Chapter 1

Introduction

Life seems to be filled with getting from one place to another. The ability to successfully navigate one's way from one place to another is a skill vital to most people's lives.

1.1 Navigation in Distributed Information Systems

The modern age has given rise to a new navigation problem in the use of Distributed Information Systems (DIS). Rocha (1999b) describes DIS as follows:

Distributed Information Systems (DIS) refer to collections of networked information resources in interaction with communities of users; examples of such systems are: the Internet, the World Wide Web, library information retrieval systems, etc.

Many modern information systems indeed consist of a heterogeneous mixture of information resources that are linked over electronic networks and simultaneously used by large groups of users. Each individual system in the DIS can contain a collection of texts, images, audio and video files, which can be linked across different platforms and databases. A document stored on the servers of a digital library can e.g. contain an electronic link or pointer to a document on a different system\(^1\).

Within DIS, users have a number of options to locate and retrieve relevant information. They can use an absolute address, databases or a number

\(^1\) I will refer to all individual containers of information as documents. I refer to documents that consist mostly of text as pages.
of search services as they are available. User can, however, prefer to locate relevant information by traversing the electronic links within the system. This process is often likened to "navigation" since from the user's viewpoint he or she actually traverses a path from one document to another, whether or not any actual change in physical location is concerned. This kind of electronic navigation can consist of the traversal of menu-structures, the use of in-text links, selection of citations within an electronic articles bibliography, etc. The main point is that the user's experience is one of navigation: choosing certain links and paths to "move" from one location in the DIS to another, more desirable one.

The issue of navigation is certainly pertinent to DIS systems such as the World Wide Web (WWW), which was specifically designed to allow large, distributed collections of documents to be linked into a single network for retrieval by user navigation. The WWW's networked nature allows the integration of a large number of information resources whose content can efficiently be made available to ever larger communities of users. At present it contains a startling number of pages and hyperlinks, and has developed into the single largest information network ever constructed.

However, due to its size, users feel overwhelmed by the amount of relevant and irrelevant information and frustrated by inadequate navigational links. Users get confused and are often not able to navigate toward information that is available yet buried in large amounts of irrelevant trivia. At the same time, our understanding of the actual mechanisms involved in DIS navigation is poorly developed.

1.2 Distributed Design

Due to their distributed nature, and the heterogeneity of their content, there is not one specific designer who elaborately crafts the interface and the structure of links and connections among the many services networked into a DIS. Rather a large number of providers each design their own services, which are then again linked to the larger existing network of services. This is certainly also true for the WWW. Given the lack of central control over content and structure, the problem arises of how we can generate DIS systems that users find easy to navigate, but conform to the constraints imposed by the specific information resources
in the system.

1.3 Two Main Questions

The focus of this dissertation is not on the general class of DIS, but more specifically on the WWW. The WWW in many ways represents the linked features of DISs (networked structure), but at the same time provides a platform strongly reduced in complexity and the idiosyncrasies associated with specific DISs. In terms of structure, it is little more than a linked collection of documents. The WWW is therefore used as a more general representation of a class of systems for non-linear text commonly referred to as "hypertext".\(^2\)

The dissertation is concerned with two research problems regarding the use and design of hypertext systems, and more specifically the WWW:

1. Development of a model of the mechanisms involved when users attempt to locate relevant information within a hypertext system's link structure.

2. Methodologies that support this kind of retrieval by adapting their link structure.

The aim is to develop a better understanding of human DIS navigation that can support novel approaches to their design.

The next introductory chapter consists of:

1. A short introduction to the basic concepts regarding hypertext systems and the WWW.

2. Introduction to the main results and problems concerning human hypertext navigation.

3. Introduction to the main results and problems concerning human hypertext design.


At the end of the following chapter, I hope the reader will not only understand the main research questions I address, but also appreciate the underlying technical issues and the state of the research as it now stands on human hypertext design and navigation.

\(^2\)The next chapter will expand on this point.