STRUCTURE AND MEMORABILITY OF WEB SITES

by

Anna Virginia Marsh

A Thesis submitted in conformity with the requirements for the degree of Master of Applied Science
Graduate Department of Mechanical and Industrial Engineering in the University of Toronto

© Copyright A. V. Marsh 1997
The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author’s permission.

L’auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L’auteur conserve la propriété du droit d’auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.
STRUCTURE AND MEMORABILITY OF WEB SITES

Master of Applied Science 1997
Anna Virginia Marsh
Department of Mechanical and Industrial Engineering
University of Toronto

Abstract

This thesis addresses the design and use of multimedia on the World Wide Web. An experiment was designed and carried out that logged the paths of users on four different Web sites, representing a combination of site sizes and structures. The more hierarchically structured sites received higher numbers of visits and accesses. Maps that subjects created after using Web sites were almost exclusively hierarchical. The number of visits to a node was a better predictor of memorability and judged importance than was the connectivity of the node. The best predictor for node visitation was the level of the node in the hierarchy.

This thesis developed a methodology for studying effects of Web site browsing on recall. The results suggest that people expect hierarchical structure on the Web, and that connectivity based measures for identifying landmarks are inapplicable to the Web. Design guidelines and future research based on this work are also proposed.
Acknowledgments

To David, thank you.

To Charles and Eunice Chaffey, thank you for your support and encouragement.

Thank you Professor Mark Chignell for your help and supervision.

Thank you to this thesis committee, Professor Robert McLean and Professor Charles Meadow.

Thank you to David K. Modjeska for his help in writing and running the computer applications that were so necessary to this thesis.
# Table of Contents

Abstract........................................................................................................................................... i

Acknowledgments................................................................................................................................. ii

List of Tables......................................................................................................................................... v

List of Figures......................................................................................................................................... vi

Introduction........................................................................................................................................ 1
  1.1 Introduction.................................................................................................................................. 1
  1.2 Research Motivation....................................................................................................................... 4
  1.3 Research Questions......................................................................................................................... 5
  1.4 Thesis Overview.............................................................................................................................. 6

Literature Review................................................................................................................................. 7
  2.1 Introduction.................................................................................................................................. 7
  2.2 Information Structure and Navigation............................................................................................. 7
  2.3 Finding Landmarks in Information Structures.................................................................................. 12
  2.4 Tools for the World Wide Web......................................................................................................... 13
  2.5 Studying the World Wide Web.......................................................................................................... 15
  2.6 Hypermedia Structures and Navigational Strategies...................................................................... 17
  2.7 Summary....................................................................................................................................... 18

Analysis of Web Sites.......................................................................................................................... 20

Experiment.......................................................................................................................................... 30
  4.1 Goal.............................................................................................................................................. 30
  4.2 Factors......................................................................................................................................... 31
  4.3 User Tasks..................................................................................................................................... 31
  4.4 Number of Trials............................................................................................................................. 32
  4.5 Subjects......................................................................................................................................... 32
  4.6 Apparatus....................................................................................................................................... 33
  4.7 Questionnaires................................................................................................................................. 33
  4.8 Data Collection................................................................................................................................. 34
  4.9 Data Analysis.................................................................................................................................. 35
  4.10 Results.......................................................................................................................................... 38
    4.10.1 Analysis of Site Effects................................................................................................................ 39


List of Figures

FIGURE 1.1 The relationship between Web, site, and node..................3
FIGURE 3.1 - Diagram illustrating site sizes and structures...............23
FIGURE 3.2 - Site Distribution graph...........................................24
FIGURE 3.3 - Home Page of Travel Montana..................................26
FIGURE 3.4 - Home Page of NyeLabs Online..................................27
FIGURE 3.5 - Home Page of the Web Museum.................................28
FIGURE 3.6 - Home Page of the Hawaii Visitors & Convention Bureau.........................................................29
FIGURE 4.1 - Site Effects on Accesses..........................................39
FIGURE 4.2 - Site Effects on Nodes Visited.................................40
FIGURE 4.3 - Scatter Graph for Hawaii.................................47
FIGURE 4.4 - Scatter Graph for Montana.................................47
FIGURE 4.5 - Scatter Graph for NyeLabs.................................48
FIGURE 4.6 - Scatter Graph for the Web Museum......................48
FIGURE 4.7 - Number of Mistakes........................................56
FIGURE 4.8 - Sample Map..................................................57
List of Tables

TABLE 4.1 - Summary of Data Collected .................................................................38
TABLE 4.2 - Correlation Results ...........................................................................41
TABLE 4.3 - Predictor variables and amount of explained variance by site for the measure judged importance .................................................42
TABLE 4.4 - Correlations of each of the predictor variables with judged importance, across the four sites (rows) .................................................................43
TABLE 4.5 - Nodes Judged Important on Hawaii ..................................................43
TABLE 4.6 - Nodes Judged Important on Montana ...............................................44
TABLE 4.7 - Nodes Judged Important on NyeLabs .............................................44
TABLE 4.8 - Nodes Judged Important on the Web Museum .................................45
TABLE 4.9 - Predictor variables and amount of explained variance by site for the measure number of mentions .........................................................46
TABLE 4.10 - Correlations of each of the predictor variables with number of mentions, across the four sites (rows) .........................................................46
Table 4.11 - Node Effects for the measure judged importance ...........................50
Table 4.12 - Node Effects for the measure number of mentions ........................50
Table 4.13 - Node Effects for the measure number of visits ..............................51
Table 4.14 - Relationship between Node Level and five other measures for the Hawaii Site .....................................................................................................51
Table 4.15 - Relationship between Node and five other measures for the Montana Site ...........................................................................................................52
Table 4.16 - Relationship between Node and five other measures for the Nye Site .............................................................................................................52
Table 4.17 - Relationship between Node and five other measures for the Web Museum Site ............................................................................................53
TABLE 4.18 - Ranked Scores of the Four Sites .......................................................54
Chapter One
Introduction

1.1 Introduction

The World Wide Web (WWW or Web), with its use of hypertext links and information retrieval capabilities, is a very large hypertext; based on the Internet. Its contents are created and maintained by thousands, if not millions, of different people. The Web's contents are stored in geographically diverse locations and are accessed with different Web client programs, which display documents in different ways. Given the lack of an overall architecture or organization, it is not surprising that Web users can become frustrated because of difficulties in navigating successfully around sites or finding documents relevant to their interests.

The recent growth and popularity of the World Wide Web is astonishing. In the space of a few years, the Web has grown from a tool used mostly by research institutions to a popular and increasingly commercial publishing medium. For instance, between June 1993 and January 1996, the percentage of Web sites with a .com suffix grew from 1.5% to 50% (Gray, 1996a). It is
important for the Web to be studied, because of its huge impact on global communications. As it grows and evolves, and as its permanency seems more and more assured, changes should be made on the basis of careful research, rather than solely in accordance with popular thought and fashion.

The World Wide Web is a gigantic mine of information that is organized into sites. Individuals put together information into larger structures called sites. The information is stored on computers that are called servers. Web users then can access the information on the servers with their own computers which are called clients. Generally speaking, a site is contained on a server, although the information may be spread over many servers. A Web site is a collection of (usually similar) information that has been selected and put together by an individual. Thus Web sites are diverse in content and style.

A Web page, or node, is a smaller unit than the web site. A Web page can be thought of a single document. When viewing a Web page with a Web browser one can scroll down the length of the document to the end. One jumps from page to page with use of hypertext links and commands that the Web browser offers. These terms will be fully defined at the end of this section.

This thesis seeks to address certain questions so that we may understand how people use the Web and what features they remember. The thesis is concerned with the Web at a site level; that is, it focuses on the design of Web sites. Further research is needed at node (page) and global (Web) levels, to investigate how people read and navigate at different levels of discrimination. Figure 1.1 illustrates the containment relationships between Web, site, and node.
FIGURE 1.1 The relationship between Web, site, and node

As the World Wide Web is a digital\(^1\) and not a physical space, discussing concepts and aspects of the Web requires a clear vocabulary. Thus, some terms used in this thesis will now be defined.

Node: a node on the Web is commonly referred to as a page. It is typically a small document that may extend over one or more screens of information.

Site: a collection of nodes or pages. Sites are typically maintained by individuals or organizations. The size of a site may vary from just a few pages to many thousands of pages.

Link (or hyperlink): a connection from one node or page to another. Links typically are shown as highlighted (e.g., by underlined blue text). The collection of nodes and links in a site form a network structure,

\(^1\)Digital space is used in this thesis to refer to the concept of space that a user has where none exists (e.g. on the World Wide Web). This term is synonymous, in regards to this thesis, with hyperspace, virtual space or information space.
which is typically hierarchical in form, although with varying degrees of crosslinking between the branches of the hierarchy (tree structure).

This thesis concentrates on Web sites, and on their structures in particular. This thesis seeks to understand what makes some nodes more memorable than others. This thesis also seeks to gather some log data on how users use the Web, as opposed to general anthropological studies. An experiment is described in which subjects navigated through Web sites while their interactions were recorded. The results are analyzed in terms of how the structure of a site affects the perceived importance of the site and the judged importance and memorability of the nodes within the site. Based on these results, design guidelines will be given, as well as recommendations for future work.

1.2 Research Motivation

The main research goal of this thesis is to develop a methodology for assessing the impact of site navigation on recall, and to obtain preliminary results on the effects of site size and structure on site and node recall. This thesis also considers the role of landmarks on a Web site and compares the effectiveness of differing techniques for identifying landmarks within Web sites.

The Pocket Oxford English Dictionary gives one definition of landmark as a "conspicuous object by which one can take one's bearings" (Oxford Press, 1942, p. 438).

Mukherjea and Foley (1995, b) proposed an algorithm which uses the connectivity measures of a node (the number of nodes that are connected to the given node by one and two jumps), and thus arrive at a measure of a landmark. The structural landmarks defined by Mukerjea and Foley are not necessarily related to actual performance. Thus there may be another behavioural landmarks that may be psychological salient or meaningful regardless of their degree of structural connectivity to other nodes.
City planner Kevin Lynch, identified urban behavioural landmarks in his classic book (Lynch, 1960) *The Image of the City*. Lynch identified as landmarks those city elements that people remembered well and used in their navigation through a city. A similar identification could be made a Web site. A combination of high memorability and high usage (during navigation) will be used in this thesis to identify behavioural landmarks on Web sites.

The experiment described later in this thesis will identify characteristics of a node that determine its importance on a Web site. Features of a node that may predict or determine its importance include a variety of measures of connectivity, as well as how far down it is in the structure of that site. The findings obtained in this thesis will be used to provide guidelines for Web site design and further research efforts.

### 1.3 Research Questions

There are few if any strong scientifically supported guidelines for designing the structure of Web sites. In spite of all the recent interest in the Web, including statistical analysis of usage and demographics, there is still a lack of research based design guidelines for Web site construction. Research questions with strong design implications include:

- How can nodes on a site be connected in such a way that important information is emphasized or highlighted?

- What features of a node and the way it is connected to the rest of the site determine how frequently it is visited and remembered?

Landmarks, i.e., nodes that are highly useful in navigation, may have an important role in organizing information within a site, and in improving the navigability and memorability of the site. However, just because a node is a structural landmark (i.e. is pointed to either directly or indirectly by many other nodes) does not imply users will treat it as a landmark, find it memorable, or judge it to be important.
Thus an additional question addressed in this thesis is:

- What is a better predictor of the memorability of a node, its position in the site hierarchy, or the degree to which other nodes point to it (i.e., its back second order connectivity)?

### 1.4 Thesis Overview

The thesis begins with a literature review covering major works on Web research and related issues in Chapter Two. The role of hierarchical structures will be discussed, along with previous work on landmarks and their identification.

Chapter Three is a summary of an analysis of Web sites that was carried out prior to experimentation. Chapter Three shows the method and reasoning behind the site selection for the experiment, introduces the four experimental sites, and compares and contrasts them.

The fourth chapter reports on the experimental study. The design and methodology are discussed, along with such topics as subject selection and user task design. The results are then presented and discussed.

The final chapter summarizes the research findings and suggests some guidelines for Web design. Suggestions for future research are also provided, along with the final conclusions for the thesis.
Chapter Two

Literature Review

2.1 Introduction

This literature review will review work that has taken place to date with the World Wide Web, including tools that have been developed for use on the World Wide Web, and research efforts to understand user demographics and browsing behaviour. Documentation of existing Web structures as well as user navigational strategies will also be examined, to gain an understanding of the effects of these structures on user navigation.

2.2 Information Structure and Navigation

In approaching Web site design, a major goal is to produce sites that are easy to navigate and use. The Web is a new information structure of unprecedented size. However, some insights into how to design for easy navigation on the Web may be gleaned from analysis of the problem of navigation in physical environments.

Kevin Lynch studied complex urban environments. One of his key concepts was legibility: "the ease with which its [an object's] parts can be recognized and can be organized into a coherent pattern." (Lynch, 1960) He goes on to explain,
Just as this printed page, if it is legible, can be visually grasped as a related pattern of recognizable symbols, so a legible city would be one whose districts or landmarks or pathways are easily identifiable and are easily grouped into an over-all pattern. (Lynch, 1960)

He also developed the idea of imageability:

that quality in a physical object which gives it a high probability of evoking a strong image in any given observer. It is that shape color, or arrangement which facilitates the making of vividly identified, powerfully structured, highly useful mental images of the environment. (Lynch, 1960)

Stern (cited in Lynch, 1960, p. 10) used this idea in 1914, applying it to artistic objects and terming it "apparency". In a sense, legibility and imageability are two terms that describe the same attribute; imageability is more specific in applying the concept to specifically to cities. A city with high imageability would be one in which residents and visitors became lost rarely, and where they retained a clear mental picture of the city.

In researching his concept of legibility, Lynch discovered five components termed "city elements". The first is a path, which is a channel "along which the observer customarily, occasionally or potentially moves." Paths in a city are subway lines, bus or street car routes, roads, bridges, anything that people would have occasion to move along by any modes of transportation. The second, an edge, is a "linear element not used or considered as [a] path by the user". Edges often define boundaries; for example in Toronto, the Don Valley or the lake front could be considered to be edges around downtown Toronto. Interestingly, an element can be considered a path at one time for a given mode of transportation (e.g., a highway for a car) but an edge at a different time for a different transportation method (e.g., on foot). In this latter case, the path becomes impassable, and will often become a boundary for the user.

The third element, a district, is an urban area with "two-dimensional extent" that has "some common identifying character". In older cities, different
"Quarters" or the "Old City" act as districts. The fourth element, a node\textsuperscript{2} is an "intensive focus", a "strategic spot in a city into which an observer can enter"; users tend to travel to and from nodes. Examples of a node are an intersection, a subway station and a square. A landmark is the last city element Lynch defined, which is an external reference point. Landmarks are used in navigation by users. In Toronto, for example, a prominent landmark is the CN Tower.

To refine his concepts and to develop the idea of city elements, Lynch used two primary methods: interviews with city dwellers and observations by trained observers. The interviews with city dwellers were lengthy and involved sketching, answering questions such as "where is ___?", describing routes, describing landmarks and their identification, and looking at photographs. Some subjects were asked to give city tours or navigate to relevant features. Some interviews were conducted on the streets, by asking passers-by for directions. Lynch was able to produce his "legible" maps of the cities by finding consensus among his subjects. Lynch also developed a number of "trained observers" familiar with his ideas and concepts. The observers could then go out into a city to produce a map of key elements, based on the training.

Lynch's idea of legibility, as it applied to the city as imageability, and to the artistic world as apparency, should also apply to digital or cyber space. A specific application to document-based hypermedia (e.g. the World Wide Web) would be useful to give users a clear sense of location and site overview. Some of Lynch's city elements may well be applicable to the design of Web sites, to the benefit of those sites.

Other researchers interested in visualization include Kaplan and Moulthrop (1994), who explore the cognitive view of digital space by examining different metaphors for digital space and multiple views of it. Their goal is to find a better foundation for representing information structure than the directed

\textsuperscript{2} Please note that Lynch's use of the term "node" differs from this thesis' use of the term. This thesis uses the term to refer to any Web page, whereas in Lynch's vocabulary only Web pages that act as intensive foci could be designated a node. Thus the thesis uses the term almost anatomically, to refer to the smallest part of a Web site, while Lynch uses the term descriptively, to refer to a set of characteristics.
graph. As digital space is not physical space, they claim that "Familiar metaphors drawn from physics, architecture, and everyday experience have only limited descriptive or explanatory value for this type of space." (Kaplan and Moulthrop, 1994, p. 206). They survey a number of alternative ideas from which they select two, termed architectonic space and semantic space, for examination and thought experiments. They conclude, "Semantic and architectonic spaces cannot be perfectly harmonized. We should aim for systems that harmonize the two as well as possible, but which acknowledge the contingent nature of any such harmony." (Kaplan and Moulthrop, 1994, p. 215).

A number of researchers have developed tools to enhance information visualization. Mukherjea and Foley (1995) designed a tool that allows a user to create visualizations of an information space. They examined the concept of binding, where a binding specifies a relationship "between the information attributes and the visual attributes of the nodes and links". For example, they uniquely colour code nodes that deal with a specific topic. They have also tried to increase the navigability of hypermedia by showing the context of a particular node. They claim that:

One of the major problems with current hypermedia software is being lost in hyperspace. ... the process of jumping from one location to another can easily confuse the user. One of the main reasons for this is that the user does not know the context of the node with respect to the overall information space.  

(Mukherjea and Foley, 1995, b, p. 1)

Other researchers have addressed the problem of providing coherent visualizations of information structure. Kim (1995) has developed a threedimensional visualization tool that allows users to gain a clear idea of the structure of Web information. Kim uses three basic information structures with his visualization tool: the tree (or hierarchy), the network, and the timeline (1995). Kim has developed this tool because he believes:

...conventional Web browsers can provide only low retrieval efficiency, and weak recognition about information structures because they depend on simple browsing methods.  

(Kim, 1995, p.1)
Kim blames poor understanding of information structure on "simple browsing methods". However, Kim provides no data to show that poor understanding of information structure follows from simple browsing methods, nor does he show that three-dimensional browsers will alleviate the problems he uses to motivate his work.

Ingram and Benford (1995) suggested that the problem of "getting lost in hyperspace" could be solved by the use of an urban planning approach. In particular, they promote the concept of legibility, i.e., "the ease with which its [a city's] inhabitants can develop a cognitive map over a period of time and so orient themselves within it and navigate through it". (Ingram and Benford, 1995, p. 210).

Ingram and Benford argue that principles which make navigation easier in physical urban environments can be applied to the digital world of hypermedia. They wish to apply Lynch's idea of legibility in information visualization, and they specifically wish to find an automatic algorithm to assign features to existing designs:

One of our main aims is to accomplish this without requiring the users of the system to perform the placement of the features manually. Essentially the system should, wherever possible, identify and place legibility features using information available from the database and visualization systems alone.

(Ingram and Benford, 1995, p. 210)

Ingram and Benford give criteria for data spaces that are suitable for legibility enhancement. The researchers point out that the development of a cognitive map takes place over time, so that the data space to which legibility enhancement is being applied should be one that is durable, relatively stable and that can be accessed repeatedly by the user. On the basis of this criterion, they feel "visualizations of large document stores, file systems or information retrieval systems such as the World Wide Web to be suited to this approach." (Ingram and Benford, 1995, p. 211).
2.3 Finding Landmarks in Information Structures

If one accepts Lynch's idea of a landmark as potentially useful for navigating through hypermedia (and the World Wide Web), then the next step is to identify these landmarks, and to highlight them and utilize them within Web sites.

Valdez, Chignell and Glenn investigated the identification of landmarks in a hypermedia. Concerns over user navigation and "getting lost in information space" (Valdez et al., 1988) encouraged researchers to find ways of assisting user navigation. Valdez et al. hypothesized that landmarks would act as markers or milestones to help users find their way. Valdez et al. proposed a number of ways to identify landmarks, and they developed an experimental methodology (path inclusion assessment) for assessing how well a node served as a milestone. They found that the best predictor of whether a node served as a milestone was its second order connectivity to other nodes.

The work by Valdez, Chignell and Glenn was extended by Mukhejea and Foley, who identified structural landmarks by using an algorithm which calculates a weighted average of the second-order connectedness (SOC), back second-order connectedness (BSOC), in degree (I) and out degree (O). This algorithm essentially counts the number of other nodes linked to the given node, in both forward and backward directions, one and two jumps away. A weighted average is taken with a cut-off level determined to identify the landmarks:

1. Calculate importance = (I+O)*wt1 + (SOC+BSOC)*wt2
2. If importance is in the top 10% of scores, the given node is a landmark.
   (Note: wt1=0.4 , wt2=0.6)

(Mukherjea and Foley, 1995, b)

Mukherjea and Foley point out, "... a major limitation of our system is that it uses just structural analysis for determining the importance of the nodes." (Mukherjea and Foley, 1995, b, p. 4) To add validity to the algorithm, they suggest using a Web access log analysis tool, such as Pitkow and Bharat's WebViz (Pitkow and Bharat, 1995). This Web access log analysis tool would be
used to look at characteristics indicating behavioural landmarks (frequency of visits, etc.).

This research provides a foundation for using a formula to calculate first and second order connectedness\(^3\), and then using the result to identify landmarks. This research would support the notion that structural landmarks exist, and that they can be identified with mathematical and graph-analytical methods. Mukherjea and Foley encourage further research with the statement, "Finding other contextual methods of determining the importance of a node is an open research issue." (Mukherjea and Foley, 1995, b, p. 4).

The concept of behavioural landmarks (i.e., those nodes that are actually used as landmarks, and whose usage can be determined from behaviour, either in terms of navigation or in terms of recall) is not as clearly documented in the research literature, but it can be gleaned from Lynch's original discussion of landmarks. Lynch defines a landmark as something used for navigation (Lynch, 1960). This claim seems to indicate that the quality of a landmark is its use by people, not its structural features.

Thus the preceding literature discusses both structural and behavioural landmarks. In the following sections research that discusses usage of the Web will be reviewed.

### 2.4 Tools for the World Wide Web

The World Wide Web is the world's largest hypertext document, it is also the most anarchical. Since there is no design document or blueprint for the Web, and since it is dynamic and changing every day, its structure on any given day can only be discovered by using specially designed software agents known as crawlers or knowbots, that traverse the Web and record its structure.

\(^3\)First and second order connectedness in this situation refers to the explicit links between nodes. Some (Meadow, 1996) have proposed another type of connectivity that exists between nodes, as links that are mentally established by the user.
For instance, Fielding (1995) developed a tool to traverse Web documents and build an index of them. This tool is called MOMspider (Multi-Owner Maintenance spider). Fielding terms the Web an infostructure, which Tilton has described as "the layout of information in a manner such that it can be navigated" (Tilton, 1995, p. 1).

Tools are clearly needed to maintain Web sites that contain vast amounts of dynamic information. The problems of Web site maintenance are compounded by the fact that many sites have multiple contributing users who are often operating in different geographic locations. Manual traversal of a Web site is tedious and time consuming, so Fielding (1995) points to automated traversal as a solution for maintenance. The MOMspider could be used for additional applications, though. One such application would be as an "explorer" or "cartographer" spider, by traversing foreign Web sites and developing an index that could be transformed into a map. This function is useful for studying the structure of existing Web sites. I have used MOMspider for this purpose in my research.

Fielding claims that a good infostructure arises from design. He does not comment on the nature of that design, however, except that it use hypertext and be non-linear. His tool, MOMspider helps users maintain a current Web site, but the tool would not help one to design a Web site. The MOMspider was used in this thesis research to develop maps of the Web sites to be studied. Detailed discussions of the use of MOMspider can be found in later sections of this thesis.

Gray (1996b) built the Wanderer, another crawler-type tool. The purpose of the Wanderer is to keep count of the number of sites presently on the Web. Gray has compiled a number of statistics on the growth of the Web, some through the Wanderer. Starting in June 1993, he claims the Web had grown from 130 sites to 100,000 in January of 1996 (Gray, 1996, a). A further paper estimates the number of sites in June 1996 at 230,000 (Gray, 1996, b), and Treese (1996) placed the number of commercial (.com) domains at over 300,000 in the October issue of Byte magazine.
2.5 Studying the World Wide Web

The Web is a new domain for research activity. Much of the early research has involved demographic surveys of Web users. For example, Pitkow and Recker recorded usage and demographics. Their 1994 survey offers a picture of early Web demographics. In January of 1994, Pitkow and Recker posted a survey on the Web for a month. Among the interesting results of the survey, "56% of respondents are between the ages of 21 and 30, 94% are male, and 69% located in North America" (Pitkow and Recker, 1994, p. 246).

Pitkow and Recker also discussed some of the shortcomings of the first survey and provided recommendations for future work. The Graphic Visualization and Usability Center (GVU) at Georgia Tech has started performing these surveys every six months. "Results from the Third WWW User Survey" by Pitkow and Kehoe were published in the Proceedings of the Third International World-Wide Web Conference in 1995, while the results from the fourth survey can be accessed via the WWW. A fifth survey will take place in April 1996. A summary of findings from the fourth survey noted the gender split on the WWW was 70% male and 30% female. The average age surveyed was 32.7 years, while the majority of users are in computer and educational occupations. The most common reasons for users to use the WWW are (in descending order of importance) browsing, entertainment, and work. Identified WWW problems include slow transmission speeds (especially downloading pages), users being unable to find information, users not being able to find a page previously visited, and users not being able to visualize where they have been and where they can go. (Pitkow and Kehoe, 1996).

The results from the fourth WWW survey by Pitkow and Kehoe bear interestingly on this research. Their section headed "What are the main problems with using the Web?" brings up some issues that this research will help to define and potentially solve. Transmission speed is a technological issue, which will be solved with time and increased bandwidth. Major problems in user navigation identified in the survey included:
• not being able to find a page known to exist (34.5%)
• not being able to find a page once visited (25.8%)
• not being able to visualize where one has been and where one can go (23.7%)

Pitkow and Kehoe commented that few users (6.5%) reported as a problem not being able to determine their location. Apparently, users are more concerned with possibilities for navigation as such than with current position. If so, this finding may explain why navigation aids as visualization maps have limited usefulness. Perhaps the answer to these navigation questions lies in enhancing the structure and design of Web pages, so that it becomes inherently obvious where one can go and where one has been.

Catledge and Pitkow (1995) argue that insufficient research is being performed on WWW usage. They tried to document the browsing strategies of users over a three week period. Catledge and Pitkow recorded users' Web movements in a log file. The authors then plotted path length against frequency of travel. They were able to characterize user browsing into three categories. The three categories mirror three broad categories found by Cove and Walsh: serendipitous browsing, general-purpose browsing and searching (Cove and Walsh, 1988). Cove and Walsh describe the three categories as follows: search browsing is a closely directed and structured activity where the product or goal is known; general purpose browsing consults specified sources, because it is highly probable the sources contain the area of interest; and serendipitous browsing is a random, unstructured and undirected activity (Cove and Walsh, 1988).

Catledge and Pitkow also observed that "users tended to operate in one small area within a particular site. This structure resembles a spoke and hub structure due to the frequent use of backtracking.". They also observed that "users rarely traverse more than two layers of the hypertext structure before returning to an entry point." (Catledge and Pitkow, 1995). These observations lead to some design recommendations, including the placement of important information within two or three jumps of a home page, and the use of
indexes to support hub and spoke usage patterns. This suggestion implies that structure influences the usability or legibility of a Web site.

The research to date on the World Wide Web has been concerned with demographics and with browsing behaviour. However, there has been no significant research on the structure and characteristics of Web sites. This thesis will present exploratory research in this area.

2.6 Hypermedia Structures and Navigational Strategies

This section of the literature review will examine different hypermedia structures and users' strategies for navigating through them. The results of this review will be used to justify the choice of Web sites for the experimental reported later in this thesis.

Parunak developed a taxonomy of hypermedia structures. He then compared the strategies and structures, showing which hypermedia topologies support which navigational strategies (Parunak, 1989). After he examined navigation strategies in the physical world, he suggested that such navigation should provide a useful model for the virtual world. Paranuk identified five basic hyperspace structures: the linear structure, the hierarchy, the hypercube/hypertorus, the directed acyclic graph and an arbitrary structure (any structure that does not fit the constraints of the first four). These structures can be used to generate nine hyperspace topologies. Parunak gives a set of constraints or equations that can be used to identify each basic structure.

Horton et al., 1996 propose that three basic structures exist on the Web: a hierarchy, a grid (network) and a sequential pattern. The sequential pattern is basically a linear structure following one major argument with side branches for sub-topics. Horton et al. then use these three major categories to define five "application-specific clusters" (e.g. catalog cluster, newspaper cluster) that users can implement for specific applications. The authors do not provide any method to classify existing Web sites into these categories, but they provide simple sketches of the structures; hence it is assumed that the reader will
recognize the structures when encountered. The focus of Horton et al.’s book is design, so it is assumed that users can construct their sites according to these prescribed structures.

Kim (1995) uses three information structures in his research: a tree, a network and a timeline. This scheme parallels the basic information structures used by Horton, if one considers the timeline to be a special case of the general sequential pattern.

Burt (1983) applied network analysis theory to the field of sociology; there he examines, for example, the role of an actor in a social structure. Some of Burt’s techniques, however, may be useful in determining the degree to which a Web site is a network or hierarchy. For example, using Burt’s idea of random vs. nonrandomness, with other measures such as density of links, may prove to be a robust method for categorizing the structure of a site.

The literature on hypertext and Web design focuses on networks and hierarchies as the two main types of structure. These structures have also been recognized as the basis of knowledge representation in artificial intelligence and cognitive science (e.g., neural networks, Rumelhart and McClelland, 1986, and semantic hierarchies, Anderson and Bower, 1980). Consequently, this thesis will examine the impact of network and hierarchical structure on the use of Web sites.

Poblete discusses a recent review of hierarchies and their use (Poblete 1995). Poblete cited research demonstrating that hierarchies have an important role in human cognition and that they are frequently used in storing and presenting information in computers.

2.7 Summary

This literature review shows that solutions to problems of navigating and visualizing hyperspace have been sought by applying principles from the physical world. The concept of landmarks has been investigated by a number
of researchers, and an algorithm has been developed to identify structural landmarks. Implications of structural landmarks, their effects on user behaviour and the correlation between structural and behavioural landmarks have not been studied. Application of architectural principles has yet to be made to Web site design. Some researchers have studied Web usage and demographics, however, results of these studies have not been tied to structural recommendations for Web site design.

This review also demonstrated that, in spite of the relative youth of the Web, there is considerable interest in developing principles for appropriate structuring of Web sites. With the home page as the access point into each site, there has been a tendency to structure sites either as pure hierarchies, or as modified hierarchies that contain "sideways" links between nodes that occur in different branches of the main hierarchy.

Given the structure of a site, Mukherjea and Foley's equation predicts which nodes will function as landmarks, based on their connectivity to other nodes. This provides a useful starting point for characterization of which nodes are important in a site. There is a clear need to determine how well the connectivity of a node predicts its perceived importance and memorability. Knowing how to design Web sites so as to highlight the most important information would be of great use to Web designers.

This thesis will describe an experiment that was run to provide some preliminary data on the relationship between site size and structure, and how that site is browsed and remembered. In addition, the specific question of how the connectivity of nodes affects their judged importance and memorability will be addressed. Prior to the description of the experiment, Chapter 3 will discuss how the Web sites used in the experiment were selected.
Chapter Three
Analysis of Web Sites

An analysis of available Web sites was carried out to identify suitable sites for the experimental study reported in Chapter four of this thesis. Between fifty and sixty Web sites were examined in this analysis. Candidate sites needed to be interesting, well designed, and critically acclaimed (by external sources). The sites required content material that was easily understood by the experimental subjects (thus ruling out sites with technical content). The sites also needed to be easily accessible, both to human subjects and to the Web crawling programme (MOMspider) that was used to extract site structure. These requirements ruled out most sites outside of North America, sites with substantial graphical content and those with poor technical implementation. The four selected sites also needed to differ in terms of size and structure.

The four sites that were eventually selected have received high scores in the Pointcom Web Reviews. Point is an online review service that claims to list the top 5% of available Web Sites. Pointcom is based in New York and is owned by Lycos, Inc., a company that has developed a popular Web search engine. Pointcom awards scores in three different categories. Pointcom describe the categories as follows:

**Content:** Just how broad, deep, and amazingly thorough is the information? Are there good links? Good clips? Is it accurate? Complete? Up-to-date?
**Presentation**: Is the page beautiful? Colorful? Easy to use? Does it lead visitors through the information nicely? Does it use video, audio, and original graphics? Does it break new ground?

**Experience**: This is the key rating. Is this fun? Is it worth the time? Will we recommend it to friends? All things considered, does this site deliver the goods?

(http://point.lycos.com/faq/, 1996)

The first of the four sites used in the experiment was Travel Montana. Travel Montana is designed to help potential tourists find information needed to plan a vacation. Information on accommodations, activities, restaurants and car rental agencies is given as well as useful information on speed limit laws, whether the water in Montana is potable, and whom to contact with various other questions. A unifying graphical theme and interesting pictures made this site a good choice. The site is the largest of the chosen four, with 1277 leaf nodes. It has a hierarchical structure, with no option to jump between branches. The site received a high score, no longer available, from Pointcom, with marks in each category in the 40's (out of 50 available points). The home page of Travel Montana is shown in Figure 3.3.

The second site used was NyeLabs Online, a site supporting the television show "Bill Nye the Science Guy". This is a North American show for school children that focuses on teaching scientific ideas. Here, fans of Bill Nye can find out when the show is broadcast in their area, learn more about the people involved in the show or explore the show's scientific ideas. The site is bright and colourful, and it is designed to appeal to school children (though many university students also find it appealing). The site has 254 leaf nodes and like Travel Montana follows a hierarchical structure. Pointcom's score for NyeLabs was:

- 44/50 Content
- 40/50 Presentation
- 42/50 Experience

The home page for NyeLabs is shown in Figure 3.4.
The Web Museum was the third site chosen for this experiment. The site is an attractive online collection of art, and it also gives information about Paris. Information on painter's lives and times are linked with their art. The site is closely interconnected as various themes are used to link pictures across time and place. Visitors to the Web Museum may also take a graphical tour of Paris and learn about the history of this city. This site has 353 leaf nodes and, though it is laid out as a hierarchy, the interconnections between branches change this site to a network. Web Museum's Pointcom score was:

39/50 Content
41/50 Presentation
42/50 Experience

The home page of the Web Museum is shown in Figure 3.5.

The final site for the experiment was the Hawaii Visitors & Convention Bureau. This site resembles Travel Montana in that both sites are designed to help potential tourists to plan vacations. The Hawaii Visitors & Convention Bureau focuses on accommodations and activities, and it also gives detailed information on each of the state's islands. It is the smallest site with only 109 leaf nodes. Like the Web Museum, Hawaii is laid out as a hierarchy, but it has a high degree of interconnectedness to allow users to navigate it as a network. Hawaii received the following score from Pointcom:

40/50 Content
40/50 Presentation
35/50 Experience

The home page of the Hawaii Visitors & Convention Bureau is shown in Figure 3.6.

There are many ways to categorize Web sites, such as size, purpose (e.g. commercial vs. academic), geographic location, and technical quality. Another way to categorize Web sites is by structure. While measures exist to place networks into structural categories, none of these measures have been applied to Web sites. In contrast to other networks, a Web site has special
considerations as the definition of boundaries, treatment of non-node links (e.g. to sound and picture files) and the ability of the user to navigate through non-link means (e.g. a Web browser's Back button or history list). In this thesis, I will concentrate on two basic structures that are found on the Web - a network and a hierarchy.

The four chosen sites represent a range of sizes and structures. They can be summarized in the following diagram:

![Diagram]

FIGURE 3.1 - Diagram illustrating site sizes and structures

The sites can be divided into two categories based on size. The size of each site was based on its number of nodes. The actual node size of the sites was as follows:

<table>
<thead>
<tr>
<th>Site</th>
<th>Nodes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hawaii</td>
<td>109</td>
</tr>
<tr>
<td>NyeLabs</td>
<td>254</td>
</tr>
<tr>
<td>Web Museum</td>
<td>353</td>
</tr>
<tr>
<td>Travel Montana</td>
<td>1277</td>
</tr>
</tbody>
</table>
The sites were also divided into two structural categories, "high" hierarchical and "low" hierarchical. To measure a site's degree of hierarchical structuring, it was decided to use the percentage of structural landmarks (as calculated by the Mukherjea and Foley algorithm). This measure was effective because a structural landmark is defined by four connectivity measures (in1, out1, in2 and out2) and represents the number of links in a site. The greater the number of links (normalized over the site), the more a site resembles a network. The percentage of structural landmarks by site was as follows:

Hawaii 79.81%
Web Museum 78.47%
NyeLabs 8.26%
Travel Montana 1.64%

The following graph shows the distribution of the four sites with respect to size and landmark percentage:

![Site Distribution Graph]

FIGURE 3.2 - Site Distribution graph
In exploring the Web and looking for suitable sites for this experiment, I discovered that the majority of sites examined (50-60) followed a hierarchical structure, with very little interconnection. I found no examples of a true network. However, sites that had a high percentage of landmarks (over 75%) were considered to be network-like. These networked sites tended to be quite small in size. The Web Museum was the largest networked site found to be suitable for use in the experiment.
Disclaimer: No endorsement is intended or made of any hypertext link, product, service, or information either by its inclusion or exclusion from this page or site. While all attempts are made to insure the correctness and suitability of information under our control and to correct any errors brought to our attention, no representation or guarantee can be made as to the correctness or suitability of that information or any other linked information presented, referenced, or implied. All critical information should be independently verified. Any questions should be directed to the administrator/s of this or any other specific sites.

This site is a service of the Montana Superhost Program
Please send comments and suggestions to: webmaster@travel.mt.gov

All contents copyright © 1996, Travel Montana. All rights reserved
Revised Friday, August 9, 1996
Please visit us again at: http://travel.mt.gov/

FIGURE 3.3 - Home Page of Travel Montana
(Travel Montana, 1996)
FIGURE 3.4 - Home Page of NyeLabs Online
(KCTS Television, 1996)
26 May 1996: Joining the WebMuseum network:
Friedrich-Alexander Universität of Erlangen-Nürnberg (Germany)

Welcome to our guests from all over the world! The ever-expanding WebMuseum network is now welcoming 200,000 visitors every week, delivering over 10 million documents! (PS: remember to click on the inlined thumbnail images to enlarge them.)

I wish you the most pleasant visit!

If this is one of your first trips on the Web, you may want to start exploring a tiny subset of the collections first: try out the medieval art exhibit. However, most artworks are exhibited in the Famous Paintings section. One final word: if you ever get lost, feel free to experiment the endpage navigation panel. Enjoy!

Special Exhibitions

Paul Cézanne

a unique set of over 100 artworks, exclusively on the WebMuseum!

Medieval art:
Les très riches heures du Duc de Berry

General Exhibitions

Explore the WebMuseum unique Famous Paintings collections

Other Resources

Visit Paris on a small tour

The Auditorium

All you ever wanted to know about the WebMuseum
(includes What's New in the WebMuseum)

Don't forget to visit our hosting sponsor home page!

© 22 Jun 1996, Nicolas Pioch - Top - Up - Info

Thanks to the BMW Foundation, the WebMuseum mirrors, partners and contributors for their support.
The Hawaii Visitors & Convention Bureau welcomes you to our Hokeo Hawaii Web Server!

Where you can enter to win your own customized Hawaii Vacation!
An experimental study was designed to address the research questions. The purpose of the experiment was to investigate how subjects used Web sites of different sizes and structures, and how the subjects perceived the Web sites. The purpose of this study was not to provide definitive answers, but rather to demonstrate the use of the methodologies developed and to develop initial findings that would form a basis for future research.

4.1 Goal

The goal of this experiment was to address the three research questions posed in the first chapter of this thesis:

- How does node position within site structure relate to corresponding node memorability?

- What features of a node and the way it is connected to the rest of the site determine how frequently it is visited and remembered?
• What is a better predictor of the memorability of a node, its position in the site hierarchy, or the degree to which other nodes point to it (i.e., its back second order connectivity)?

4.2 Factors

Two different site structures were used - strongly hierarchical and less strongly hierarchical. Two different structures were used to discover whether site structure affects the user's navigational paths. The sites were also divided into two groups by size - large and small.

4.3 User Tasks

The experimental method was designed to give the user a good exposure to the site, in both breadth and depth. After some pilot tests, it was found that five minutes of browsing, combined with ten search tasks (in the form of questions), took users about 20 to 25 minutes to complete. The ten questions were designed to move users throughout the Web site and not to concentrate on any particular area. The prior five minutes of browsing allowed a user to explore the Web site at their own pace, and to stop to examine any interesting items.

The question design had as a primary objective the subject's exposure to the site. The questions did not focus on the subject's search efficiency. The subjects were given a questionnaire after using each site, and they were also given a summary questionnaire after completing the tasks on all four sites. The questions were presented in random order (the experimenter shuffled

---

4As noted in Chapter 2, all the sites used in this study were hierarchical to some degree. The "hierarchical" sites were strict hierarchies, whereas the other sites were "tangled hierarchies", with links occurring between nodes located in different branches of the hierarchy.
the questions before each trial) and the order in which the subjects used each site was balanced.

The questions used in this experiment, the instructions for the subjects and the subject consent form can be found in Appendix A.

4.4 Number of Trials

Each subject completed four trials, one trial per site. Twelve subjects participated, for a total of 48 trials to collect data. Twenty-four possible sequences for site presentation order were possible, of which twelve were used (one per subject). The presentation order was completely balanced so that any learning effects or other subject order influences would be counter-balanced.

4.5 Subjects

Subjects selected were required to speak English well enough to read and understand the questions and to find the answers, and subjects needed to have previous experience using Netscape. Thus no subject training was necessary. Subjects were found by word of mouth, and were paid $20 for the two hours of experiment time. Twelve subjects participated, all of whom were university students or recent graduates. Four of the subjects were female, and the subjects' backgrounds were evenly split between technical (engineering, computer science) and non-technical (arts, psychology, etc.) disciplines.
4.6 Apparatus

The experiment was performed on a Macintosh PowerPC running Netscape 2.0 with a high-speed connection to the Internet backbone of the University of Toronto. The same computer terminal and network connection was used for all twelve subjects. Although the speed of the network connection varied with the time of day, using the same computer terminal minimized this disruption; as no time-specific data were being collected, the slight variations in network speed did not affect experimental results. Use of the same terminal also minimized variance in the data due to environmental influences (lighting, noise, etc.).

4.7 Questionnaires

At the end of each trial (for each site), subjects were administered a questionnaire by the experimenter. This questionnaire asked the subjects to describe the site; to name the most distinctive pages; to recall nodes; to draw a map of the site, and to indicate which nodes were most important. The questions about the distinctive pages and site description were to warm up the subjects and to start them thinking about the site and the visited nodes. Each subject answered this questionnaire four times (once per trial) and gave answers for each site. At the end of the experiment a second questionnaire was administered. In this questionnaire, subjects were asked to comment on their favorite site (of those visited) and to give some guidelines about site design on the basis of their experiences. The subjects were also asked to rank the sites subjectively by size. For copies of the actual questionnaires, see Appendix B.
4.8 Data Collection

Two types of data were collected. First, the subjects' movements around the Web site were logged electronically. The logging programme captured the AppleEvents sent by Netscape with each selection of a URL, and it listed the URLs that the subjects visited. The second type of data collected was subjective answers in the questionnaires.

To gather data about the structure of the four Web sites, I employed the MOMspider Web crawler. Information about the development of MOMspider can be found in chapter 2 of this thesis.

The Web crawler ran as a background process over a period of several weeks. It generated a set of four files in HTML format. Among other information, each file included a list of site pages with name, URL (Uniform Resource Locator), and child nodes (outgoing links). This information was sufficient to construct a reasonably complete graph of each research site.

To record the Web browsing of the subjects, a simple AppleScript utility was developed. This utility took advantage of Netscape's client API (Application Programmer Interface) on the Macintosh platform. This API enables run-time communication in both directions between Netscape and client programs. For these purposes, it was sufficient to log the EchoURL events sent by Netscape. The small AppleScript utility was developed using a standard development environment with public-domain XCMD (Hypercard external command) extensions.

The results of the experiment were collected and analyzed. Each URL was assigned a unique identifying number. Then the subjective data was summarized, so that the most frequently recalled pages, most frequently drawn pages and the pages judged important could be analyzed with respect to the differences between sites. From the subjects' logs, the most frequently visited nodes were collected. From the earlier crawls by MOMspider, it was possible to determine the structural landmarks and the level of each node in
the site graph hierarchy. All of this information was considered during data analysis.

4.9 Data Analysis

Data collected in the experiment covered two different levels, site and node, and three general categories: structural, log, and subjective responses. (A summary table showing data collected at each of these levels and categories can be found at the end of the section.) The first data category was structural. This data describes the structure of a site's nodes. At the node level, the data took the form of connectivity measures, identification of landmarks (according to Mukherjea and Foley's algorithm) and the node's hierarchy level.

Four connectivity measures were used; out degree, in degree, second order connectedness and back second order connectedness. Out degree describes the number of other nodes that can be reached directly from the current node; in degree describes the number of nodes that link directly to the current node. For the purposes of brevity, first and second order forward connectivity will be referred to as Out1 and Out2, and first and second order backward connectivity will be referred to as In1 and In2, respectively, in the remainder of this thesis. Second order (forward) connectedness is the number of nodes that can be reached in two steps from the current node, and back second order connectedness describes the number of nodes that point to the current node within two steps.

To determine a node's level in the structure, a breadth-first traversal was used. The top node (home page) was designated level 0. Node that were directly linked to this level were designated level 1. Nodes that linked directly to level 1 were designated level 2, etc. This technique worked well with the two hierarchical sites. For the networked sites, I assumed the structure to be a

---

5 The terminology used here to describe these measures follows that used by Mukherjea and Foley.
6 i.e., the number of nodes that point to nodes that in turn point to the current node.
hierarchy with many interconnections between branches. Ignoring the interconnections revealed a hierarchical structure, and thus a method of measuring a node's level.

The structural data at the site level mirrors somewhat the data at the node level. Data in this category include the average connectivity measures (in1, out1, in2, out2) for the whole site, the depth of the site, its number of levels, the number of nodes that appear at each level, the total number of nodes, and the number of structural landmarks.

To process the data from the structural traversals and Netscape logging, a series of UNIX awk, grep, sed, and C++ utilities were developed. These utilities fall into three general categories: data refinement, usage pattern analysis, and landmark analysis. Input files for these utilities fell into three categories: structural, behavioural, and aggregate behavioural.

The research reported in this thesis used a number of software utilities developed by David Modjeska at the University of Toronto. One of these utilities carried out a landmark analysis by assembling connectivity graphs, which were used to generate output files listing and analyzing graph node levels and landmarks. Landmarks were identified according to the formula developed by Mukherjea and Foley. Node levels were also assigned, based on a breadth-first traversal from the root node of each site.

Two types of data were collected in the experiment: a set of subjective data based on subject's responses to questions and tasks, and a set of log data from the logs the subjects generated by moving through the Web site.

After using each site, subjects were asked to name all the nodes they could remember from the site. Pooled across subjects, this data provides a count of subjects' node recall. The drawing task (where subjects were asked to draw a map of a site) produced similar data. The two measures (nodes recalled and drawn) were combined in a third measure, i.e., how often a node was mentioned. The fourth measure (judged important) was based on which nodes subjects felt were most important.
Within each site, subject responses were pooled, yielding the number of nodes recalled, drawn, mentioned and judged important for that site. In addition, the subjects were asked to rank each site according to their perception of its size, producing another observational measure (subjective size).

A number of measures were derived from the log data including the number of visits per node. The subject's path was examined, and the nodes visited were collected into a "virtual site". Each node the subject had visited became a node in the virtual site, while the nodes the subject did not visit were left out of the virtual site. The nodes on the virtual site were connected as they were on the original site. Thus, each subject generated four different virtual sites (one per site) and a total of 48 virtual sites were generated in the experiment. Connectivity measures from the virtual site were produced (in1, out1, in2 and out2) and landmarks were calculated for the virtual site (referred to as virtual landmarks).

At a site level, relevant measures included the number of accesses per site (i.e., the number of times any node on the site was accessed), the total number of nodes visited on the site, and the percent of the site that was visited. The number of virtual landmarks per site, and the percent of nodes on the site that were virtual landmarks, were also computed.

The collection of this data allowed two levels of analysis to be conducted, at a site level and at a node level. Table 4.1 summarizes the measures used in this study.
### TABLE 4.1 - Summary of Data Collected

<table>
<thead>
<tr>
<th>Structural Measures</th>
<th>Log Measures</th>
<th>Subject Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Node Identification of Landmarks</td>
<td>Number of Visits</td>
<td>Node Recalled</td>
</tr>
<tr>
<td>Node Level</td>
<td></td>
<td>Nodes Drawn</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Nodes Mentioned</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Nodes Judged Important</td>
</tr>
<tr>
<td>Site Average Connectivity Measures</td>
<td>Number of Accesses</td>
<td>Number of Nodes Recalled</td>
</tr>
<tr>
<td>Depth of Site</td>
<td>Number of Nodes Visited</td>
<td>Number of Nodes Drawn</td>
</tr>
<tr>
<td>Number of Nodes at Each Level</td>
<td></td>
<td>Number of Nodes Mentioned</td>
</tr>
<tr>
<td>Total Number of Nodes</td>
<td></td>
<td>Number of Nodes Judged Important</td>
</tr>
<tr>
<td>Number of Structural Landmarks</td>
<td></td>
<td>Subjective Size Ranking</td>
</tr>
</tbody>
</table>

### 4.10 Results

Two main analyses were carried out on the experimental data. The first analysis was carried out at a site level; the goal here was to identify and to explain differences between sites. The second analysis (at a node level) assessed which characteristics tended to be associated with those nodes that served as behavioural landmarks. A regression analysis was carried out to identify the best predictors of which nodes would be mentioned or judged important. Analysis of variance was then performed to determine what characteristics of a node predicted how frequently it would be visited.
The statistical analyses for this thesis were carried out using SPSS for the Macintosh (Version 6.1), an alpha level of 0.05 was used in all tests of significance reported in this chapter.

4.10.1 Analysis of Site Effects

The sites were grouped by their size, structure (network vs. hierarchy) and the order in which the subjects had seen them. This division of sites by order was to detect any learning effects among subjects (e.g., that the subjects tended to mention more nodes in the last-seen site).

A repeated measures ANOVA was carried out with size and structure as factors and number of accesses as the dependent variable. Figure 4.1 shows that the total number of accesses (number of times any page was requested by the subjects) is higher for the strongly hierarchical sites ($F[1,23]=127.76$, $p<.001$). There was no significant interaction between size and structure ($F<1$), nor was there any main effect of size ($F<1$) with respect to number of accesses.

![Accesses Graph](image)

FIGURE 4.1 - Site Effects on Accesses

---

7 with the exception of the Friedman two-way ANOVA which was calculated by hand (with a calculator)
A repeated measures ANOVA was then carried out with number of nodes visited as the dependent variable. As with number of accesses, there was no significant interaction between size and structure (F<1), and there was no significant size effect (F<1). However, as shown in Figure 4.2, there was a significant effect of structure (F[1,23]=133.53, p<.001), with a greater number of different nodes being visited on the hierarchical sites.

The number of nodes mentioned (the inclusive union of recall and drawing) showed no significant interaction (F[1,23]=2.14, p>.1), and no significant effects from size (F<1) or structure (F=1).

The interaction for number of nodes recalled was not significant given the statistical power available in this study (F[1,23]=3.99, .05<p<.10) and there were no significant effects from size (F<1) or structure (F[1,23]=2.04, p>.10).

The nodes drawn measure showed no significant interaction (F<1) or significant effects from size (F<1) and structure (F<1).

For the number of nodes judged important, there were also no significant effects.
4.10.2 Learning Effect

To test for a learning effect, the sites were divided into four groups by order of subject viewing (first, second, etc.). A repeated measures ANOVA was then used to determine whether there was any significant learning effect. The variables tested in this ANOVA were dependent measures collected from the log data (number of accesses, subject’s path’s connectivity measures, number of nodes the subject visited and the number of virtual landmarks created by the subject’s path) and from the subject’s responses (number of nodes the subject recalled, drew, mentioned or judged important). There was no significant effect of order in which subjects saw the sites and thus no evidence of any learning effects across sites.

4.10.3 Analysis of Node Effects I: Predictors of Memorability

To begin the analysis at a node level, regression analysis was used to see which variables would predict well the subjects’ responses. Correlation analysis was used to assess interdependencies between variables. It was found that the recall, drawn and mentioned data were highly correlated. Table 4.2 shows that the correlation between mention and drawn, and mention and recall were both greater than .96 across all four sites. The correlation between mention and judge were much lower (ranging between .53 and .68).

<table>
<thead>
<tr>
<th></th>
<th>Hawaii</th>
<th>Montana</th>
<th>NyeLabs</th>
<th>Web Museum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drawn</td>
<td>0.96</td>
<td>0.98</td>
<td>0.98</td>
<td>0.96</td>
</tr>
<tr>
<td>Judge</td>
<td>0.61</td>
<td>0.68</td>
<td>0.53</td>
<td>0.68</td>
</tr>
<tr>
<td>Recall</td>
<td>0.97</td>
<td>0.99</td>
<td>0.98</td>
<td>0.99</td>
</tr>
</tbody>
</table>

TABLE 4.2 - Correlation Results
Two forward stepwise regression analyses was carried out using the SPSS statistical package (SPSS, 1996). The purpose of the first analysis was to determine which variables best predicted the nodes that were judged important. The second analysis developed a predictive equation for the memorability of nodes.

The results of the analysis are shown in the following tables (4.3 and 4.9). For each of the two regression analyses performed, the predictor variables are shown by site. The number beneath the predictor variables indicates the beta value (slope) assigned to that variable. Where more than one variable is used to predict, the order in which the variables entered the equation is shown in parentheses beside the beta value. Negative beta values are shown by parentheses. The final column shows the amount of variance that is explained by using these predictor variables.

<table>
<thead>
<tr>
<th>Judged Importance</th>
<th>Totvis</th>
<th>In1</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>HAWAII</td>
<td>0.043</td>
<td></td>
<td>73.6%</td>
</tr>
<tr>
<td>MONTANA</td>
<td>0.022</td>
<td></td>
<td>77.8%</td>
</tr>
<tr>
<td>NYELABS</td>
<td>0.031</td>
<td></td>
<td>64.7%</td>
</tr>
<tr>
<td>WEBMUSEUM</td>
<td>0.034(1)</td>
<td>0.030(2)</td>
<td>70.0%</td>
</tr>
</tbody>
</table>

TABLE 4.3 - Predictor variables and amount of explained variance by site for the measure judged importance

The total visits measure (totvis) is by far the most powerful predictor of judged importance. It was the first variable used (and therefore the best predictor) for the stepwise regression analyses for all four sites. The other

---

8 Stepwise regression was used to avoid problems of multicollinearity when a large number of correlated predictors are used in multiple regression. Forward stepwise regression has generally been regarded as the most economical and statistically defensible method for choosing the right combination of predictor variables in multiple regression (e.g., Draper and Smith, 1966, p. 172). For the data obtained in this thesis, similar results were obtained using both a forwards and backwards selection procedure; consequently, only the results for the forwards stepwise regression are reported here.

9 For this analysis, memorability was represented by how often the node was mentioned. As shown in Table XX, this variable was found to be closely related to other measures of memorability such as recall and the number of times the node was included in the sketch (drawing) of the site.
predictor variables included in the analysis were the structural connectivity measures for each node (in1, out1, in2 and out2), and node level and slmark (whether the node was a structural landmark, as predicted by the Mukherjea and Foley equation).

Examining the data in another way, it is clear that the total visits measure correlates the best with judged importance across the four sites, as is shown in Table 4.4.

<table>
<thead>
<tr>
<th></th>
<th>IN1</th>
<th>IN2</th>
<th>OUT1</th>
<th>OUT2</th>
<th>LEVEL</th>
<th>SLMARK</th>
<th>TOTVIS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hawaii</td>
<td>0.290</td>
<td>0.250</td>
<td>0.502</td>
<td>0.444</td>
<td>-0.430</td>
<td>0.099</td>
<td>0.858</td>
</tr>
<tr>
<td>Montana</td>
<td>0.832</td>
<td>0.694</td>
<td>0.056</td>
<td>0.017</td>
<td>-0.481</td>
<td>0.505</td>
<td>0.882</td>
</tr>
<tr>
<td>NyeLabs</td>
<td>0.055</td>
<td>0.027</td>
<td>0.576</td>
<td>0.433</td>
<td>-0.372</td>
<td>0.311</td>
<td>0.804</td>
</tr>
<tr>
<td>Web Museum</td>
<td>0.502</td>
<td>0.471</td>
<td>0.431</td>
<td>0.428</td>
<td>-0.434</td>
<td>0.199</td>
<td>0.733</td>
</tr>
</tbody>
</table>

TABLE 4.4 - Correlations of each of the predictor variables with judged importance, across the four sites (rows).

Tables 4.5 through 4.8 show the nodes that were judged important and the number of visits by each site\(^\text{10}\).

<table>
<thead>
<tr>
<th>HAWAII</th>
<th>Page Name</th>
<th>Judged Important</th>
<th>Number of Visits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learn About the Islands</td>
<td>8</td>
<td>90</td>
<td></td>
</tr>
<tr>
<td>Activities</td>
<td>5</td>
<td>76</td>
<td></td>
</tr>
<tr>
<td>Accommodations</td>
<td>3</td>
<td>40</td>
<td></td>
</tr>
<tr>
<td>Homepage in Japanese</td>
<td>1</td>
<td>4</td>
<td></td>
</tr>
</tbody>
</table>

TABLE 4.5 - Nodes Judged Important on Hawaii

Table 4.5 shows that the Hawaii site had only four nodes that were judged to be important. It can be seen that there is a strong relationship between the number of visits to a node and the frequency with which it was judged to be important.

\(^\text{10}\) The data point representing the homepage has been removed from each of the tables. This is because the number of visits was high for each homepage (200-300) and the trend is more clearly shown with this data point omitted from the tables.
Table 4.6 shows the corresponding relationship for the Montana site. Here it is also evident that the more frequently a node is visited the more frequently it tends to be judged important.

<table>
<thead>
<tr>
<th>MONTANA</th>
<th>Judged Important</th>
<th>Number of Visits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Page Name</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Camping and Lodging</td>
<td>5</td>
<td>99</td>
</tr>
<tr>
<td>Events and Activities</td>
<td>4</td>
<td>88</td>
</tr>
<tr>
<td>Recreation and Adventure</td>
<td>3</td>
<td>83</td>
</tr>
<tr>
<td>Travelers Services</td>
<td>3</td>
<td>77</td>
</tr>
<tr>
<td>State Parks</td>
<td>1</td>
<td>53</td>
</tr>
<tr>
<td>What's New</td>
<td>1</td>
<td>30</td>
</tr>
<tr>
<td>Hotels and Hot springs</td>
<td>1</td>
<td>27</td>
</tr>
<tr>
<td>Welcome from the Governor</td>
<td>1</td>
<td>17</td>
</tr>
<tr>
<td>Private Campgrounds</td>
<td>1</td>
<td>9</td>
</tr>
<tr>
<td>Glacier National Park</td>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td>Yellowstone National Park</td>
<td>1</td>
<td>6</td>
</tr>
</tbody>
</table>

**TABLE 4.6 - Nodes Judged Important on Montana**

The relationship between number of visits and frequency of importance judgments also holds for the NyeLabs site (Table 4.7), although with some anomalies (for instance the Map of the NyeLabs is judged important most frequently and yet is visited slightly less frequently than two other nodes).

<table>
<thead>
<tr>
<th>NYELABS</th>
<th>Judged Important</th>
<th>Number of Visits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Page Name</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Map of NyeLabs</td>
<td>7</td>
<td>63</td>
</tr>
<tr>
<td>Information</td>
<td>3</td>
<td>70</td>
</tr>
<tr>
<td>U-Nye-Verse</td>
<td>2</td>
<td>72</td>
</tr>
<tr>
<td>Goodies</td>
<td>2</td>
<td>50</td>
</tr>
<tr>
<td>TV Listings</td>
<td>2</td>
<td>17</td>
</tr>
<tr>
<td>Nye York Times</td>
<td>1</td>
<td>28</td>
</tr>
</tbody>
</table>

**TABLE 4.7 - Nodes Judged Important on NyeLabs**
Although there are some anomalies in the Web Museum results as well (Table 4.8) the overall relationship between number of visits and frequency of judged importance also holds.

The results are not as clear for mentions (memorability). In this case, there are a number of variables that correlate with mention. Since these variables also correlate among themselves it is difficult to develop a definitive regression model for the relationship. Furthermore, the pattern of relationships between the mention variable and the predictors appears to change across the four sites studied, as shown in Table 4.911. This is in contrast to the more stable pattern of predictor-criterion correlations for judged importance (at least for total number of visits) shown earlier in Table 4.4.

---

11 This instability of regression equations across sites is most likely due to the high degree of multicollinearity amongst the predictor variables.
The correlations are negative for the correlation variable because number of mentions tend to increase at higher levels of the site (which were coded with lower numbers, 0, 1, etc.).

For each of the graphs, the point representing the homepage has been removed. This was because the large number of visits to the homepage (in the 200-300 range) skewed the graphs. The removal of the homepage allows a clearer picture of the remaining data points.

<table>
<thead>
<tr>
<th>Mention</th>
<th>Totvis</th>
<th>In1</th>
<th>In2</th>
<th>Out1/2</th>
<th>Level</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hawaii</td>
<td>.065(2)</td>
<td></td>
<td></td>
<td></td>
<td>(1.44)(1)</td>
<td>63.4%</td>
</tr>
<tr>
<td>Montana</td>
<td>.056(3)</td>
<td>(.001)(4)</td>
<td>.005(1)</td>
<td></td>
<td>(.431)(2)</td>
<td>82.2%</td>
</tr>
<tr>
<td>NYELABS</td>
<td>.160(1)</td>
<td></td>
<td>(1.11)(2)</td>
<td>(.665)(3)</td>
<td>69.5%</td>
<td></td>
</tr>
<tr>
<td>WEBMUSEUM</td>
<td>.049(1)</td>
<td></td>
<td>.010(2)</td>
<td></td>
<td></td>
<td>52.4%</td>
</tr>
</tbody>
</table>

TABLE 4.9 - Predictor variables and amount of explained variance by site for the measure number of mentions

Table 4.10 shows that while most of the correlation of the measures with mention vary considerably across the sites, the corresponding correlations for level and total visits are more consistent between sites (ranging between .6 and .712).

<table>
<thead>
<tr>
<th></th>
<th>IN1</th>
<th>IN2</th>
<th>OUT1</th>
<th>OUT2</th>
<th>LEVEL</th>
<th>SLMARK</th>
<th>TOTVIS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hawaii</td>
<td>0.426</td>
<td>0.383</td>
<td>0.499</td>
<td>0.627</td>
<td>-0.6968</td>
<td>0.237</td>
<td>0.685</td>
</tr>
<tr>
<td>Montana</td>
<td>0.525</td>
<td>0.865</td>
<td>0.04</td>
<td>0.083</td>
<td>-0.68</td>
<td>0.785</td>
<td>0.695</td>
</tr>
<tr>
<td>NYELABS</td>
<td>0.189</td>
<td>0.16</td>
<td>0.306</td>
<td>0.663</td>
<td>-0.601</td>
<td>0.561</td>
<td>0.673</td>
</tr>
<tr>
<td>Web Museum</td>
<td>0.304</td>
<td>0.334</td>
<td>0.323</td>
<td>0.602</td>
<td>-0.607</td>
<td>0.314</td>
<td>0.651</td>
</tr>
</tbody>
</table>

TABLE 4.10 - Correlations of each of the predictor variables with number of mentions, across the four sites (rows).

Scatter plots were constructed for the four sites, showing the relationship between the number of mentions and the total visits per node13. While total visits may not correlate as well with mentions as with judged important, these graphs show that there is a distinct relationship between the two. The scatter plots are shown in the following Figures 4.10, 4.11, 4.12 and 4.13.
HAWAII

Total Number of Visits

Number of Mentions vs. Total Number of Visits

MONTANA

Total Number of Visits

Number of Mentions vs. Total Number of Visits
Overall, the best predictor of nodes remembered and judged important is how many times a user has seen them. This result makes sense intuitively since repeated visits should increase the chance of recall. The question then becomes: what makes a person visit some nodes more often than others? The next section will describe an analysis intended to help answer this question.

4.10.4 Analysis of Node Effects II: Identification of Significant Discriminators

The preceding analyses compared the aggregate behaviour across different individuals. For instance, total accesses were found to be greater when a person worked on a strongly hierarchical site than when on a weakly hierarchical site. In this section I will examine the effects of node attributes on performance. Relevant node attributes include its level within the site structure, and its degree of connectivity to other nodes, which is reflected in the structural landmark measure. Other attributes help to define the structural landmarks (in1, out1, in2 and out2) or are derived from the two variables mentioned above.

The nodes mentioned and the nodes judged important were chosen as the main dependent variables for the analysis. Drawing and recall were not included because they were found to be highly correlated with number of mentions (see Table 4.2). Other dependent measures used in the analysis included total number of visits per node and the number of virtual landmarks.

An ANOVA was used to examine how a node's type (i.e., its level in the site structure and whether it was a structural landmark) affected the extent to which it was visited, judged to be important, etc. Separate analyses were carried out for each of the four sites.

The results of the analyses are summarized in the following tables. Table 4.11 shows the results for testing the judged important variable; Table 4.12 shows the results for testing the nodes mentioned; Table 4.13 shows results from the
number of visits per node; and Table 4.14 shows results from nodes identified as virtual landmarks. In all tables NS refers to Not Significant, while the F test and p (probability) values are shown for the significant effects. Only level and Slmark are included in these tables as none of the effects involving an interaction between the two proved to be significant.

Table 4.11 shows that level was a significant discriminator for all four sites at the .05 level or better. Structural landmark (yes or no) was only significant for the Montana site.

<table>
<thead>
<tr>
<th>Judge</th>
<th>Level</th>
<th>Slmark</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hawaii</td>
<td>F[4,103]=20.73, p&lt;.001</td>
<td>NS</td>
</tr>
<tr>
<td>Montana</td>
<td>F[4,1272]=537.84, p&lt;.001</td>
<td>F[1,1272]=21.42, p&lt;.001</td>
</tr>
<tr>
<td>NyeLabs</td>
<td>F[6,246]=20.34, p&lt;.001</td>
<td>NS</td>
</tr>
<tr>
<td>Web</td>
<td>F[3,34]=3.43, p&lt;.05</td>
<td>NS</td>
</tr>
<tr>
<td>Museum</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.11 - Node Effects for the measure judged importance

Table 4.12 shows that the results for mention parallel those for judged important, with level being a significant factor across all four Web sites. The only difference with Mention (Table 4.12) is that Slmark is significant for the NyeLabs sites as well as the Montana site.

<table>
<thead>
<tr>
<th>Mention</th>
<th>Level</th>
<th>Slmark</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hawaii</td>
<td>F[4,103]=23.21, p&lt;.001</td>
<td>NS</td>
</tr>
<tr>
<td>Montana</td>
<td>F[4,1272]=505.89, p&lt;.001</td>
<td>F[1,1272]=7.18, p&lt;.005</td>
</tr>
<tr>
<td>NyeLabs</td>
<td>F[6,246]=7.22, p&lt;.001</td>
<td>F[1,246]=7.89, p&lt;.010</td>
</tr>
<tr>
<td>Web</td>
<td>F[3,34]=5.48, p&lt;.005</td>
<td>NS</td>
</tr>
<tr>
<td>Museum</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.12 - Node Effects for the measure number of mentions

Table 4.13 also shows that level tended to be a better discriminator (in this case for Total Visits) across all the sites, whereas Slmark was significant only for Montana.
These results indicate that a node's hierarchy level is the best way to determine whether the node will be used often, remembered well, or judged important.

The next four tables show the relationship between level and other measures across each of the four sites used in the experiment.

<table>
<thead>
<tr>
<th>Total Visits</th>
<th>Level</th>
<th>Slmark</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hawaii</td>
<td>$F[4,103]=86.38$, $p&lt;.001$</td>
<td>NS</td>
</tr>
<tr>
<td>Montana</td>
<td>$F[4,1272]=3307.75$, $p&lt;.001$</td>
<td>$F[1,1272]=8.67$, $p&lt;.005$</td>
</tr>
<tr>
<td>NyeLabs</td>
<td>$F[6,246]=133.86$, $p&lt;.001$</td>
<td>NS</td>
</tr>
<tr>
<td>Web Museum</td>
<td>$F[3,348]=417.69$, $p&lt;.001$</td>
<td>NS</td>
</tr>
</tbody>
</table>

Table 4.13 - Node Effects for the measure number of visits

For the Hawaii site (Table 4.14), most of the structural landmarks are in the lower levels of the Site, whereas the nodes that are judged important are in the top two levels (0 and 1) of the site. The results for number of mentions are less clear-cut, with number of mentions being greatest for levels 1 and 2 (note that level 2 has a lot of structural landmarks whereas level 1 does not.)
Montana (Table 4.15) is a hierarchical site and thus has few structural landmarks which are concentrated at the top of the hierarchy. Number of mentions is again greatest in levels 1 and 2 of the site.

<table>
<thead>
<tr>
<th>Level</th>
<th># of slmarks</th>
<th># other nodes</th>
<th># of mentions</th>
<th># judged important</th>
<th># of visits</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
<td>12</td>
<td>7</td>
<td>369</td>
</tr>
<tr>
<td>1</td>
<td>8</td>
<td>0</td>
<td>84</td>
<td>16</td>
<td>478</td>
</tr>
<tr>
<td>2</td>
<td>12</td>
<td>8</td>
<td>66</td>
<td>6</td>
<td>286</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>124</td>
<td>13</td>
<td>0</td>
<td>142</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>1124</td>
<td>2</td>
<td>0</td>
<td>62</td>
</tr>
</tbody>
</table>

Table 4.15 - Relationship between Node and five other measures for the Montana Site.

The Nye site (Table 4.16) is another hierarchical site with only a few structural landmarks that are concentrated at the top of the site. The nodes mentioned most frequently are again those in levels 1 and 2.

<table>
<thead>
<tr>
<th>Level</th>
<th># of slmarks</th>
<th># other nodes</th>
<th># of mentions</th>
<th># judged important</th>
<th># of visits</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
<td>12</td>
<td>6</td>
<td>225</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>8</td>
<td>46</td>
<td>4</td>
<td>189</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>22</td>
<td>130</td>
<td>13</td>
<td>380</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>206</td>
<td>14</td>
<td>0</td>
<td>72</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 4.16 - Relationship between Node and five other measures for the Nye Site.
<table>
<thead>
<tr>
<th>Level</th>
<th># of slmarks</th>
<th># other nodes</th>
<th># of mentions</th>
<th># judged important</th>
<th># of visits</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
<td>12</td>
<td>6</td>
<td>224</td>
</tr>
<tr>
<td>1</td>
<td>16</td>
<td>0</td>
<td>88</td>
<td>17</td>
<td>402</td>
</tr>
<tr>
<td>2</td>
<td>170</td>
<td>8</td>
<td>86</td>
<td>12</td>
<td>256</td>
</tr>
<tr>
<td>3</td>
<td>90</td>
<td>68</td>
<td>1</td>
<td>0</td>
<td>50</td>
</tr>
</tbody>
</table>

Table 4.17 - Relationship between Node and five other measures for the Web Museum Site.

The Web museum site (Table 4.17) is a network, and like the Hawaii site it has a lot of structural landmarks that are concentrated in the lower levels of the site. In spite of this, most of the mentions again occur for nodes in levels 1 and 2 of the site.

### 4.10.5 Results from Questionnaires

**Subjective Size Ranking**

I also summarized the ranks that subjects assigned to the sites based on their perceived size. It appears that the perceived size of a site does not depend solely on its number of nodes. The two sites in this study that resembled a network (Hawaii and Web Museum) were judged to be bigger than the two strongly hierarchical sites used in this study. Table 4.18 shows the size rankings (1-4) assigned to each of the site by each of the subjects.
The ranks that the subjects assigned to each site are listed in rows 1-12 of the table. The mean assigned rank is also calculated and shown, along with the true rank. A Friedman two-way ANOVA by ranks was performed to see if the differences in assigned ranks between sites were significant. The statistic \( F_r = 16.9 \) was calculated according to a formula given in Siegel and Castellan (1988, Section 7.2), which is significant at the .001 level with three degrees of freedom, as occurs in this case. This result shows that there are significant differences between the four sets of ranks. Note that even though Travel Montana has three times as many nodes as the Web Museum and ten times as many nodes as Hawaii, it is ranked behind these two sites in terms of perceived size.
**Subject's Maps**

The site maps drawn by the subjects give an interesting picture of the subject's perception of the sites. The overwhelming majority of the subject maps were drawn as strict hierarchies, regardless of site structure. Of the 48 maps drawn, only five showed interlinking between branches. Of these five, three were incorrect representations of Travel Montana (which did not have any interlinking between branches but did have a recurring graphic which three subjects showed as a link between branches). Only two maps showed correct links between branches - one map of Hawaii and one of the Web Museum. The five maps with network indications were all drawn by different subjects.

For the weakly hierarchical sites, thus, only one subject of the twelve (8% of total) showed crosslinking between branches. This result would seem to indicate that even when crosslinking is present, users prefer to visualize the site as a strict hierarchy.

The number of correct links and number of incorrect links drawn by subjects in their site maps were counted. This was to see if the mental image the user had of the site matched the actual structure. Of the links that were drawn, the number of mistakes and the number of correct representations were counted. The subjects did not draw all the links on each of the sites, but tended to draw the top levels of each site.

From an accuracy standpoint, the percentage errors ranged by site from 2% to 10% of the total number of links drawn. This is a small error rate for a recall task, and would seem to indicate that the subject's mental images of the site were based on the actual site, and not on a fictitious site inside their minds.

A repeated measures analysis was run with SPSS, and it showed that there was no significant effect from size or structure on the number of mistakes. (For size, F<1, for structure, F=1). However, there was an interaction (F[]=5.81, p<.05) with the small, weakly hierarchical site having few mistakes and the large, hierarchical site having few mistakes. This can be seen in Figure 4.7.
The number of correct links drawn were also analysed. The results show that there was no significant interaction for size (F<1), no significant effect for structure (F[1]=2.21, p>.05) and no significant interaction.

A sample site map is shown in Figure 4.8. The remaining site maps from all the sites and subjects can be found in Appendix C.

**Anecdotal Evidence**

Anecdotal comments from the subjects show that they prefer good graphics or icons, the use of colour, consistent page layout and clear labeling. None of these comments is surprising. The preference for sites varied over the group of subjects, probably because the site's content material most influenced subject's choices of favorite site.
FIGURE 4.8 - Sample Map
4.11 Discussion

Several interesting results come from the data analysis, the effects of site structure, the non-effects of site size, the reasons certain nodes were remembered better than others or judged as more important and the ranking of the site's size by the subjects.

The structure of a site has a significant effect on user navigation. The strongly hierarchical sites generate significantly more "activity" (number of accesses, and number of nodes visited). This result is most likely because a user in a hierarchy needs to use several steps to get from the bottom of one branch to another. The user also needs to back up the hierarchy before moving down a different branch. In a tangled or weak hierarchy, however, the user is able to move quickly between branches and thereby eliminate the extra steps.

Weakly hierarchical sites had significantly greater second order connectivity and more structural landmarks. However, there was no significant difference between the strong and weakly hierarchical sites in terms of memorability. This suggests that larger amounts of connectivity do not of themselves increase the memorability of nodes.

None of the measures showed a significant effect of size. This result may occur because the differences in site size may not have been large enough to show significant experimental differences. However, given that all of the F scores for this test were less than one, it is more probable that the size of a site does not significantly affect the user's navigation or the site's memorability. Thus the site structure is more important than size in determining user reaction.

It was shown by the regression analysis that the nodes remembered well tended to be those that were visited the most often. The fact that the effect of structure on memorability was only borderline significant may have resulted from extraneous (for the purpose of this study) inter-site variations (e.g., in terms of node content, aesthetics or the Web sites, etc.) The number of visits measure was also by far the most powerful predictor of judged importance.
This indicates that the more a user views a node (by visiting), the more likely the user will remember the node and judge it important.

In looking for a discriminator to predict the number of visits, it was found that placement of nodes in higher levels of the hierarchical structure are far more important in defining the importance of a node and the memorability of a node than are its connectivity and function as a structural landmark. The results from the analyses of variance indicated that the node level has a significant effect on its perceived importance whereas whether or not it is a structural landmark (as defined by the Mukherjea and Foley equation) does not. This indicates that a node's placement (in terms of being near the top of the structure) is more important than the number of other nodes connected to it (both one and two jumps away). Tables 4.5 to 4.8 showed that judged importance is related to the frequency of visits. Tables 4.15 to 4.18 show that level is an important determinant of both mentions and judged importance.

Structured landmarks tended to occur lower down in sites (as shown in Tables 4.11 to 4.14), whereas it was the nodes higher in the hierarchy that were more likely to be memorable and to be judged important. Thus the present results suggest that Mukherjea and Foley's equation for predicting structural landmarks is probably not very useful when applied to the types of hierarchical structures used in this study (and these hierarchical structures are in turn a representation of many Web sites). For all four sites, nodes in levels 1 and 2 were mentioned far more frequently than nodes in other levels. For hierarchies there are few structural landmarks, and those landmarks are concentrated near the top of the site. For the networks there are many structural landmarks, and these landmarks are located in the lower levels of the site.

The NyeLabs site had the most nodes mentioned. This site had very distinctive pages and made good use of colour and icons. In contrast, the Montana site had a very similar background for each page and used a similar graphic on many of its pages. Such differences may have influenced the subjects' responses. It is expected, though, in a situation where sites with different structures had similar content and graphics, or if this experiment were repeated on a larger scale with more sites, that it would be found that
the hierarchical sites would generate a greater number of nodes mentioned and judged important. This is because the strongly hierarchical sites generate more activity, visit node more frequently, which will lead to increased site memorability.

The structure of a Web site appears to have a very significant effect in terms of the ranking of subjective size. It is interesting to note that the weakly hierarchical sites appear bigger to users; the largest site that was used in this experiment was judged to be smaller than the smallest site used because of differences in structure. This finding is very useful to Web site designers, because they can use the structure of their site to influence the user's perception of the site. A designer with a lot of information to put on a site, who wants users to perceive a small and manageable site should choose a strongly hierarchical structure. Crosslinking between branches should probably be used judiciously.

The maps that the subjects drew show an interesting trend in that the majority of the maps (for both types of structure) reflect a hierarchical mental map of the sites. The results of the number of correct links drawn on the maps and the number of mistakes seem to indicate that the subject's mental maps of the sites are based on reality and not on their own mental representations. However, further research in this regard is needed.

The trend across all the results is that hierarchical sites are easier to use. This can be seen from the higher number of accesses and nodes visited on the hierarchical sites, the fact that hierarchical sites are perceived as smaller and the uniform tendency for users to draw their mental maps of Web sites as strongly hierarchical structureless. The question arises, why are hierarchical sites easier to use? This result could just indicate that the better Web site designers, who are more organized, who think through their designs better, etc., tend to build hierarchical sites. Or is it that a hierarchical site fits better with human nature and perception? The results from the subject maps seem to indicate that users already have a hierarchical view of Web sites. Thus it seems hierarchical sites are better because they correlate better to the user's mental image of the Web site. This idea though needs further research and investigation.
In general, because only four sites were used, one to represent each size/structure combination, the results of this experiment may be reflecting idiosyncratic tendencies between the sites (e.g. the difference in use of colour between the sites). In particular, any tendencies involving the interaction between size and structure of sites must be interpreted cautiously because the interaction may be due to the idiosyncratic behaviour of one site.
Chapter Five

Conclusions

5.1 Introduction

This thesis consisted of three main activities: review of theoretical and practical issues relating to construction and navigation of Web sites; review of current Web sites and selection of four sites for further study based on both their quality and their relative size and structure; conduct and analysis of an experiment that demonstrated the effect of various site and node parameters on the judged importance and memorability of nodes, and the apparent size of sites.

The experiment was concerned with how the structure and organization of a Web site affects the way that people browse it when attempting to answer questions or find information. Four different Web sites were studied, a large and a small hierarchical site, and a large and a small network site. Twelve subjects each browsed all four of the sites, in different orders. Their activity in browsing each site was logged in terms of which nodes were visited. In addition, at the end of interacting with each site, subjects were asked to: draw a picture of each site (which in almost every case was a tree structure, i.e., hierarchy); indicate which nodes were "important"; to name the nodes that they could remember. When they had browsed through all four sites, subjects were asked to rank order the four sites in terms of how big each site was (i.e., how many nodes it had).
The measures collected were related to the size and structure of the sites and to the connectivity of individual nodes within the sites. Basic connectivity was assessed in terms of four measures (first and second order backward, and forward, connectivity). In addition a measure of the degree to which each node was a structural landmark was calculated, based mainly on the second order connectivity of that node. The main results obtained in the experiment will be listed in the following section, along with other contributions made in this thesis.

5.2 Contributions

This thesis identified four Web sites that provide a good contrast in terms of two levels of size and two levels of structure (which can be viewed either in terms of whether that structure is more or less hierarchical, or in terms of the percentage of nodes that are landmarks).

A methodology was developed that used logging of navigation through a site, along with collection of judgments concerning each site and the nodes therein, to assess the effects of site size and structure on how people responded to it in terms of memorability and other parameters.

The experiment yielded a number of key results that have implications for Web design.

5.2.1 Results pertaining to the Site Level

- Site structure has greater effect on user navigation than site size specifically:
  - Hierarchical sites have a greater number of node accesses
  - Hierarchical sites have a greater number of nodes visited

- Site structure affects a user's perception more than size does specifically:
• Strongly hierarchical sites are perceived as being smaller than weakly hierarchical sites (in this study a site with a couple of hundred nodes was judged to be bigger than a size with over a thousand nodes).

• Users visualize sites as hierarchies, as indicated in their drawings of the sites after navigating them.

• The summary of all these findings is that users tend to find strongly hierarchical sites easier to remember.

Results at the node level

• The frequency of node visitation is the best predictor of memorability (judged important and mentioned) and

• Node level is a better predictor of node visitation than structural landmarks

5.3 Applications

These findings have applications for Web site designers. First, they give the designers an educated way to choose the structure of a site. A designer with much information who nevertheless wants his site to "feel" small and not overwhelming, can use a hierarchical structure.

In general, though, it appears that a strongly hierarchical site is preferable, if the goal is to increase the chances that nodes will be remembered. A strong hierarchy also seems to conform to people's expectations about how a Web site should be structured. Providing links between branches may facilitate navigation, but it may also confuse users who prefer to navigate vertically. However, it may be possible to construct links between branches that would not interfere with those users who wish to keep using a tangled hierarchy.
Important information should be placed in nodes near the root, for there it is most likely to be remembered by subjects. A node's level is the best predictor of its being remembered and visited, but structural landmarks did prove significant for some variables; so important information should be placed on nodes that are both near the top of the hierarchy and well connected to other nodes.

### 5.4 Recommendations for Future Work

The World Wide Web is an interesting and relevant area to concentrate research efforts on. It is also a good area to work in on account of the receptiveness of experimental subjects. As the area of research is one of such interest, the subjects enjoyed the experiment and were keen to provide their opinions on the Web tested sites. This importance of subject participation cannot be overestimated.

Future work is certainly needed in this area to understand better how users use and perceive the Web. With regard to this thesis, further work is needed in the effects of site size and structure on user navigation and perception. A larger experiment with more sites would help to pinpoint differences caused by size and structure. If another experiment were to be done, sites that were as similar as possible in terms of content and graphical layout would help to magnify existing differences. Also, sites covering a larger size range would help to show existing differences. This would counter any negative effects this experiment suffered from idiosyncratic differences in Web sites used in this experiment. While the results from this experiment seem reasonable, verification of the results and extension of the ideas developed in this thesis are needed.

One interesting experiment would be to take an existing site and to modify its structure (for example changing a strongly hierarchical site to a weakly hierarchical site or vice versa) and then examine the effect of the differences between structures in terms of user navigation.
Experimental work also needs to be done to see what effect differences at a node level play on user navigation and perception. For example, what is the role graphics have in this area? Are nodes with different background colours perceived differently from those with the same colour?

Further work on the mental images of Web sites users have would also be interesting. All of the subjects used in this experiment had had previous experience with the Web. It would be interesting to see if there was a difference in the way novice and expert users perceived the structure and size of Web sites. Further work could also be done to verify that user's mental maps of Web sites mirror the actual, as opposed to users mentally creating their own Web site. This ties into the other interpretation of connectivity between nodes, where the links between nodes are a result of user's perceptions instead of the links built in by structure.

Taking the other extreme, there also needs to be research at a Web-wide level to see how users navigate through the Web as a whole, and how they perceive the digital space of the Web as a whole.

### 5.5 Conclusions

This thesis has attempted to do some exploratory research into an area where there has been little scientific research, and none in the area this thesis touched on. It is an area where research is needed, because of the increasing popularity and use of the Web. It showed that the effect of actual site size upon user's navigation and perception appears to be less important than the structure and organization of the site. The structure of a Web site has a large impact on how users navigate in the site, the memorability of the site, and its perceived size. The strongly hierarchical sites receive more navigation activity, are better remembered and are perceived as being smaller than the weakly hierarchical sites.

This thesis has examined differences at a site level in terms of size and structure. It had also tried to show what features about nodes in a site make
them more likely to be used as behavioural landmarks. The idea of structural landmarks has already been documented in literature. The idea of behavioural landmarks is derived from work done by Lynch, and though having been used before in application to the Web (by Ingram and Benford, etc.) there has been little research on what causes behavioural landmarks in Web sites, nor on how users use the Web sites.

In terms of predicting judged importance and memorability, whether a node is a structural landmark is not as critical as its level in the structure. This suggests a useful heuristic for Web design, for example, place important nodes that need to be remembered in the upper levels of hierarchical structures. This strategy will likely work better than simply increasing the backward connectivity of "important" nodes.

The results obtained in this thesis fit well into the human/computer interaction theory concerning hierarchies and their importance in cognition. This thesis and experiment were designed as exploratory research into a new area, and it is hoped that other researchers will be able to verify these findings and build on them in future research.

Areas for future research include looking in detail at what causes a node to be treated differently from other nodes, and also how users navigate and perceive the Web as a whole.

The results obtained in this research call into question the application of the Mukherjea and Foley equation to Web site design. This research shows that hierarchical structured sites do work well and that a greater degree of hierarchical structure seems to improve the memorability and navigability of a site. However, the Mukherjea and Foley equation loses much of its effectiveness with hierarchies, where it tends to predict that nodes in the lower levels of the hierarchy will tend to be landmarks, whereas it is the upper level nodes that are more frequently visited and that are more memorable. In fact, as shown in the results of the experiment, although structured landmarks tend to be associated with level of hierarchy, level remains a superior predictor of importance and memorability, and also is
better at accounting for differences in variables such as the total number of visits to each node.

The results of this thesis should assist Web site designers in choosing site structure and information location. This thesis represents exploratory research into the structure and use of Web sites. It provides information that will enable Web site designers to make an informed choice on their use of site structure, and provides a useful heuristic to guide the placement of important information.
References


Gray, Matthew, b. "Growth and Usage of the Web and the Internet". Published online at URL: <http://www.mit.edu/people/mkgray/net/>. 1996.


Meadow, C. *Personal Communication with Author*, 1996.


Appendix A

Subject Consent Form
Subject Instructions
Experimental Questions
Subject Consent Form

Information Retrieval from Web Sites

I hereby agree to act as a subject in an experiment entitled "Information Retrieval from Web Sites" that is being carried out under the auspices of the Industrial Engineering Department at the University of Toronto.

I have been given a full description of what I shall be required to do in this investigation. I am aware that I may withdraw from the investigation at any time, and that I have the right to ask in that case for any data collected about my performance to be given to me or destroyed.

I understand the experiment will last approximately one hour and I will be paid $10 CDN per hour for my participation. I consent to take part in this experiment voluntarily and without any coercion.

Name: ____________________________

Signed: ____________________________

Date: ____________________________

Investigator's Name: Anna V. Chaffey

Supervisor's Name: Prof. Mark Chignell
Subject Instructions for Experiment

(Note: These instructions were read out loud by the experimenter to the subject.)

Here is how the experiment works. Each experimental trial takes place on a web site. The experiment is like a scavenger hunt for information. There will be four trials, each lasting about 25 minutes. Please remain within each web site for the duration of that trial. If you accidentally venture off, the I will ask you to return to the most recent page on the web site. Please use features of Netscape as usual, (for example, the back button, the history list or other features). You will be given 5 minutes at the beginning of each trial to explore the web site, and to find your way around it. Then I will ask you ten questions, the answers to which will be on the web site. When you find the answer, please read it out loud (or point to it on the screen), I will then give you the next question. If you are having real difficulty in finding the answer to a question, you may "pass" on the question (as on a game show). If you have not found an answer within 3 minutes, I will proceed to the next question.

After each trial I will ask you a few questions about the web site. At the end of the entire experiment I will ask some summary questions about your perceptions of each web site.

Your web navigation will be recorded by the computer. After the experiment is complete, I will analyse the navigation logs to see what pages you visited.

Do you have any questions?
Experimental Questions for the Hawaii Visitors & Convention Bureau Site

How old is the Big Island (Hawaii)?

What are the three different ways to eat in Hawaii?

Where can you obtain a marriage license?

How many places are there for you to stay in Hawaii?

What is the first link on the links page?

What information do they ask for on the feedback form?

What kind of island is Kauai?

What does Hokeo mean?

What other languages are the pages available in?

Who was the latest winner of the customized Hawaiian vacation?

Experimental Questions for the Travel Montana Site

Who is the governor of Montana?

What are the six tourist regions in Montana?

How many visitors visit Montana's state parks each year?

How many hostels are available in Custer Country?
What address would you write to to find information on Montana's Highway Patrol?

What events took place in the State Parks during March 1996?

How many State Parks are in Missouri River Country?

How close to Glacier Park are the Northern High Plains Outfitters, Browning?

Which city in Gold West Country begins with the letter "O" and has outfitters?

Does the restaurant in Custer take reservations?

**Experimental Questions for the NyeLabs Site**

How many photos are on page two of the Nye Labs Photo Album?

When did Bill write the memo?

Who is the Assistant Production Co-ordinator for the show?

When does "Bill Nye the Science Guy" show in Toronto?

How much does a Bill Nye home video cost?

What were the new items added to the web site February 29 1996?

What is the current "random home" demo?

Who wrote Bill's favorite book on extinction?
What is today's PBS show topic?

Where is the Nye York Times printed?

**Experimental Questions for the Web Museum**

When was Edgar Degas born?

What is WebMuseum's definition of "classicism"?

What is the Trés Riches Heures?

Who took the photographs in the Paul Cézanne exhibition?

What is the time span covered by the term "Japanese Art and Architecture"?

What is the current name of the city of Edo?

What size is the picture "The Lictors Bring to Brutus the Bodies of His Sons " by Jaques-Louis David cover?

The bones of how many people are in the Paris catacombs?

What new pages were added on February 4 1996?

When is the WebMuseum open?
Appendix B

Questionnaire for Use After Each Site
Questionnaire for Use at the End of Experiment
Questionnaire for Use After Each Site

Subject ID:_______ Site:_____________ Start Time:
End Time:

1. Please name or describe all the pages you can remember from this site.
   (list answers below)

2. If you were describing this site to a friend, what would you tell them?

3. Which one or two pages on this site were the most distinctive? Why?

4. Please draw a small map (or representation) of the site (on back).

5. Please circle on your map the pages you feel are the most important.

PAGES:
Questionnaire for Use at the End of Experiment

Subject ID:____________

1. Which site was your favourite? Why?

2. Which site did you find easiest to use? Why?

3. Which site do you think was designed the best? Why?

4. After your experience with these different sites, in your opinion what features make a site easier to use than others?

5. Please rank the sites in order of size:

   (Largest) 1.

   2.

   3.

   (Smallest) 4.
Appendix C

Subject Site Maps
Hawaii Visitors and Convention Bureau Subject Maps
HOME - (ISLANDS) - HAWAI'I
  - KAUA'I
  - NI'AU'I
  - HONOLULU

- ACTIVITIES
- ACCOMMODATIONS
- EVENTS
- FEEDBACK
- JAPANESE PAGES
- HOLEO

(JAPANESE (unexplored))
CHAMBER OF COMMERCE
still under construction
(not all links work)
Hawaii / Honolulu / Japan

Activities
Accomodations

Green Island

What you can do in Hawaii

What to live
Travel Montana Subject Maps
NyeLabs Subject Maps
The Web Museum Subject Maps
Department of Mechanical and Industrial Engineering

CLEARANCE SLIP

Name of student: ___________________________ Date: __________________
Degree Due: _____________________________
Supervisor: ________________________________

This is to certify that no item is outstanding toward the above-mentioned person. On completion of a graduate degree ALL keys must be returned. Any specific key requirements will be reissued by the Graduate Studies Office on the basis of post-degree studies/employment.

Stores - Dan Kalra, MC 121
Technical Services - Peter Loite, MC 110

Computer A/c - Alan Milner, MC 229
Electronics Lab - Miro Kalovsky, MC 402B

Pending bills -
Mary-Rose Naudi, MC 123A

Supervisor

Thesis on-loan - Graduate Studies Office
RS 214

Coffee Club -
Julie Verduci, MC 136

Keys returned: ________________________________
NOTE TO USERS

The original manuscript received by UMI contains pages with indistinct and slanted print. Pages were microfilmed as received.

This reproduction is the best copy available

UMI