for History and Computing, August 30-September 2 1994, Nijmegen (The Netherlands)

[Utting89] Utting K., Yankelovich N.: *Context and orientation in Hypermedia networks.*, ACM Transactions in Information Systems, Vol. 7, No. 1, January 1989, Pages 58 - 84.

© Archives & Museums Informatics, 1995