HYPERTEXT AND THE WORLD WIDE WEB: PROBLEMS AND PROPOSED SOLUTIONS

DAVID RIBEIRO LAMAS

ABSTRACT

Today, the World Wide Web is an all-inclusive information system in which any piece of information which is accessible as part of a seamless hypertext information space. Such a large volume of information has an inherent application potential but, unfortunately, besides inheriting and enhancing most of hypertext's qualities, the Web also incorporates and stresses common hypertext problems. It is the Web's disorganised nature that mainly prevents its most effective use and, as such, research is needed to develop the means that will help to transform the available data into meaningful information able to support productive activities. This document proposes an adaptive navigation support as a way of increasing the Web's value specifically addressing problems such as disorientation, cognitive overhead and information overload.

RESUMO

Hoje em dia, a World Wide Web é um sistema de informação de âmbito global que faz com que toda a informação nele disponível parece uma parte de uma espaço de informação praticamente sem limites. Tal espaço de informação tem um potencial quase ilimitado. Infelizmente a Web não implementa só os aspectos positivos do hipertexto mas também traz consigo muitos dos seus problemas. É a natureza desorganizada da Web que mais prejudica a sua utilização eficiente e assim é necessário perceber como se podem transformar todos os dados disponíveis em informação útil e com sentido. Este documento descreve os problemas principais e termina com a proposta de uma linha de investigação para lidar com os problemas tais como a desorientação, o esforço cognitivo e a sobrecarga de informção.

KEYWORDS

Hypertext, world wide web, problems, proposed Solutions

INTRODUCTION

Instruments are at hand which, if properly developed, will give man access to and command over the inherited knowledge of the ages... wholly new forms of encyclopedias will appear, ready-made with a mesh of associative trails running through them (Bush, 1996)[1]

Hypertext as a concept and the *World Wide Web*, also known as the *WWW*, the *Web*[2] or the *W3* are instances of this vision.

HYPERTEXT

There are several not so convergent or overlapping (Conklin, 1987b; Nielsen, 1995b; Woodhead, 1991) hypertext[3] definitions in the literature but, in the end, the basic characteristic is the absence of an explicit order or sequence in which to read or retrieve information.

If traditional text is *sequential*, hypertext, on the other hand, is *nonsequential ie*, whereas with sequential text there is an explicitly set linear way of reading it, no such thing exists in hypertext.

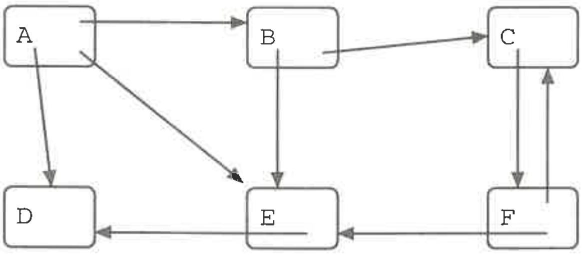

Figure 1: Simplified view of a small hypertext structure having six nodes and nine links (Nielsen, 1995b)

---

[1] this is a reprint from the July, 1945 The Atlantic Monthly original and it can also be found at http://www2.theAtlantic.com/atlantic/flashbks/computer/tech.htm

[2] the term *Web* has other common connotations such as (Rein et al., 1997): the physical network of servers and gateways, the actual information content contained in the distributed web of links and nodes, the internet technologies as an operating system or computing environment with its tools, or the internet interface but within this document, except if otherwise stated, the meaning is the one prescribed by Berners-Lee et al. (1995)

[3] word introduced by Nelson (1981)

Typically, hypertext is made out of *nodes* or chunks[4] of information connected by *links* (see figure 1 above) in a variety of possible topologies such as *webs*, *meshes* and *hierarchies*.

Following a path through the nodes, traversing the links that connect them, is called *browsing* (Woodhead, 1991).This is the atomic search activity in a hypertext.

Navigation, on the other hand, is composed out of a sequence of browsing steps complemented, in some way, by the use of abstract features or structures for general orientation or by other cues and tools that direct a change of focus or motion in the hypertext structure (Woodhead, 1991).

As presented in table 1, hypertext has already a long history of conceptualisation, research and development which started with *Memex* (Bush, 1996) about 50 years ago and was massified with the World Wide Web.

Year	***System***	**Highlights**
1945	Memex	"... a sort of mechanised private file and library ... in which an individual stores his books, records, and communications, and which is mechanised so that it may be consulted with exceeding speed and flexibility..."
1962	Augment/NLS[5]	Though not developed as an hypertext system, it was the first major work in office automation and text processing with an hypertext structure that supported cross-referencing among stored documents
1965	Xanadu	Intended to be a repository for everything that anybody has ever written in which it is possible to address any substring of any document from any other document. This system would also be able to keep a permanent record of every document's version enabling temporal scrolling useful for version management and critiquing
1967	Hypertext Editing System	The first working hypertext system to provide linking and jumping to other documents with a text oriented interface
1968	FRESS[6]	A follow-up of Hypertext Editing System

[4] term used by Woodhead (1991) to convey the idea that hypertext nodes can be clusters of multiple media and information and not only *simple* pieces of text
[5] oN-Line System
[6] File Retrieval and Editing System

Year	*System*	**Highlights**
1975	Zog	A research system with linked frames of on-line text field tested in applications such as a maintenance manual for weapons elevators
1980	Enquire	A notebook program which allowed bidirectional typed links to be made between arbitrary arbitrary nodes
1983	KMS[7]	A direct descendant of Zog designed to manage large hypertexts and to work across local area networks, base on a single type of node and optimised for speed of navigation. It provided a home frame as a referential starting point and a special action language
1983	Hyperties[8]	This system supported both text and graphic interfaces and provided users with prospective views of what would happen if links are followed
1985	NoteCards	A hypertext system with about 50 types of nodes linked with user typed links which provided structural overview diagrams and hierarchical node nesting .
1985	Symbolics Document Examiner	Symbolics Document Examiner & An interface to the on-line documentation of the Symbolics workstation written with Concordia[9]
1985	Intermedia	Based on the scrolling window model also provided a linking protocol enabling other applications to be linked to and from Intermedia documents
1986	Guide	Supports replacement, pop-up and jump links which are revealed to the user by changing the cursor's shape. A later release also supports scripting and provides it as fourth link type
1987	Hypercard[10]	Supports read-time information generation through the use of a general programming language. Its nodes are cards and links can basically take the reader to anything the system can compute

---

[7] Knowledge Management System
[8] originaly named TIES as an abreviation for The Electronic Encyclopedia System
[9] a system that provided authors with slotted templates for structured text production (Walker, 1988)
[10] other compatible products exist such as *SuperCard*, *Plu*, and *MetaCard* which expand in a way or another Hypercard's basic features

Year	System	Highlights
1990	ToolBook	This package supports most of the accepted hypertext structuring metaphors such as interactive maps, text and graphic cards, buttons and backtracking and also includes a scripting language[11].
1991	World Wide Web	Conceived to be a pool of human knowledge it is the first global hypertext
Today		Hypertext features are present in most contemporary desktop applications from text editing systems to electronic mail tools and drawing packages

Figure 2: Hypertext development overview[12]

Today, hypertext features can be found virtually on every personal computer's desktop provided either by applications such as word processors and personal information managers, by smaller programs like electronic mail handlers or by the operating systems themselves[13] which are just about starting to integrate local with global networked information space browsing.

As a medium, either standing alone or on a networked environment, hypertext's[14] applications are at least as diverse as any other medium and include (Nielsen, 1995b):

Computer applications from system prototyping to on-line documentation and operating systems' user interface;

Business applications from procedure manuals to dictionaries, reference books, product catalogs and advertising materials;

Intellectual applications from idea organization and brainstorm support to journalism, research and scientific publications;

---

[11] see http://www.assimetrix.com
[12] there are a number of other hypertext applications that were left out not because they are less relevant in any way but because the goal is just to provide an overview on hypertext's evolution history
[13] after Sun's first attempts with the *Answerbook*, the first attempts in this direction have started to show-up in wide use operating systems such as *Microsoft Windows* and *Mac OS*
[14] used here as suggested by Nielsen (1995b) where it also includes multimedia able hypertext systems otherwise named *hypermedia* systems

Educational applications from teaching foreign languages to museum guides and history tours; and

Entertainment applications from tourist guides and interactive fiction to news, newspapers and magazines.

Other classification taxonomies exist that classify hypertext systems from a functional point of view such as (Conklin, 1987b):

Macro literary systems - Technologies that support large on-line libraries in which interdocument links are machine supported.

Problem exploration systems - Which embraces tools to support unstructured thinking on a problem when many disconnected ideas come to mind.

Structured browsing systems - Similar to macro literary systems but smaller in size.

General hypertext technology - General purpose systems designed to support a variety of hypertext applications.

THE WORLD WIDE WEB

The World Wide Web is a networked and distributed hypertext system implemented over the Internet, that merges a number of technologies resulting in a global information system in which any information accessible over the network is perceived as part of a seamless hypertext information space (Berners-Lee, 1992).

It was originally developed to be a pool of knowledge, for the physicists and engineers at CERN[15], that would enable and promote the sharing of ideas and information related to collaborative projects (Berners-Lee et al., 1992). Its evolution is shown in the chronogram in table 2.

---

[15] the European particle physics laboratory

Year	**World Wide Web sequel**
1980	Enquire-Within-Upon-Everything notebook hypertext program supporting bidirectional typed links between arbitrary typed and titled nodes
1989	"Hypertext and CERN" and "Information management: A proposal" circulated for comments at CERN
1990	the proposal was reformulated and by the end of the year, a line mode browser and a NeXTStep browser were demonstrable providing hypertext files, Internet news and specific CERN systems access
1991	general WWW release on central CERN machines and anonymous Telnet service started
1992	first X Windows browsers available together with a list of 26 reasonably reliable servers
1993	Mosaic releases Web browsers for all common platforms and more than 200 HTTP servers are on-line
1994	first international WWW conferences and first World Wide Web Consortium meeting - CERN transfers the WebCore project to INRIA
1995	the Web Society is founded in Graz
1996	almost every potential information publisher offer some sort of World Wide Web access to their repositories
1997	core WWW specification move towards standardisation

Figure 3: WWW Time line

Today, the World Wide Web's architecture enables a highly advanced information system using the Internet as its infrastructure. As an open system, it includes, in its data model, ways of integrating most of the information available on the Internet.

PROBLEMS

As shown before, the World Wide Web merges the concepts of hypertext and networked information to provide a powerful global information system-

Apart from the relationships established before between previous hypertext systems and the World Wide Web, viewing the WWW as a *traditional hypertext system* may be a mistake.

The Web's heterogeneity and its cross-domain links and dynamic nature mean, in most cases, that many assumptions connected with typical hypertext systems do not apply.

Hypertext research has primarily focused on a single document or set of related documents converted to hypertext by a single individual or team but, while the WWW bears some similarity to classic hypertext systems, there are important differences:

Links across documents and sites[16] - World Wide Web links, as opposed to most of the other hypertext systems links, do not know any boundaries. While in most of the systems presented in table 1 links stay within a specific environment, context or document, WWW links bring together almost every bit of information available in the Web.

Repeated or missing information - While within a stand-alone hypertext system, information concerning a given topic either exists in exactly one place or it doesn't exist at all, in the Web, the same subject can be addressed in several places in a variety of ways.The consequence is that information can be redundant, contradictory or incomplete.

Constant change - Again, unlike classic hypertext systems, the WWW is constantly changing both in content and in structure which means that information is endlessly being added, deleted and moved. While this temporal dimension is exactly what keeps the Web alive, it is also what prevents any static approach to content description in terms of maps, tables of content or other overview techniques.

These are not, however, the only differences and Shun (1996) highlights others:

Speed - As the WWW is a global and networked hypertext infrastructure, users' frustration related to bandwidth problems is a major problem specially when dealing with graphics, audio and video.

Interactivity - Simple text and images are fine for publishing a wide variety of information and it is this simplicity that enabled the Web to be as successful as it is today. Nevertheless, regarding interactivity, the WWW is still far from most current standards.

Heterogeneous end-user environment - As everything is available to everyone, interface design cannot keep on relying on *a-priori*

---

[16] a Web site is a set of pages among which some special relation holds

assumptions regarding user profiles or hardware and software features.

From these remarks it may seem that the World Wide Web is something to be avoided. That is surely not the case as what must be done is to *know how to use it effectively and efficiently.*

As frequently happens when dealing with large problems, the approach here is to break it down into smaller and more manageable issues and to deal with them in a well defined context.

CLASSICAL HYPERTEXT PROBLEMS

As listed by Nielsen (1995b), usability is associated with several attributes:

- Easy to learn;
- Efficient to use;
- Easy to remember;
- Few errors; and
- Pleasant to use;

Nevertheless, these parameters cannot be taken literally without further refinement. When discussing hypertext usability,a distinction must be made between *the usability hypertext system* and the usability of the hypertext information content and structure.

As some hypertext systems were built as vertical solutions, in a number of cases it is difficult to separate the two constituents.

Navigation

Navigation is both the greatest strength and the greatest weakness of hypertext (Jonassem, 1992; Nielsen, 1995b, Rivlin et al., 1994a) but overall it is a positive attribute because it provides opportunities for incidental learning since users can explore the information space and find information which they actually would not be able to request using a formal query.

In one of the earliest surveys on hypertext, Conklin (1987b) identifies two broad classes of problems with hypertext:

Implementation problems - Including visualisation delays, restrictions imposed by the underlying infrastructure, each of node and link properties and browser deficiencies or lack of functionality.

Endemic hypertext problems - Including *disorientation* and *increased cognitive overhead.*

As the former class either depends on available technology or reflects design considerations, attention will be on the latter.

Disorientation in hypertext can be defined as the tendency for the user to lose her sense of location and direction in a nonlinear document.

The disorientation problem also exists in traditional linear text but, in this case, the reader only has the option to search the desired information either earlier or later in the text. With hypertext, a variable number of dimensions or degrees of freedom is available, resulting in a greater potential for the users to become lost or disoriented.

With hypertext, to be lost or disoriented means that the users:

- Do not know where they are in the information structure;
- Do not know how to get to some other location in the hypertext which they know, or believe to exist:
- Cannot return to previously visited information;
- Do not manage to remember the key points covered and sometimes even forget what they are looking for.
- The hypertext cognitive overhead is the additional effort and concentration necessary to maintain several tasks or trails at one time. It mainly results from the large number of choices about which links to follow and which to leave behind. It usually gives rise to questions such as (Conklin, 1987b):
- Is following the suggested path worth the distraction?
- Is the label or any other information associated with the link enough to make such a decision?

This problem, also named *informational myopia*, is an overhead that is absent in linear text where the author has already made most of these choices.

Disorientation and cognitive overhead constitute the navigation problem, and results reported by (Shum, 1996) provide added evidence that it was already an important issue with classical hypertext systems embracing items such as:

- Enabling user orientation;
- Supplying navigation mechanisms;
- Contributing to a sense of context and location; and

- Providing search devices.

Approaches to address these problems comprise *guided tours*, *common sense metaphors*, *visual clues*, *hotlists*, *maps*, *fisheye views* and *query or search mechanisms*.

Other problems

Apart from navigation related problems, others were reported such as (Woodhead, 1991):

- Creating and maintaining consistent annotations among related pieces of information;
- Transforming existing information into hypertext structures as a way of incorporating the vast quantity of available data in traditional linear form;
- Handling a large amount and variety of information;
- Supporting many types of users with different skills and information needs;
- Using goto links since they are a major source of disorientation;
- Dealing with presentation constraints such as screen size or scrolling windows (the chunking problem); and
- Conveying subject boundaries when links exist to nodes on another subject (the framing problem).

Most of these fits within the category of implementation problems, briefly mentioned before while others, like the framing problem, can be addressed by authoring techniques and careful design of hypertext content and structure.

World Wide Web specific issues

The World Wide Web affected and affects the *modus operandi* of many people in areas that range from scientific work to business, including education and entertainment and Web usability problems affects many users today.

The fact that the Web is currently perceived as an accessible medium in which anyone can create content results in a community of developers where the majority of authors are new to system development and have little or no knowledge or understanding of human-computer interaction issues.

Further, poorly designed hypertexts are also cited as a problem source since, just as early programmers lacked experience, so do the majority of current Web authors who are, among other things, said to be responsible for:

- Poor or inadequate graphical design;
- Overuse of technology in pursuit of clever effects but with no attempt to redesign for the new medium;

- Lack of an overall structure and presentation style.

As a result, most of the effort is invested in making hypertexts look *technically good* without much careful consideration being given to the information organisation or to node and link modelling.

The consequence is that many Web sites are developed by multiple authors with different goals which tends to result in incremental expansion of their information space without coordinated planning.

Still, the early mixture of grey text and large images that characterised the Web has changed and better layout capabilities are now available. Many WWW authors have now learned to minimise their use of graphics and to make better use of the available technology, saving download time and slightly increasing usability.

However usability problems endure. Some of them identified by Levi and Conrad (1997), such as *unclear labelling* and *unreasonable memory demands on the user*, are also associated with conventional software whereas others like *cluttered interfaces*, *overused graphics* and *inadequate links* to other Web nodes are definitely more characteristic of the World Wide Web.

An approach to classify the variety of Word Wide Web usability problems is also proposed by Levi and Conrad (1997) who identify seven broad Web usability problem categories:

Management and maintenance of broken links and obsolete data;

Technology constraints imposed by the underlying software and hardware infrastructure;

- Available navigation mechanisms and hazards such as back buttons, dead ends and scrolling;
- Matching the user's mental model;
- Content adequacy;
- Mismatch between the user's and designer's goal or intention; and
- Inconsistent Web page layout.

Moreover WWW development differs from conventional software usage because the identity of the users can be difficult to establish and consequently, the *know thy user* design principle is hard to apply.

PROPOSED SOLUTIONS

If on one hand, the Internet enables the existence of a world wide hypertext system, the World Wide Web, which provides access to a considerable amount of continuously evolving information, unfortunately, on the other hand, World Wide Web navigation is not an easy task and research on ways to maximise information navigation quality without interfering with the WWW is needed (Lamas, 1998).

Maximising the quality of information navigation means reaching the reader's exploratory goals while avoiding common World Wide Web problems such as *disorientation*, *increased cognitive overhead* and *information overload*. Further, maximising information navigation quality means providing some sort of *weak hypertext* linearisation in order to allow the user to follow a line of thought while reading an optimised sequence of WWW pages.

In order to achieve such a goal, at least three main issues must be addressed:

Adaptive navigation - In this case, adaptive navigation deals with improving the goal directed interaction between the user and the Web by providing an adaptive guidance device that reacts to the evolution of the information base and to the evolution of the user's interests, preferences and experience based on previous interaction sessions. Several studies (Espinoza and Hook, 1996) identify the user's background knowledge and spatial ability as major influences on understanding a given domain.

Social navigation - This is the process by which users share their experiences. It goes beyond traditional posting and alerting approaches providing a more specialised mechanism for sharing successful browsing experiences and feedback information that enables a common information awareness or in other words, a community wide memory.

World Wide Web modeling - In order to provide adaptive and collaborative navigation, a model of the Web must be conceived that goes beyond bookmarks or link collections andholds all the necessary information enabling a systematic account of its contents and qualitative attributes in a way similar to traditional knowledge representation techniques such as semantic networks or frame-based models.

This proposed solution does not, however, pretend to be a solution for every situation and its aimed context is the support of learning or research tasks within common-context research groups or student teams.

Although the application of this line of research to other similar contexts such as general project teams or common interest user groups, such as shoppers in e-malls, will eventually be addressed, Web related tasks such as generic searches or other browsing activities are currently beyond its scope.

Adaptive navigation

One way to provide such direction is through the use of adaptive navigation techniques used in some hypertext systems and identified by Brusilovsky (1996) as being able to provide features including:

Direct guidance - Usually in the form of a *best link* selected by the system from the range of links available on the current hypertext node or from the set of notes available in the information space. This selection is usually based in some user model or profile and presented as a *next* or *best choice* button.

Adaptive annotation - Annotations can be provided as textual or visual clues and the most common example is the colour change observed on visited links on a Web page. These add information to the links in order provide individualised information about the document on the other side of the link.

Adaptive link re-reordering and hiding - The former is a technique that sorts the available non-contextual links on the basis of a user profile whereas the latter hides or disables less relevant links and applies to all sorts of links, contextual or not.

However, in order to enable adaptive navigation, apart from modeling the user, information must be available to the system about the underlying information space.

Social navigation and Web modeling

A complementary approach to providing a friendly environment for performing learning or research activities over the Web is to use the research in *Computer Supported Cooperative Work* which provides numerous interesting findings about groups of cooperating individuals, about the process of cooperation and about phenomena like group awareness.

Related results have already been adopted in areas such as database browsing and education and recently begun expanding over the WWW as it is highly suitable for collaborative working. Current related applications include:

Active annotations - Active annotations allow each and every word of any WWW page to be annotated and merge the asynchronous behaviour of reading a Web page with a synchronous notification system that informs all viewers of the same page about new annotations.

Social Web navigation - Support for asynchronous, semi-synchronous and synchronous collaboration on the Web.

Shared workspace - A shared information repository which supports the concept of a user community and enables the group to organise and coordinate their work.

Collaborative techniques should thus be explored as a way of gathering the information needed to enable adaptive navigation on the World Wide Web while the research on Web content description, should be the basis for the necessary account of the Web's known and relevant contents, structure and qualitative attributes.

FINAL REMARKS

This proposed line of research does not mean imply that *Web-surfing is dead* but, Instead, tries to define an initial standpoint to address Web usability problems. It is thus our beleive that:

- It is possible to provide a meaningful record of successful Web navigation experiences in the form of a restricted domain model;
- It is possible to model users by representing their individual preferences, their relevant background domain knowledge and its relative persistence and their relations with other users; and
- It is possible to possible to provide a set of rules to combine the relevant Web navigation record and user model information enabling effective navigation guidance.

This approach is a reactive one as the effort goes into providing a way to deal with the vast amount of useful information available on the WWW as is, rather than to try to improve or alter the Web's infrastructure in any way.

This reactive top down stand point is, however, not taken lightly because other approaches that tackled the problem from other angles were also considered. However, all the pro-active bottom up alternatives seem to reduce the Web's successful characteristics resulting in either a decrease in hypertext production, a bypass of publication rules or a to complex publication procedure.

Bibliography

Conklin, J.: 1987b, Hypertext: an introduction and survey, IEEE Computer, 20(9), 14-41.

Berners-Lee, T.: 1992, An executive summary of the world-wide web initiative [online], Avaliable at http://w3c,org/Summary.html [Accessed: 11 March 1998].

Brusilovsky, P., 1996, Methods and techniques of adaptive hypermedia, *User Modeling and User-Adapted Interaction*, 6, 87-129.

Bush, V.: 1996, As we may think, Interaction 3(2), 35-46.

Espinoza, F. and **Hook, K.**: 1996, An interactive WWW interface to an adaptive information system [online}, Available at http://www.sics.se/~kia/espinoza_hook.html [Accessed: 12 April 1998].

Jonassen, D.: 1992, Designing hypertext for learning, *New direction in educational technology*, Vol. 96 of NATO ASI Series, Springer-Verlag, Berlin, Germany.

Lamas, D. 1998, Adaptivity on the Internet, *University of Portsmouth research seminars*, University of Portsmouth, UK.

Levi, M. and **Conrad, F.**: 1997, Usability testing of World Wide Web sites: a CHI 97 workshop, *ACM SIGCHI Bulletin* 29(4), 40-43.

Nilesen, J.: 1995b, *Multimedia and hypermedia: the internet and beyond*, Academic Press, London, UK.

Rivlin, E. et **al.**: 1994a, Navigating in hyperspace: designing a structure-based toolbox. Communications of the ACM 37(2), 87-127.

Shum, S. 1996, The missing link: hypermedia usability research & the Web, *ACM SIGCHI Bulletin*, 28(4), 221-248.

Woodhead, N.: 1991, *Hypertext and hypermedia: theory and applications*, Addison-Wesley, Wimslow, UK.