

Appeared in the *Proceedings of the WebNet 2001 – World Conference on the WWW and Internet, 2001.*

Navigation is Law: The Conceptual Architecture for Layered Web-based Information Systems

Thawatchai Piyawat, Jantawan Noiwan

Department of Business Administration
Faculty of Management Sciences
Prince of Songkla University
Hatyai, Songkhla 90110 Thailand
E-mail: {piyawat, noiwan}@umbc.edu

Anthony F. Norcio

Department of Information Systems
University of Maryland Baltimore County
1000 Hilltop Circle
Baltimore, Maryland 21250 USA
E-mail: norcio@umbc.edu

Abstract

Navigation capability is considered as the most important facet of large information space since it is the indicator identifying the ease of locating information. Therefore, the navigation structure of information space must be well designed and instructed to minimize learning requirements and maximize efficiency of use. This paper proposes a new hypertext system called the Retriever information system in which it is designed to promote ease of navigation in wide information space. The Retriever introduces separation of the navigational content and the actual content of documents. It employs validation of navigational content to ensure the correct integration of documents in the Retriever navigational structure. Currently, the Retriever is planned to be a layer on top of the current Web infrastructure to raise the navigation capability of the World Wide Web.

Keyword: Navigation, Information Retrieval, Web Usability

1 Introduction

Undoubtedly, the World Wide Web is expected to be the medium to drive the global economy of the 21st century. With its tremendous accessibility, it tears down the world's trading boundary and brings people from around the globe to meet as they are sitting next to the others. Nevertheless, because the organization of Web contents is not designed in the first place for dealing with large amounts of information, the structure of the Web obstructs itself from showing its highest possible benefits.

The Web structure employs the concept of directories, subdirectories, and filenames to construct Uniform Resource Locator (URL), which is used to allocate a document called a Web page. However, without a required method to emphasize the logical construct of Web contents explicitly, users easily lose their sense of the position they currently view on Web sites. In addition, its filename-based structure does not give necessary information for users to recall the address of Web pages they have visited. Furthermore, currently most Web pages are dynamically generated, and URLs are employed as a means to convey information from one page to another. This situation consequently disables the meanings of URL, which is the only required tool for navigating Web content outside the content itself.

There are a number of studies presenting methods for solving this problem. Nevertheless, all available solutions to overcome the navigational difficulty are based on the current structure of the Web. This paper considers the Web navigational problem laid on its

primary foundation in which the only way to resolve is to redesign its basis. Therefore, this paper presents a new hypertext system entitled Retriever Information System (Retriever) that is designed to employ in a large information space, similar to the Web. In this paper, the concept of Retriever and the fundamentals it is built upon are presented.

2 Background

The structure of the World Wide Web employs a tree-like structure. It starts from the “root” directory, which is referred to as its beginning point of a Web site. Then, using filenames and subdirectories from the root directory, a URL (Uniform Resource Locator) is formed to specify an address of a Web page. However, the length and complication of an URL are unpredictable since they are solely dependent on the decisions of Web designers. Therefore, the URL of a Web page does not give predictable information for users to recall and allocate the page on the Web space.

Furthermore, since a Web page is a hypertext document, it can link to an unlimited number of other Web pages by referring URLs of the others. The links, with the loose structure of the Web, create a chaotic network of Web pages in which its structure cannot be visually perceived. Without an imaginable structure, users can easily lose their sense of the position they visit on the Web space.

Acknowledging the problem, several papers present solutions that are aimed to raise Web navigational performance. For instance, Lynch and Horton (Lynch & Horton, 1999) assist Web designers to organize Web information into units and arrange them to be Web sites. They also provide some basic information structures that can be applied to the organization of Web sites. Spool, et al. (Spool, Scanlon, Schroeder, Snyder, & DeAngelo, 1999) provides design tactics to improve site navigation. Whitaker (Whitaker, 1997), admitting the unstructured environment of Web sites, suggests ten principles of navigational design.

Nielsen (Nielsen, 2000) emphasizes site design as a primary part of the success of Web sites, although the site structure is not explicitly presented to users. He provides advice for site developers to improve navigational performance by designing the site structure according to users’ tasks and goals. He also reviews different styles of site designs of existing Web sites. Nevertheless, he admits that Web browsers need to provide better tools for structural navigation. Additionally, he proposes the concept of “active sitemap,” which spots the user’s current position and gives the user’s trail of traveling throughout the site. This concept indicates the need of navigational structure similar to the Retriever discussed in Section 2. Nielsen also shows his concerns about the complicated nature of URL structures, which can be considered as the only tool required for aiding Web navigation outside the Web content.

Methods presented in the literature give guidance on how to minimize problems caused by the deficient nature of the Web structure. Following those methods, Web designers are responsible to cope with the problems by themselves. As the result, navigational quality of Web sites is varied according to expertise of their Web designers. Therefore, to improve benefits of the Web as a whole, the Web structure needs to be redesigned with an understanding of the problems and being aware of other potential problems that may arise when applied to larger information space.

3 Navigation is Law

3.1 Overview

Navigation is the most important facet for information space. In the large information space such as the Web, users navigate throughout streams of information looking for a piece of information at a time. Generally, users spend their time searching for information than using the information. Therefore, the navigation capability is the ability to minimize time used to locate information.

Navigation is considered as the law of an information space. Users learn navigation law and search for information according to the law. Normally, with less restricted law, navigation difficulty occurs, and, with too strict law, flexibility is decreased. Hence, the primary task of creating an information space is to design the law of navigation that aids users to use the information space effectively.

The importance of navigation may not primarily be recognized in designing the Web in the first place. It results navigation structures of Web sites can be different from one site to another. This causes users to relearn a new navigation structure every time they visit a new Web site. The navigational differences are unnecessary information that overloads the cognitive memory of users for the Web space and derives difficulties to use the Web in general. These difficulties can be imagined as to drive a car to different towns that have their own unique traffic systems.

This paper considers that the Web navigational deficiency is located at its fundamental level of the Web. Therefore, it introduces the Retriever that is designed to emphasize on the navigational capability. Currently, the Retriever aims to implement as a layer on top of the current Web to promote navigation efficiency. The design of the Retriever system follows the heuristics of designing a navigational structure as proposed in the following section.

3.2 Heuristics for Designing Information Navigation

The navigational architecture of an information space is considered in this paper as the most important part that contributes to the ability of users to search for information in the space. This paper proposes heuristics to create a navigational structure such that a navigational structure must be designed to (1) *minimize learning requirements*, and (2) *maximize efficiency of use*.

There are several issues to consider in minimizing learning requirements. Topics that are related to the design of the Retriever are discussed here. The first concern is the *level of navigational consistency across the entire information space*. The Web shows minimal level in this issue. For the Web, it is frequently found that users must examine pictures and texts on a Web page whether or not they are links, as links can be differently customized in each Web site. The second issue involves the *complexity of navigational structure*. Intuitively, complex navigational structures are harder to learn than the less complex ones. To overcome this issue, the third requirement for ease of navigation, i.e., the *predictability of navigational behaviors* must be well implemented. The Web contains low ability in this topic. For the Web, it is almost unpredictable about the changes of navigational structure when users click at links. Furthermore, users are frequently required to learn new navigation structures when they navigate from one Web site to another. The Retriever attempts to overcome these issues by separating navigational content from document content.

Issues for maximizing efficiency of use cover several topics. The issues that are relevant to the design of the Retriever are discussed here. The first item includes *flexibility for content arrangement* in which requirements for navigation should not prevent content designers to arrange information according to their intentions. However, the requirements must enforce in a certain level to ensure the integration of navigational ability across the information space. The second issue is the *availability of information*. It involves steps users are required to accomplish in order to access a piece of information. URL shows a great example in this issue such that it needs only one step to access information, as it is a complete point to the information. Nevertheless, the availability of information must balance with the third issue, i.e., the *accessibility of information*. This topic deals with sequence to access a piece of information as it often that, in order to maintain its meaning, information needs to access according to its sequence. The Retriever employs URL to specify locations of documents in conjunction with access flags to indicate the access methods required for documents.

As navigation is considered the main issue of a large information space, this section discusses heuristics for designing a navigational structure in which they are employed to design the Retriever information system, which is presented in the following section.

4 Retriever Information System

4.1 Overview

The design of the Retriever is largely influent from the Web and the Gopher information system. The Gopher is a hierarchical information system that is predated the Web. The Gopher, unlike the Web, restricts its navigational structure. Users navigate within the Gopher's hierarchical path in order to retrieve information, and users are not able to jump from one node in the hierarchy to another that has no explicit hierarchical link with the previous one. With less capability on enriching user interactions than the Web, the Gopher is now considered being obsolete, although there is still an effort to promote the use of the Gopher, such as Gopher Manifesto (Karger, 2000).

The Retriever is an attempt to combine the explicit navigational structure, as in the Gopher, with information presentation and the hypertext ability of the Web. Ease of navigation is the main concern on designing the Retriever.

4.2 Structural Concept of Retriever

In the Retriever information space, documents are arranged in the hierarchical fashion as illustrated in Figure 1. The Retriever aims to promote the well organization of hypertext documents by implementing a number of restrictions. A Retriever document consists of *navigational content* and *information content* parts. In this current stage of development, the navigational content is primarily studied, and the information content at this time is specified only as it contains texts and links to other documents. Additionally, each part is planned to display to users in different areas to prevent the fusion of the user perception of navigation and information contexts. The visualization methods to present both parts to users are currently under study.

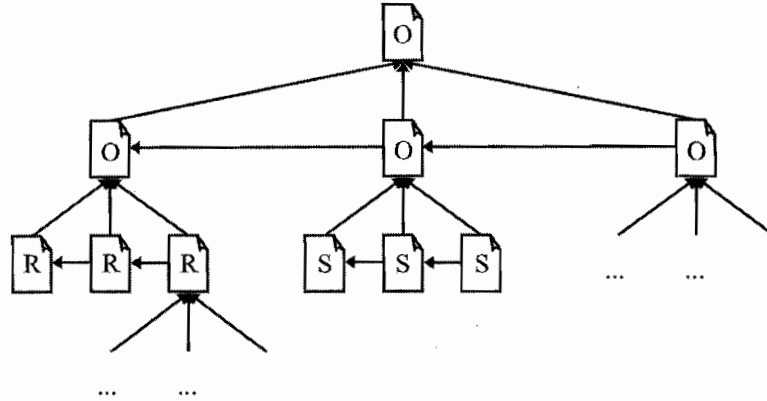


Figure 1: Logical structure of documents in the Retriever

The navigational content of a document consists of several items to aid efficient navigation in the Retriever information space. All items are required in order to produce a valid Retriever document. As the Retriever employs strict rules to prevent navigational difficulties, an invalid Retriever document is not allowed to display.

Navigational content of each document is unique. Thus, it also serves as the full address of the document. As a document is a node of the Retriever hierarchical structure, navigational content of a document contains (1) document handle, (2) document title, (3) access flag, (4) parent node address, (5) previous node address, and (6) list of links in the document. For the previous node address, the null value is used when a document is the beginning of a sequence. However, every document must have its parent node address. For the root document of a site, its parent node address is the root document of the Retriever information space. The document handle of a document is its URL. The access flag of a document can be O (open), R (root-protected), or S (sequence-protected). The open documents can be accessed from any documents in the Retriever space; the root-protected documents must be entered from their parent documents only, and, the sequence-protected documents must be only accessed from their preceding documents. The root-protected and sequence-protected documents are designed to support several application potentials, e.g., category explanations or checkout procedures of e-commerce sites.

The navigational content of a document, which is also used as the address of the document, contains two addresses, namely, its parent node, and its previous node. This requirement results accumulation of addresses in which one address of a document encloses the address of another document that in turn contains the address of another document and so on. Hence, as a document resides deeper in the hierarchy, its navigational content tends to be exponentially larger than the shallower one. As a result, the Retriever as a whole consumes more communication bandwidth than does the Web or the Gopher. Nevertheless, with the advance of telecommunication technology nowadays, the benefit of having complete navigational information embedded in each document, which can be visualized and validated, is expected to overcome the moderately high use of communication bandwidth.

4.3 Navigational Validation

In the Retriever, only valid documents can be displayed. Document validation is achieved by analyzing the navigational content of a document. Currently, the validation consists of the *navigational validation* phrase and the *access validation* phrase. In the first phrase, the document is verified whether it is properly placed in valid location in the

Retriever information hierarchy. It is done by recursively analyzing the previous node address and the parent node address in the navigational content of the document whether they also exist in valid location in the hierarchy. In the second phrase, the access flag is interpreted. If the navigational content of the document does not contain the “O” (open) access flag, the flag and the access method to the document are further validated. In case of the “S” flag, the document must be entered from its previous document, and the “R” flagged document must be accessed from its parent document. After the validation, the valid document or an appropriate error message is displayed to users.

5 Future Directions

The Retriever information system is currently in its early stage of development. Presently, the study is primarily conducted on its conceptual architecture and its presentation of information and navigational contents. In the next stage, the communication protocol of the Retriever will be designed, and the client and server applications of the Retriever will be implemented for demonstration. At this time, they are planned to implement as a layer on top of the existed infrastructure of the Web. The eXtensible Markup Language (XML), which is currently the standardized language for information representation and interchange (Bos, 1999), is aimed to employ in the design and development of the Retriever communication protocol and document presentation.

As the design of Retriever emphasizes on ease of navigation in large information space, it will be experimentally tested against the navigation capability of the Web and the Gopher. The study will utilize between-subject design. Subjects will be randomly assigned into three groups. Such groups include Retriever system, Gopher system, and Web system. Subjects in each group will be required to navigate through the system aiming to accomplish the series of tasks. Such tasks are to answer a series of questions in which they require to navigate through the system. The observed outcomes will be measured by three dependent variables. These variables include time to accomplish each task, task accuracy, and numbers of paths used to accomplish each task. All outcomes will be automatically recorded into the log data. This step is expected to conduct when the design and development of the Retriever are done.

6 Reference

- Bos, B. (1999) XML in 10 points, <http://www.w3.org/XML/1999/XML-in-10-points>.
- Karger, B. (2000) gopher:// manifesto, <http://www.scn.org/~bkarger/gopher-manifesto>.
- Lynch, P. J., and Horton, S. (1999) *Web Style Guide*, Yale University Press, New Haven.
- Nielsen, J. (2000) *Designing Web Usability*, New Riders Publishing
- Spool, J. M., Scanlon, T., Schroeder, W., Snyder, C., and DeAngelo, T. (1999) *Web Site Usability: A Designer's Guide*, Morgan Kaufmann Publishers, Inc., San Francisco, CA.
- Whitaker, L. A. (1997) Human Navigation, in C. Forsythe, E. Grose, and J. Ratner (eds.), *Human Factors and Web Development*, Lawrence Erlbaum Associates, Inc., Mahwah, NJ.