Redefining the Hyperlink

Dominic S J Stanyer

Doctor of Philosophy

University of Edinburgh

2001

Declaration

I declare that this doctoral thesis was composed by myself and that the work contained therein is my own, except where explicitly stated otherwise in the text. The following articles were published during my period of research. Certain material and concepts from these publications will necessarily be presented within the body of this work.

1. Stanyer, D. and Procter, R. Improving Web Usability with the Link Lens. In Mendelzon, A. et al. (Eds.), Journal of Computer Networks and ISDN Systems, Vol. 31, Proceedings of the Eighth International WWW Conference, Toronto, May, 1999, Elsevier, p 455-66.

2. Stanyer, D. and Procter, R. Human Factors and the WWW: Making sense of URLs. In Cockton, G., Cawsey, A. and Brewster, S. (Eds.), Volume II of the Proceedings of INTERACT'99, Edinburgh, September, 1999, p. 59-60.

3. Stanyer, D. and Procter, R. Link Lens: An Enhanced Link User Interface for Web Browsers. In Stephanidis, C. and Waern, A. (Eds.) Proceedings of the 4th ERCIM Workshop on User Interfaces for All, Stockholm, October, 1998, p. 34-47.

Acknowledgments

I would like to thank Dr. R Procter for his help, guidance, inspiration and encouragement, and my parents for all their support.

Abstract

The number of people using the World Wide Web (WWW) is increasing. These new users have a range of educational backgrounds and belong to diverse cultures from all over the planet. The user interface to the WWW must therefore exploit the fundamental principles of human-computer interaction and not rely on the use of learnt conventions if it is to be successfully used by everyone.

By advocating an approach to interface design based on the work of Suchman and the observation that actions are situated in the particular circumstance in which they occur, this thesis argues that the simple hypertext abstractions of the Web page and the hyperlink, prevent the system from revealing the Internet technology below. A principled investigation into WWW use shows that this seamless hypertext abstraction is very rarely sustained in practice and users struggle to explain dynamic system disturbances and prevent worthless hyperlink activations. The investigation specifically reveals the problems that users encounter and the strategies they employ in their attempts to circumvent these abstraction barriers and probe the system for more appropriate information.

The findings of this investigation highlight the Universal Resource Locator (URL) as the most important resource available to the WWW user when resolving a system disturbance or predicting the content of a referenced Web page. Experiments are subsequently undertaken to uncover first, what type of system information users can extract from the URL and secondly, how confident and accurate users are at extracting such information.

The understanding of WWW use elicited through the observational investigations and the directed experimentation, is then employed in the design of an extension to the WWW browser – *the hyperlink lens* - which redefines the hyperlink abstraction by revealing to the user more appropriate information about the data transfer process and the referenced Web page.

Finally, the implications of this interface design philosophy are discussed and evaluated and the impact that a new interface would have on WWW use is evaluated.

Table of Contents

Chapter 8 The Hyperlink Lens

List of Figures

List of Tables

Chapter 1

Introduction

There is a large body of literature concerning human computer interaction (HCI). It arises from a range of disciplines in science and the social sciences and there has traditionally been an open policy of accepting contributions from all areas. The result is a diverse literature where many different perspectives and methods are applied to the problems and questions surrounding the topic.

The region where interaction occurs between two systems is the interface. The *user interface* of a system is the area where the human user and the system interact: the place where the user and the system ought to be able to reciprocate actions. The concerns of HCI fall into two broad areas: the study of the major phenomena surrounding the interaction between a human and a computer and the design and implementation of the user interface where the interaction occurs. This statement and the term human-computer interaction may lead to a focusing on a single user operating an application on a single computer. It is clear from the literature, however, that this is not the only situation that HCI addresses. The use of the word *computer* in the term does not restrict the devices under study to stand-alone workstations. The interaction between humans and all forms of embedded computational machines are within the research remit. Similarly, the interaction between groups of humans through distributed systems, the cooperative nature of work and the communication between them, is another area that is well within the sphere covered by HCI.

The work presented in this dissertation covers the interaction between a user and a distributed information space that uses a *network* of computers to store and transfer data. The term human-computer interaction is therefore used throughout this dissertation to signify a perspective that focuses on the relationship between humans and distributed computer systems: the particular interface to the system where the interaction occurs; the language of the interaction; and the capabilities of the human to deal with the system, and the system to deal with the human. This dissertation is concerned with the human-computer interaction of one particular distributed system, an application that uses the world-wide computer networks of the Internet to transfer data. This system is called the World-Wide Web (WWW).

1.1 Motivation

The WWW is a globally distributed information space. The authorship of WWW information objects – *Web pages* – is decentralised, unregulated and technically simple. These factors have contributed to the production of an extensive, culturally and contextually diverse mass of information. Information about anything can be published by anyone and read by everyone. The mechanism provided to maintain order in this mass of information is the *hyperlink*. The hyperlink defines an association between two information objects by being situated in one and referencing another. The user of the WWW is able to reveal this association by explicitly *activating* the hyperlink. By activating a sequence of hyperlinks, the user can peruse and digest a trail of associated information through the information space; the user can *"browse the Web"*.

The infrastructure of the WWW is the Internet, a world-wide inter-network designed to connect different forms of communicating networks. The technology behind the Internet provides a method for addressing every computer connected to the Internet and a standard protocol for sending data between each of them. Internet technology can be used when there is a need to transfer data between remote computers. The WWW is just one of the applications that employs this technology.

The number of Internet users around the world is rising constantly. The Computer Industry Almanac reported that by the end of 2000, 375 million people around the world would have Internet access (Petska-Juliussen and Juliussen, 2000). The main reason for the increase in Internet users is the introduction of applications such as email and the WWW. The WWW has the potential to be used by a vast number of users from a range of cultural and educational backgrounds. Making the WWW easy to use for everyone is a big challenge to those who work in HCI.

This dissertation presents the argument that the user interface of the WWW unhelpfully hides the mechanisms of its underlying technology behind the rigid and opaque abstractions of the Web page and the hyperlink. When using the WWW, users experience system behaviour that is directly related to the workings of these underlying mechanisms and consequently the user often needs to question the workings in order to elicit an explanation. This is never more apparent than when system disturbance or failure occurs and the user is left to cope with a limited amount of accurate system information for acting appropriately and effectively. It will be argued in this dissertation that the WWW interface is not a region where interaction occurs naturally between the human user and the system, but rather a region where interaction is constrained. Users often make poor decisions because of the limited system information made available to them; this results in ineffective use of the system and bad usability.

The research work undertaken for this dissertation focuses on **when** users need to circumvent this abstraction obstacle to uncover the mechanisms below and **what** diverse strategies they actually employ to do so. The result is an extensive and principled investigation into WWW use, the findings of which lead to an understanding of the interaction between the WWW user and the system. These findings are subsequently employed to propose a new design approach for the WWW user interface.

1.1.1 The London Underground Diagram

An illustrative example of an abstraction restricting the interaction between a user and a system is given by the London Underground Diagram (Figure 1.1). The diagram provides an interface between the *tube* traveller and a potential journey across London. The diagram's purpose is to help the traveller get on the right train at the right station, make the right connections and to get off at the right destination. The first map of the London Underground was produced by cartographers in 1906 and although some topographical distortions were introduced in the subsequent years, it was not until 1933 that the draughtsman Henry Beck conceived the now famous abstract schematised diagram (Garland, 1994). The emphasis of the diagram is on the connections between stations rather than the stations' geography (Figure 1.2). This is achieved through the straightening of the lines into horizontals, verticals and diagonals, and by evening out the distance between stations. An abstraction of connectivity is imposed on the geography of the stations and tracks.

The diagram is an excellent piece of graphic art with its bright, clean and colourful design, coupled with the skilful distortion of geography, and although the diagram has undergone continual improvements and growth since 1933, the basic features of the abstraction remain the same. The outlying portions of the diagram are compressed in contrast to the central area, all surface detail is absent (with the exception of a stylised River Thames) and interchange stations are presented differently from standard stations. No single feature dominates the diagram ensuring that there is no focal point to be considered as the centre of the railway system, or the centre of London. The diagram is a simple and consistent interface to the underground railway that can be used by a traveller to plan any journey on the railway network.

It is however, no more than a simple diagram, a single static abstraction imposed on a complex railway system that is prone to disturbances and failures. Although the user, in some sense, can operate the diagram by planning routes and connections across the railway network, the diagram cannot interact with the user. The diagram cannot provide any reciprocation to the actions of a user. The diagram cannot adapt and change the abstraction it presents; it cannot reveal any more (or less) detail of the underground railway system or of the geography of London to the user should they, in a particular circumstance, require it to.

A first-time visitor to London, arriving at Heathrow and wishing to travel to Trafalgar Square is immediately confronted with an abstraction barrier from the diagram. Where is Trafalgar Square and which is the nearest station to it? The diagram cannot help. The traveller must circumvent the abstraction by referring to a map, a guidebook or asking a helpful passer-by. Only after discovering the nearest station, can the visitor plan and commence their journey.

Figure 1.1: The London Underground Diagram

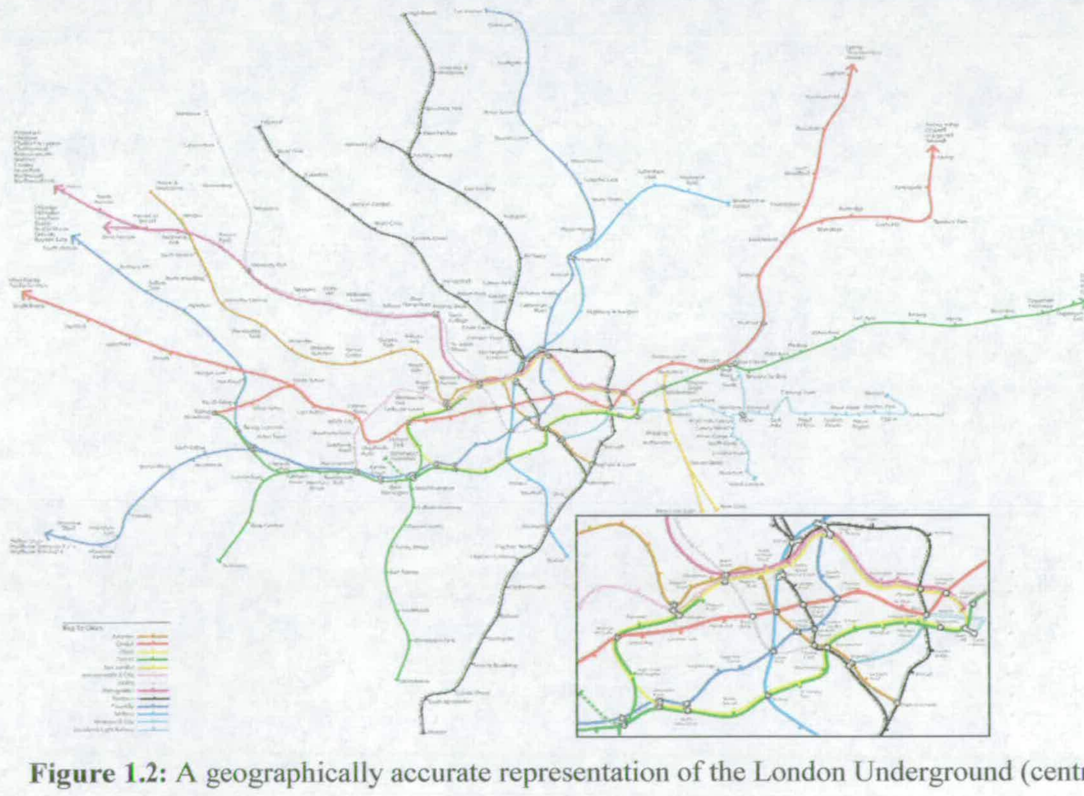


Figure 1.2: A geographically accurate representation of the London Underground (central area enlarged).

The visitor decides to take the Piccadilly Line eastbound through nineteen stations, changing at Piccadilly Circus to the Bakerloo Line and travelling one station southbound to Charing Cross. After thirty minutes of the journey, our visitor will only be at Northfields. After referring to the diagram, our visitor may well be shocked to notice how little they appear to have travelled. The geographical distortion of the abstraction making the extensive journey right through the suburbs of London appear to be only a very small (and simple) part of the journey. Tired from travelling and now expecting a long journey, our visitor's eyes may close. Twenty-two minutes later as our visitor sleeps the train is speeding through Piccadilly Circus, Leicester Square, Covent Garden on its way north to Arsenal and Cockfosters. The abstraction that is imposed by the diagram to help users select the right connections hides the spatial and related temporal properties of the actual rail travel. There is no way for users to reveal these temporal properties should they wish to do so.

At least this could have been the case if our weary visitor had not been woken by an announcement that due to engineering works everyone should alight at South Kensington and find an alternative route that avoided the Piccadilly Line. Having disembarked, our visitor plans another route taking the Circle Line five stops to Embankment and then taking either the Northern or Bakerloo Line northbound. Having arrived at Embankment our visitor changes to the Bakerloo Line and awaits a train. After a considerable wait, our visitor alights at Charing Cross and finds that they must walk some distance underground before entering the main station complex (our visitor is in fact in the old Trafalgar Square station that was closed and linked to Charing Cross station). On leaving the station, our visitor immediately realises that Embankment station is less than one hundred metres down the road, and by alighting there and walking, they could have saved a considerable amount of time.

Would a geographical map of London been more useful to our visitor? Could a timetable have helped to complete the journey more effectively? The answer is that both of these alternative interfaces to London and the underground railway system could have been appropriate at some point in this journey and for this particular user. For another user (perhaps a more experienced one) and for another journey the diagram may have sufficed. The point is not that the diagram is not useful, it most definitely is. It is just that the diagram cannot provide any other perspective on the underground railway system. The diagram is a static abstraction.

1.1.2 The World-Wide Web interface

In contrast to the example of the London Underground Diagram, the WWW user interface is in a position to reciprocate actions, to receive inputs from a user and the network, and react with dynamic outputs relevant to the particular circumstance. The argument presented in this dissertation is that the interface does not achieve its full potential. In a similar fashion to the London Underground Diagram, which hides the realities of a complex railway system, the WWW interface presents a single static abstraction of an information space that conceals the underlying dynamic mechanisms of the Internet technology. This in turn, can be the cause of unsatisfactory usability; for example, the retrieval of inappropriate information and the unexplained existence of long temporal delays.

The user interface of the WWW presents an abstract model of the information space imposed on the underlying Internet technology. The two fundamental parts of this abstraction are the Web page and the hyperlink. The Web page is an abstraction of a single information entity imposed on a collection of data files of different types and stored in different locations. The hyperlink is an abstraction of an association imposed over the mechanisms for locating and subsequently transferring all the data that constitutes the referenced Web page, from their various storage locations to the user's rendering system. These two abstractions are central to the interface hiding the real workings of the system. By redefining the model of a Web page and a hyperlink, removing the abstraction when appropriate and presenting a more adaptable user interface, the usability of the system could be transformed.

The remainder of this first chapter presents a review of the relevant background material that will aid the reader in understanding the main issues concerning abstraction and the improvised nature of user action. From this understanding, the reader will better appreciate the decisions that were made throughout this work that ultimately led to the reported findings.

1.2 Affordance

Norman (1988) introduced the concept of the designer's conceptual model as one of the major factors in understanding how users operate devices. Norman argued that the designer's conceptual model was crucial to the success of the interface as the presentation of the interaction must always be consistent with this model. The usability of the system was then determined by how easily a user accepted and understood the designer's conceptual model. In conjunction with the conceptual model, Norman also introduced the concept of affordance to HCI. Affordances specify the full range of possible manipulations provided by a device, but these are constrained through interface design to present only the desired and relevant actions to the user in order to support the conceptual model.

The theory of affordance was introduced by Gibson (1966) to describe the relationship between the actionable properties of any object and an actor. All the information necessary for understanding the *meaning* of all things in an environment can be perceived directly. Norman used this theory to explain how people managed to operate the objects they encountered in everyday life, even if they have not seen them before. The information required to operate water taps, light switches, and doors is already provided by the world through their appearance. The buttons on a mouse afford clicking, the cursor on the screen affords pointing, and the keys of a keyboard afford pressing. Norman (1999) argues that these objects provide *real* affordances. In contrast, most objects on the computer screen only provide *perceived* affordances that are under the control of the interface designer to advertise the existence of the real affordance. Real and perceived affordances are therefore two separate design concepts both of which can be independently manipulated in system design. Some real affordances of the system may be concealed by an *abstraction* and some perceived affordances might highlight nonexistent real affordances. Good design is therefore defined as the manipulation of these affordances: *"the art of the designer is to ensure that the desired, relevant actions are readily perceivable"* (Norman, 1999).

For an object on a computer screen to provide any perceived affordance, it must allow the user to anticipate what actions can be performed upon it. A well-designed object affords what manipulations can be performed on it and therefore conveys what manipulations to avoid. A classic example from the real world is of the door handle that affords whether one should push or pull in order to open the door. There exists, of course, a real affordance to both push and pull the door, but the good design of the handle must hide one and allow the user to readily perceive the other. On the computer screen, the visual representations must do the same: should the user click, or double click on a particular object. If the user can perceive which is the right action to perform to manipulate this object then it provides an affordance to the user. If, however, the user has simply learnt through experience to double click on this particular object then the visualisation affords nothing. The fact that the user can still easily use this object is because of what Norman refers to as a convention or cultural constraint.

> *"Cultural constraints are conventions shared by a cultural group. The fact that the graphic on the right-hand side of a display is a 'scroll bar' and that one should move the cursor to it, hold down a mouse button, and 'drag' it downward in order to see objects located below the current visible set (thus causing the image itself to appear to move upwards) is a cultural, learned convention. The choice of action is arbitrary: there is nothing inherent in the device or design that requires the system to act in this way." (Norman, 1999)*

It is fair to say, therefore, that the fact that a Windows user must double click on an icon on the *desktop* to start a program is simply a convention. There is nothing provided by the icon, the screen or the mouse, to lead a user to perceive this information for themselves. It is simply prior knowledge that allows the user to understand what action to perform. Affordances are relationships between objects and actors, and they exist as properties of the environment. A convention is simply a learned set of actions, shared by a cultural group.

Norman's philosophy on interface design insists that good interface design must use all three of these concepts – real affordance, perceived affordance and conventions – to convey the designer's conceptual model of the system to the user. One method for manipulating these three concepts is abstraction and its use in metaphor.

1.3 Abstraction

Abstraction is used in system design to selectively hide the complexities of a system, to strip away and conceal its fundamental mechanisms. A change of abstraction allows a designer to consider a system at a different level of detail. Traditionally, the primary concern of system design has been the conception, placement and manipulation of abstractions - the binary signal imposed on a continuous voltage; the instruction set and register files imposed on the transistors and electrical circuits; the abstractions employed in the programming languages used to write the code to control the computer. Each abstraction creates a *black box* that is defined, not by how it implements a task, but rather by how it interacts with the outside world. Software engineering's '*need to know*' principle (Somerville, 1992) states that every component is only authorised access to data that it directly needs to implement a particular function. The advantages to system design of the abstraction approach are that it simplifies the task: carefully chosen abstractions hide unnecessary details. Complex systems can therefore be built using a network of simple units, and any two units can be thought of as equivalent if they present the same interface. The abstraction approach facilitates straightforward reuse and replacement of each unit. A direct consequence of this approach, however, is that the internal mechanisms of each abstraction are not always available for inspection.

The success of the abstraction approach in system design has led to a similar technique being used in the design of the user interface. Abstraction is employed in user interface design as a means of reducing the complexity and increasing the consistency of the interaction. The user is presented with an interface that shows some aspects of the system's functionality but often hides the actual mechanisms employed behind an opaque barrier. The rigid and absolute nature of these abstraction barriers is what makes them so useful in system design. Unfortunately, the very properties that make the abstraction approach so appealing to system design can cause problems when applied to the user interface. The inability of the user interface to adapt to the user's needs is now recognised as having a fundamentally damaging effect on the communication between the user and the system. The abstractions used by the interface to simplify the interaction may actually cause the mystery and confusion that results in poor usability.

> *"Once upon a time, technology was mostly mechanical. Everything was built of levers, gears, cogs and wheels. Workers who operated tools could view many of the parts and could see the effects of their actions. People had some hope of understanding how large machinery and small gadgets worked, because the parts were visible. The operation of modern machines and the concepts behind their design are invisible and abstract. There may be nothing to see, nothing to guide understanding. Consequently, workers know less and less about the inner workings of the systems under their control, and they are at an immediate disadvantage when trouble erupts." (Norman, 1995, p 159)*

1.3.1 Metaphor

One of the major applications of abstraction in the design of user interfaces is in the use of metaphor. A metaphor persuades a user to transfer some prior knowledge of a particular situation or structure, into their present situation. The metaphorical comparison is made only with respect to certain characteristics of this relationship, and it is left up to the user to determine which characteristics are appropriate. The user must extract the salient issues from the features of one situation and then form them into a coherent model that can structure the new situation or problem. The application of a metaphor forces the user to draw equivalences between abstractions of the two situations.

Metaphors help the user learn and understand how to act in novel situations by building a bridge between old knowledge and new knowledge. The resilience of this bridge, however, is not guaranteed and the metaphor must fulfil several requirements in order to ensure that the user transfers knowledge across successfully. The metaphor must present a pertinent and recognisable relationship between the two situations: it must provide the abstraction at an appropriate level and, the old situation must be understood sufficiently to make the extraction of the appropriate characteristics apparent (Borgman, 1984). For a metaphor to work, an intuitive abstraction of the known situation must relate to an appropriate abstraction of the novel situation.

What a metaphor can never do is explain any real underlying mechanisms of the new situation. The metaphor is prone to collapse when an understanding of the underlying mechanisms that are hidden by the abstraction becomes important to the user. The interface abstraction, implicitly applied in the form of a metaphor or directly applied by the interface designer, encapsulates the system and, unlike the real world, the system is no longer open to exploration by the user. It is not simply a case of finding the right abstraction for every system but rather it is a fundamental requirement for human action to have the opportunity to probe and improvise.

> *"The real world is always available to be pushed and prodded to explore how it works, and human actors allow us to query their actions and motivations. In other words, we organize our actions not simply around abstractions of possible actions, but around the detail of the production of action and behaviour in particular circumstances." (Dourish and Button, 1998, p 415)*

1.4 Situated action

Lucy Suchman in her book: *Plans and Situated Actions* (Suchman, 1987) introduced the HCI community to the idea that human-computer interaction, like human-human interaction, is organised not only around plans, but also around actions that are improvised from the immediate circumstances in which users find themselves. The user of a machine does not blindly act following a predetermined system of rules that can be observed, collected and formalised, but rather users' actions are situated in the local circumstances of the moment.

> *"The initial observation is that interaction between people and machines requires essentially the same interpretive work that characterizes interaction between people, but with fundamentally different resources available to the participants. In particular, people make use of a rich array of linguistic, nonverbal, and inferential resources in finding the intelligibility of actions and events, in making their own actions sensible, and in managing the troubles in understanding that inevitably arise. Today's machines, in contrast, rely on a fixed array of sensory inputs, mapped to a predetermined set of internal states and responses. The result is an asymmetry that substantially limits the scope of interaction between people and machines." (Suchman, 1987, p 180-181)*

Suchman's concerns with user interface design revolved around this asymmetry between the desire of the human user to probe a system in order to reveal explanations for the system behaviour currently being experienced, and the actual information offered by the interface. The interaction between the system and the human user is limited by the inadequacies of the human-system communication due to the static and rigid nature of the user interface. The user interface designer must therefore learn from the nature of human-human interaction, which Suchman suggested was itself not only orchestrated through purposeful action and shared understanding but also involved actions situated in circumstance[1]. The theory, or

[1] One more issue that is directly relevant to this dissertation was raised by Suchman. When the system is in normal working order, the system interface is transparent to expert users. It is only when part of the system is temporarily broken or unavailable (or when the user is a novice) that the interface constricts the human users as they start the processes of inspection and practical problem solving. When system breakdown or disturbance occurs, users may start to employ rules or procedures in a goal-oriented manner to repair the system or eliminate the disturbance in order to once again resume normal practice. Suchman's important point was that the goal-oriented procedures that users employed in such circumstances were not *fundamental* but rather, were *"contingent on and derived*

rather perspective, on human-human interaction that grounded Suchman's belief was a branch of sociology called ethnomethodology.

1.4.1 Ethnomethodology

Ethnomethodology (Garfinkel, 1967) is an analytical perspective on the problems of social order. It sets out a policy for study that aims to uncover settings through the social action of a society, which are continually revealed in the activities and actions of the society's participants. Historically, ethnomethodology confronted the then traditional perspective of sociological thought by redefining the relationship between observed social action and the problem of social order. What was rejected by ethnomethodology was the assumption that the coordination of social activities was achieved through society's participants performing according to shared rules. Ethnomethodology insists that stable social order does not exist, and social action is not directly determined from it. Instead, social order is made to work in the actions and interactions of society's participants. Ethnomethodology focuses on the interaction between people to uncover how this social order phenomenon manifests itself.

Everyday social interaction involves not just engaging in rational social behaviour, but also being seen by other social actors to be so engaged. For ethnomethodoloigists, the procedures for producing social action and the procedures for observing and learning rational social action are the same; social action is *reflexively accountable* (Dourish and Button, 1998). The organisation and production of social action provides the method for other social actors to recognise it as social action (accountable), and this recognition is a direct consequence of the actual production of social action (reflexive). The acceptance of the reflexively accountable nature of social action means that ethnomethodology requires the undertaking of investigations, which consider specific actions in detail and within a particular circumstance.

---

from" the situated action. It is not the situated action that is defined by the goal-oriented procedures, but rather the procedures that develop a meaningful context when situated action becomes problematic.

1.4.2 Ethnomethodology in HCI research

In recent years, the HCI community has begun to turn to an ethnomethodological perspective on social order as part of the requirement capture process in system design (Randall, Hughes and Shapiro, 1994; Heath, Jirotka, Luff and Hindmarsh, 1993; Jirotka and Goguen, 1994). The belief that the requirements engineer, or system analyst, needs to uncover information embedded in the social interaction of the system's users and the organisation's managers, and the realisation that the interpretation of this interaction is situated in its social context, is based within ethnomethodology. The contrast between the elicitation of the social requirements of the system, being informal and situated, and the more traditional technical requirements, being formal and context insensitive has been a central issue of this research. Goguen distinguished these two aspects as "*the wet*" and "*the dry*" of requirements capture (Goguen, 1992) and has argued that requirements engineering is the reconciliation of the two: the social and the technical (Goguen, 1994). Research that has used an ethnomethodological analysis has generally focussed on the relationship between the technologies used in implementing a system and the accomplishment of some aspect of work. This research used investigations based on fieldwork to understand the complex interactions of an organisation and the situatedness of the practices that people employ.

In summary, HCI research has employed ethnomethodology to analyse work practice to uncover and study the interactions between people and to produce a set of requirements to improve, or guide, the design of technology for that particular purpose. This is not, however, the use that Suchman (1987) proposed for ethnomethodology. The insights gained from the situatedness of human-human interaction epitomised by ethnomethodology needs to be used to improve the design of the interface between the human and the system. The *'asymmetry'* that presently dominates human-computer interaction (and it is argued in this dissertation dominates the interaction between the user and the WWW) must be balanced by drawing from the insights of human-human interaction.

Some work that has addressed the problem of bringing ethnomethodology to the design of the user interface is that of Dourish and Button (1998). Their approach, termed *technomethodology*, is claimed by the authors to be a foundational relationship between ethnomethodology and system design. The approach intends to be the analytical framework within which ethnomethodology leads interface design and although the work is relatively new and undeveloped, it does capture a number of the issues and principles that informed the work presented in this dissertation.

> *"The goal of our work ... has been to seek a new position on the relationship between computer science and ethnomethodology in the design of interactive software. This position regards the relationship between our disciplines [computer science and ethnomethodology] as a foundational, analytic concern rather than simply a practical one, and so it emphasizes how it is that the ethnomethodological position on the problem of social order can inform, respecify, and reconceptualize foundational elements of system design." (Dourish and Button, 1998, p 428)*

To clarify the position and further stimulate the argument, the next section describes the perspective on interaction and the design of the user interface used in this dissertation.

1.5 Perspective on interaction

The background to the work presented in this dissertation is Suchman's assertion that the course of action employed by a user when interacting with a computer[2] is situated in the circumstances of the present interaction. The implications of this for the user interface are that the system must be able to reveal to the user the intent behind its actions and behaviour. This may, however, be wrongly interpreted as making the interface abstraction self-explanatory, so that the user can discover the use that the interface designer intended. The correct interpretation is that the interaction between the user and the system should not be constrained by the interface designer. It is the very *nature* of human action to improvise and probe, to prod and explore, to circumvent and break down any abstraction barrier in order to reveal the actions and behaviour of the system. The aim of the user interface designer must be to harmonise the asymmetry present within human-computer interaction, a course of action proposed by Suchman over ten years ago.

This dissertation presents a course of work that undertook to apply this perspective on interaction to the WWW. A series of detailed studies and experiments uncovered that the interaction between the user and the WWW possesses an undesirable asymmetry between the needs of the user to improvise and the behavioural explanations provided by the user interface. Using the results generated through these studies and experiments, this work proposes an enhanced user interface that starts to tackle the problems of this asymmetry by presenting to the user a more appropriate model of the WWW and the underlying Internet technology.

1.6 Thesis overview

The main elements of the thinking that form the background to this work have now been presented. The remaining chapters are organised as follows:

Chapter 2 (Background): this chapter introduces the fundamental distinction between the complex communication mechanisms of the *Internet* and the abstract information space of the *World-Wide Web*. The chapter explains that the WWW is a hypertext system consisting of two components; information nodes and links between nodes; and a defining principle; rapid information retrieval.

Chapter 3 (The Technology): using the distinction presented in Chapter 2, this chapter presents the standards, protocols and languages that are now universally accepted in order to implement the whole system. The chapter starts from the WWW and then delves deeper, to describe some of the lower level protocols of the Internet. As the investigation descends the abstraction, the model of the WWW is replaced by the realities of the communication mechanisms. Of particular interest is the implementation of the Web page and the hyperlink.

Chapter 4 (Interacting with the WWW): this chapter presents the argument that the boundary between the WWW and the Internet is not sufficiently explained to the WWW user through the interface, resulting in unsatisfactory usability. The WWW browser presents an interface designed to help the user navigate through the WWW, with the central task of the interaction being the activation of *hyperlinks*. Usability problems arise due to users experiencing the system behaviour directly related to the communication mechanisms of the Internet that lie under the abstraction barrier. The evidence used in this chapter is based on a review of the literature.

Chapter 5 (Mental Models): an overview of the literature concerning mental models is followed by the results of an experiment undertaken to show the characteristics of WWW

---

[2] And more specifically a network of computers.

users' understanding of the system. Using a classical psychology experiment as the basis for the investigation, the difference between the novice and expert user is highlighted using drawings of WWW user's browsing sessions.

Chapter 6 (The WWW User): in this chapter and the two that follow is presented the empirical work undertaken to reveal and justify the findings of this dissertation. The subtle nature of uncovering human-computer interaction required a range of experiments, with a range of experimental methods. The chapter presents the method and results of the first set of experiments aimed at generating an initial focus for the research; to reveal the salient issues of the interaction between the WWW and its users so that the subsequent experiments could become progressively more focused.

Chapter 7 (The Universal Resource Locator): in this chapter the URL, which was identified in the first set of experiments as an important and rich source of information concerning the system, is looked at in some depth. The set of experiments undertaken to investigate this process more fully, together with their findings are presented.

Chapter 8 (The Hyperlink Lens): the penultimate chapter presents the hyperlink lens, which applies some of the findings of the experiments undertaken in the course of this research, to improve the WWW user interface

Chapter 9 (Conclusions): in the final chapter the results of the research programme are summarised, the original contribution to knowledge is discussed, and areas for further research are identified.

Chapter 2

Background

The World-Wide Web is a large information space that uses Internet technology to transfer data from the server where it is stored, to a user's client where the information can be displayed. This chapter presents a simple distinction between the WWW and the underlying Internet. Drawing from research that estimated the size and connectivity of the network, this chapter starts by presenting the dimensions of the Internet. Following this is an introduction to the abstract model of the WWW and an investigation into what fundamental properties it has inherited from the large body of research concerning hypertext systems. Of particular interest is the *hyperlink* – a connection between information objects – and the hyperlink's relationship with the object it references and the data transfer process that it defines.

In this chapter it is argued that the world-wide distribution of the information means that the behaviour a user experiences is strongly influenced by an unpredictable communication with an associated temporal delay. Traditional hypertext literature has failed to appreciate this and other usability problems related to the decentralised, highly distributed nature of the information space. This chapter, therefore, lays the foundation for the argument that the WWW on its own does not provide users with a sufficient model to explain the behaviour they encounter. The behaviour that users experience is a result of the hypertext model of the WWW and the entwined communications governed by Internet technology. This statement reveals the dissertation's fundamental proposal: *the presentation of the relationship between the WWW and the underlying Internet technology through an enhanced user interface will improve the usability of the WWW.*

2.1 The Internet

The Internet was a technology designed specifically for the United States military, but as the potential benefits of the technology became apparent, the Internet grew to serve the needs of the academic and research communities. Subsequently, commercial interest drove the rapid expansion of the Internet across the globe (Leiner, 1994). The Internet is now a world-wide inter-network that is designed to connect different forms of communicating networks together. The network protocol used by the Internet to achieve this inter-networking is the *Internet Protocol* (IP), which provides a message passing service and a mechanism for addressing every computer connected to the network. Any two computers connected to the Internet, irrespective of the technology used by their local networks or any intermediate networks, can communicate using the IP.

The Internet Domain Survey (RFC 1296), which is carried out every six months, estimates how many computers are connected to the Internet. It uses a method based on the premise that any computer that wished to communicate using the IP must have an IP address so that messages can be sent to it. An IP address uniquely identifies a host and has traditionally been a 32-bit number, giving a theoretical maximum of 4.3 billion (2^{32}) IP addresses. IP addresses, however, are hierarchically allocated so the usable total is considerably less (the IP address is discussed in more detail in Chapter 3). The Internet Domain Survey distinguishes the IP addresses that are in use from the ones that are not, by counting only those that are advertised in the Domain Name System (also explained in Chapter 3). The premise is that an IP address is only in use if it has been given one of these *domain names*, which allow it to be used by applications like the WWW and email. The survey concluded that in January 2000 there were 72 million IP addresses advertised in the DNS. The growth of these Internet *hosts* over the past seven years is shown in Figure 2.1.

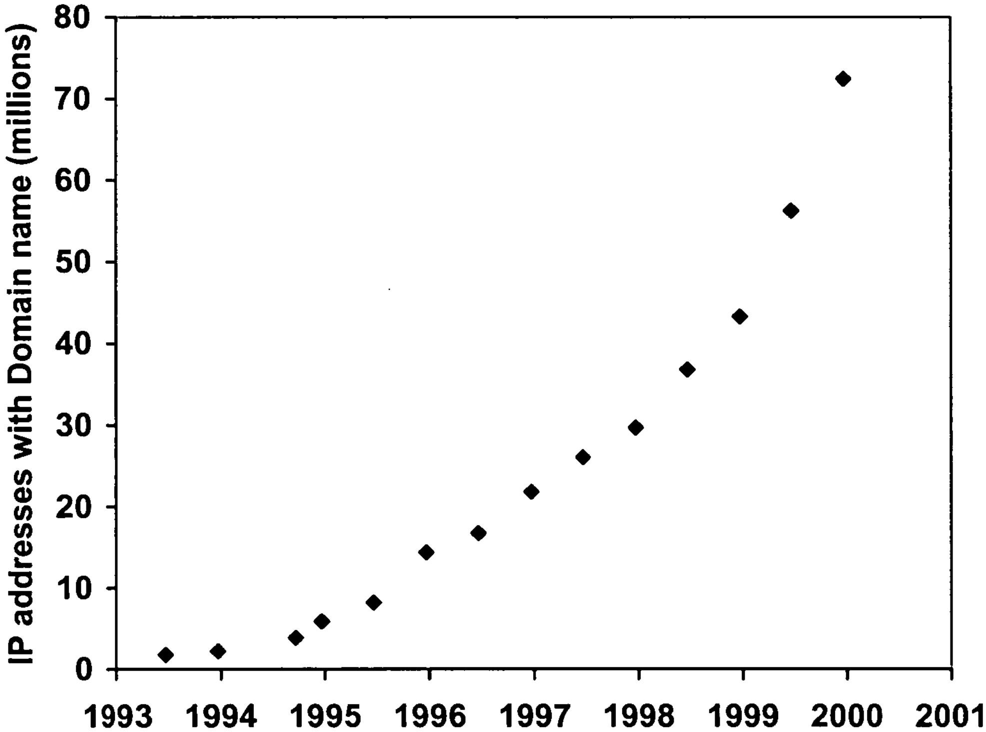


Figure 2.1: Increase in IP hosts (Source: Internet Software Consortium (http://www.isc.org/))

If two computers are connected to the Internet one can communicate with the other if the sending computer knows the IP address of the receiving computer. The IP address enables the system to construct a route through the network that can be used to transfer the message. The Internet is designed to be robust and if sections of the network fail, two hosts can still communicate by finding an alternative route for all or part of a message. This simple property means that the network must have a high connectivity, and it is this connectivity that is illustrated by the Cheswick/Burch map of the *"full Internet"* (Cheswick and Burch, 1999). The map (Figure 2.2) captures how local networks are connected together to form the Internet: the end of each line being a connection to a small local network. The map was constructed by probing and then tracing, the route that small communications[1] took to get to random Internet entities from a single host at the Bell Laboratories (Dodge, 1999). The picture is coloured using a strategy based on IP addresses.

[1] The communications were actually UDP packets sent to random high-numbered ports.

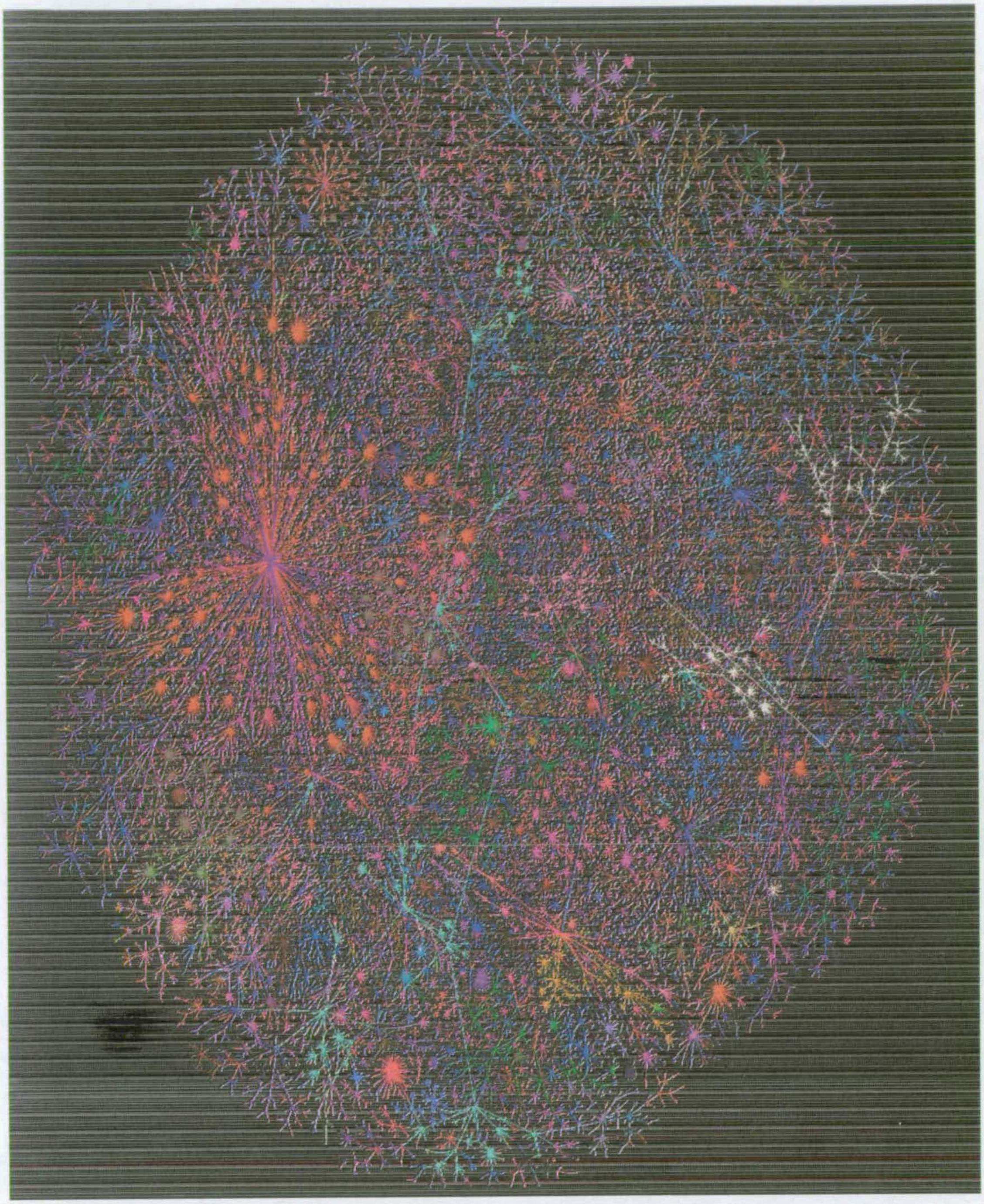


Figure 2.2: Cheswick/Burch Internet map as of 26 June 1999.

With nearly 90,000 nodes, 100,000 edges, and about 50,000 leaves, the tree-like structure that the figure presents is complex. Communicating a message from a computer in a network, at one of the leaves, to another, connected at another leaf, can involve routing through a number of branches. Each of these branches can communicate using different low-level protocols and yet the overall communication can be thought of as a single Internet communication due to the use of the IP.

There is a range of applications that use Internet technology to transfer data between computers. The first applications that used Internet technology were for transferring files and led to the development of electronic mail. Recently, a distributed information space has become one of the major applications to use Internet technology.

2.2 The World-Wide Web (WWW)

Tim Berners-Lee developed the WWW at the European Centre for Nuclear Research (CERN). The original aim was seamlessly to transfer information between independent projects at CERN. If similarities between projects could be uncovered, and there existed a relationship between information that could be advantageous to one or other of the projects, the information could be smoothly reshaped to represent the new state of the knowledge (Berners-Lee, Cailliau, Luotonen, Nielsen and Secret, 1994). The fundamental philosophy behind the original idea was that there existed a pool of knowledge that could be perused easily and updated by everyone. The fact that the initial Web was scalable and the protocols were simple to implement, were major factors in allowing the Web to grow by means of the Internet and expand world-wide.

Although Berners-Lee in his original design of the WWW hoped to realise an interactive means for collaboration and augmentation of ideas, the world-wide adoption and subsequent commercialisation of the design created de-facto standards that are now difficult to change. The uptake of these standards has led to the WWW becoming a new medium for published information: a single author has control over the content of a Web page and readers are, in general, unable to augment the Web page with their ideas (Wright, 1997). The WWW describes a model of boundless information where all items can be retrieved by use of a unique reference. What makes the WWW a powerful information resource is the facility to *link* information objects together by embedding an information object's unique reference (called a Universal Resource Locator and introduced in full in Chapter 3) within another information object. An individual or institution can provide any number of links to *any* other information object from its own published information objects.

Unlike the traditional means of publication the WWW allows people to publish work unregulated with respect to both its content and quality. This lack of regulation, combined with the ease with which information objects and links can be added or removed, together with the growth of the WWW has led to a complex, continually changing information space. The WWW can be thought of as a very large directed graph whose vertices are information objects and whose edges are links, pointing from the *referencing* information object to the *referenced* information object (Figure 2.3).

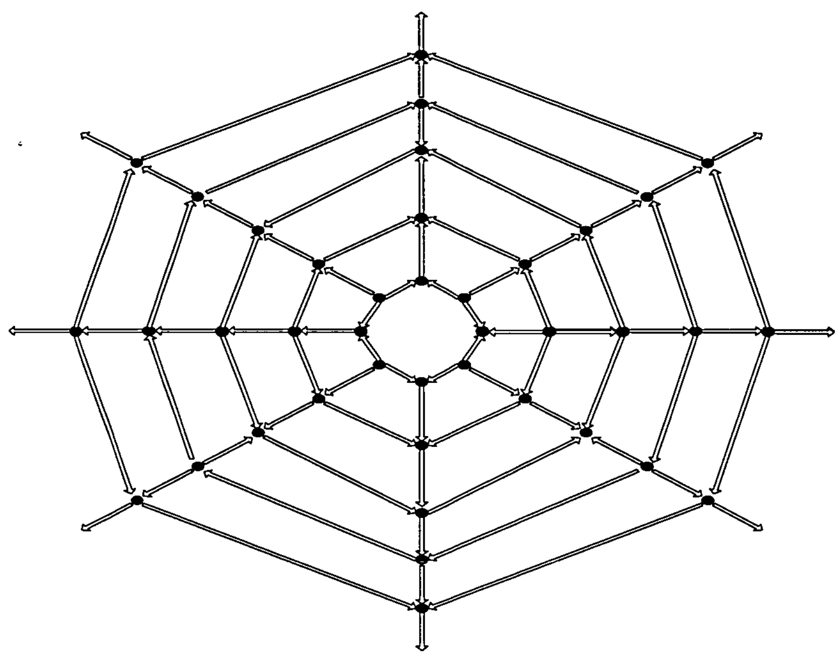


Figure 2.3: A topology graph of the WWW where the nodes are information objects and the directed edges are links.

In February 1999, it was estimated that the WWW consisted of at least 800 million publicly accessible pages. The information contained on these pages was calculated to be six terabytes (6×10^{12} bytes) of textual information[2] and three terabytes of images (Lawrence and Giles, 1999).[3] All this information is highly connected and one mathematical model of the WWW has concluded that any two of these 800 million WWW pages are, on average, only nineteen links away from each other (Giles, 1999). These, and many other studies to determine the characteristics of the WWW, are described in depth in Chapter 4. For now,

[2] This estimate is adjusted to mean the text that is available for display and does not contain any markup text. The total including markup was estimated at 15 terabytes.
[3] To put these figures in perspective, Ritchie suggested (Ritchie, 1999) that if somebody recorded all the speech that they uttered and heard in their lifetime, it would take up about one terabyte of storage.

however, this account is sufficient to justify the view that a visualisation of the WWW's topology, as the directed graph suggested earlier, would be extremely complex.

To further complicate this model of the WWW, the information objects and the links between them are **not** invariant in time. The author of an existing information node can change its content at any time. An author can add new links and remove existing links as Web pages are updated. New information nodes can be made accessible and operative information nodes can be eliminated. The WWW has the potential to change constantly with respect to both the content of, and the connections between, the information nodes it encapsulates. The WWW is not an information landscape where users can place a reliable signpost and learn to navigate through recognition, but rather a fluctuating sea of information where information objects can constantly move and routes through the information objects can change continually.

2.3 The WWW-Internet relationship

The Internet is a technology that allows communication between computers connected to networks that use different technologies. The WWW is a constantly changing, highly interconnected web of information pages distributed world-wide. The relationship between the two is that the Internet physically connects a WWW server to a WWW client (both located at one of the leaves of the Cheswick/Burch map). A WWW information object resides on a WWW server, and is referenced by a link embedded in another WWW object. This referencing object may reside on another WWW server. A WWW client, having obtained the referencing information object, can traverse the WWW link by starting a data communication process to transfer the referenced information from its server to the client. Chapter 3 will present an in depth look at how WWW information objects are stored and how they are transferred using Internet technology. This chapter, however, will concentrate on the abstract model of the WWW and how it relates to the discipline of hypertext.

2.4 Hypertext

Hypertext is an approach to managing information where data is stored as a network of nodes connected together by links. The essential feature of hypertext is that users must be able to navigate freely through the information in accordance with their own needs. This navigation is achieved by a sequence of explicit user actions, each of which *activates* the system to retrieve the information connected to by a particular link. Nelson coined the word "*hypertext*" in 1965 (Nelson, 1965) to imply a non-linear approach to presenting textual information, and the term "*hypermedia*" was subsequently introduced to stress the multimedia aspects of information spaces.[4] The original idea for the associated linking of ideas (and hence hypertext) can, however, be traced back a further twenty years.

At the end of the second World War, Vannevar Bush published a paper, under the title "As we may think" (Bush, 1945), in which he optimistically suggested that technology could help solve some of the problems of the post-war society, in particular the problems of storing, retrieving and communicating ideas. He proposed that his philanthropic vision could be achieved by the construction of a device that stored a vast collection of information items that could be consulted flexibly and quickly. This device[5], or system of devices, and the accompanying description of an information environment, has deeply influenced researchers in a number of HCI related disciplines.

Bush was the Director of the Office of Scientific Research and Development in the United States and had co-ordinated the activities of a large number of leading American scientists in the direct application of their scientific research to the war effort. His experience of co-ordinating such a large and diverse number of science specialists led him to highlight the problem of sharing copious amounts of scientific information, both that which was newly breaking and that which was already established, within and across disciplines. He felt that the systems for accessing records at the time were unsuitable owing to problems that stemmed from the rigid mode of traditional indexing;

> *"When data of any sort are placed in storage, they are filed alphabetically or numerically, and information is found (when it is) by tracing it down from subclass to subclass. It can be in only one place, unless duplicates are used; one has to have rules as to which path will locate it, and the rules are cumbersome." (Bush, 1945)*

It was Bush's contention that the indexing system was the main cause for the slow selection of items and the indexing concept restricted the tying together of related items. The need for

---

[4] We use hypertext and hypermedia interchangeably from now on.
[5] Bush called this device the memory extender or memex.

accuracy and speed when compiling results of scientific research was very real to Bush as he had experienced the need for both in the cauldron of a war effort. Time spent searching for previously published results and theories was time lost from the design of instruments of war. Bush's vision was of a device that would rapidly select and present information, and provide links to related information. What we now refer to as hypertext was conceived consisting of two major components:

- information *nodes,*
- associated *links* between these nodes,

and driven by a defining imperative:

- *rapid* information retrieval

2.4.1 Hyperlink

The tying together of associated pieces of information is the crucial mechanism of hypertext and the hypertext link, or more simply *hyperlink*, is the construct that facilitates this. Bush's frustration with indexing was justified by a belief that the human mind operates by association.

> *"With one item in its grasp, it snaps instantly to the next that is suggested by the association of thoughts, in accordance with some intricate web trails carried by the cells of the brain. It has other characteristics, of course; trails that are not frequently followed are prone to fade, items are not fully permanent, memory is transitory. Yet the speed of action, the intricacy of trails, the details of mental pictures, is awe-inspiring beyond all else in nature." (Bush, 1945)*

The hyperlink is designed to simulate the cognitive processes of the mind. The hyperlink connects associated ideas and when desired, rapidly retrieves these ideas.

2.4.2 Models of linking in classical hypertext

The literature concerning the history of hypertext is comprehensive and the numerous academic and commercial hypertext applications, each with its own twist on the theme, are well documented (see for example, Nielsen, 1995 and Cotton and Oliver, 1993). Rather than revisit this area, the following section looks at two more informal models of hypertext. There then follows the formal Dexter reference model that was introduced to encapsulate this research field by proposing a consensual terminology and semantics for hypertext.

2.4.2.1 The Nelson model

In 1965, Nelson defined the word hypertext thus,

> *"Let me introduce the word hypertext to mean a body of written or pictorial material interconnected in such a complex way that it could not conveniently be presented or represented on paper. It may contain summaries, or maps of its contents and their interrelations; it may contain annotations, additions and footnotes from scholars who have examined it. Let me suggest that such an object and system, properly designed and administered, could have great potential for education, increasing the student's range of choices, his sense of freedom, his motivation and his intellectual grasp." (Nelson, 1965)*

Nelson's view on augmentation of information was crucial to his vision and so he introduced a way of associating ideas that was slightly different from the original idea of the hyperlink presented by Bush. He called this method *transclusion* (Nelson, 1995). Nelson argued that a link is simply a connection between different things, whereas transclusions are connections maintained by the system that connect instances which are the same in different contexts. The exact same information object can exist in multiple places. Rather than copying and subsequently pasting to create an identical copy, transclusions act as pointers that connect the original copy to all the places that use it. The data that are transcluded are *alive,* since any change in the original is propagated to all transclusions. An information node in such a system is the author's original material plus a collection of transclusions from a variety of sources. It is, therefore, easy to edit the information, since users can transclude any parts of an information node and add their own comments. This model of hypertext is very different from the one presented previously in the explanation of the WWW. Associated links in the Nelson model are not simply pointers to other information nodes, they are alive inclusions which can be of any type or size of information; a whole encyclopaedia, a picture, a paragraph or simply a word may be a transclusion in another information node.

As well as allowing the hypertext reader to augment and view other authors' augmentations, Nelson also wanted **readers** to decide to what level of detail they wished to peruse a particular subject. This was in stark contrast to traditional books where the author had control.

> *"Systems of paper have grave limitations for either organizing or presenting ideas. A book is never perfectly suited to the reader; one reader is bored, another confused by the same pages. No system of paper – book or programmed text - can adapt very far to the interests or needs of a particular reader or student." (Nelson, 1965)*

Hypertext readers (or users) control the information they wish to expand on and what information they wish to summarise. This led Nelson to conclude that the system must have information about everything on-line together, since all information is intertwined. In 1981 Nelson proposed a system called Xanadu (Nelson, 1981), which was a repository for everything that anybody had ever written. Implementation of this system is still on going.

2.4.2.2 The Nielsen model

The Nielsen model of hypertext is synonymous with the initial presentation of the WWW. Nielsen defined hypertext to be a text that can be read non-sequentially (Nielsen, 1990b; 1995). The non-sequencing is achieved by interlinks between pieces of text (or other information). In Nielson's model the hyperlink connects two pieces of text or nodes and is normally directed, meaning that the link points *from* one (*anchor*) node to another (*destination*) node (Figure 2.4). A link is anchored to a departure point in the anchor node to provide users with some explicit object to activate when they wish to follow that link (the activating object). Nielsen defines these anchors as taking the form of an "embedded menu" where part of the primary text or a graphic double up as both information in itself and as a link anchor. The anchor must then be activated to follow the link to its destination node (Nielsen, 1990a).[6]

---

[6] Nielsen refined his model of a hyperlink by introducing two different types of link; implicit and explicit links (Nielsen, 1990a). An explicit link is defined in the hypertext whereas implicit links are a property of the information. For example, a hypertext system may use explicit links to connect ideas, but the system may also provide an automatic glossary that will retrieve a definition of every word, an implicit link.

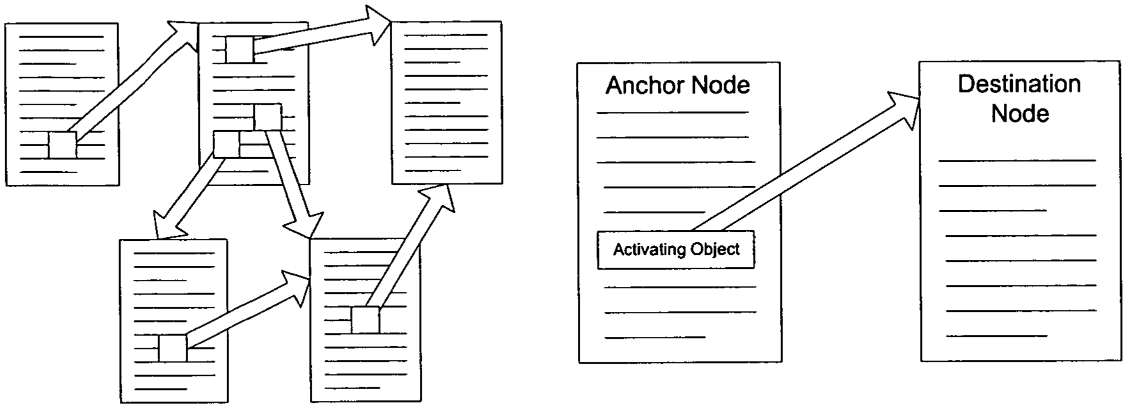


Figure 2.4: Nielsen's model of the hypertext and the hyperlink

Although Nielsen conceded that this was a very broad definition of hypertext, he stated in numerous sources that he preferred to define a system as hypertext by what it is like to use and whether a user can move freely through the information space.

> *"When asked whether I would view a certain system as hypertext, I would not rely so much on its specific features, command, or data structures, but more on its user interface "look and feel"." (Nielsen, 1995, p 5)*

> *" ...true hypertext should also make users feel that they can move freely through the information according to their own needs. This feeling is hard to define precisely but certainly implies short response times and low cognitive load when navigating." (Nielsen, 1990b, p 298)*

2.4.2.3 The hyperlink within the Dexter model

The Dexter hypertext model was developed in 1988 in an attempt to form a consensus among hypertext researchers, on the terminology and semantics of basic concepts (Halasz and Schwartz, 1994). The hope was to provide a reference model that could be used to compare different hypertext systems. The basic entities of the Dexter model are *components* that can be an atom of information, a link, or a composition of other components. A link is defined in the Dexter model as an entity that captures the relationship between other components.

Every component has a unique identity that provides a guaranteed mechanism for addressing any component in the hypertext. A link, however, does not rely on the explicit use of the unique identifier, as this is potentially restrictive, since documents change over time. When a link is followed, its specification is resolved by the system to a possible set of identifiers that can then be used to access the correct component or components.

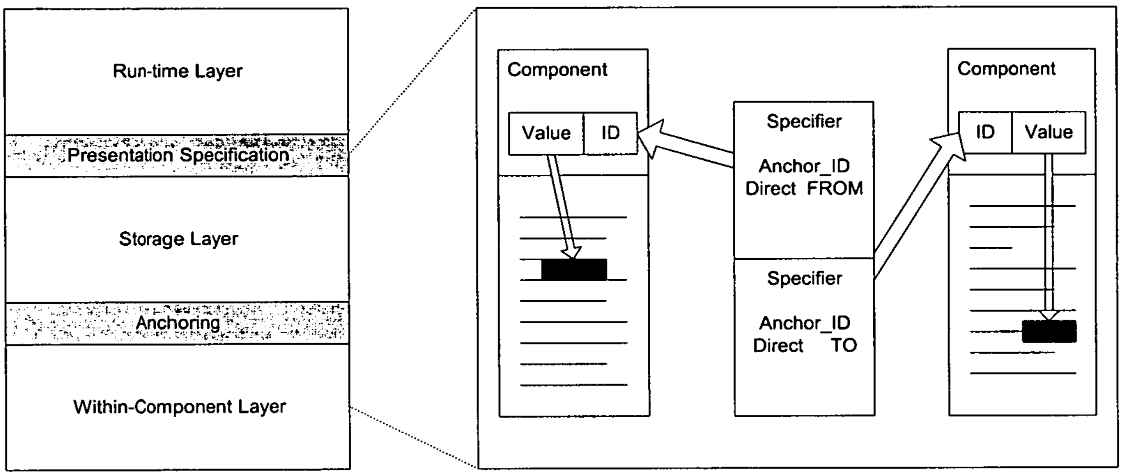


Figure 2.5: The hyperlink within the Dexter model

Using these definitions, a link component can connect a component to a number of other components. Hypertext, however, must provide the facility to link between sub-components, but this mechanism cannot rely on knowledge about the internal structure of an atomic information component in order that the Dexter model can separate the hypertext 'network' from the internals of the components. To address this restriction, an indirect addressing entity is used which is called the *anchor* that consists of two parts: an id and a value. The *anchor value* specifies a location or region within a component that can be interpreted only by the within-layer application. The *anchor id* simply identifies the anchor within a component as each component can, of course, contain a number of anchors. A link's source or destination can be identified uniquely using the surrounding component identifier and anchor pair. Any change to the internal structure of a component through editing only needs a change to the anchor value for the link to be preserved.

An anchor is needed as an end-point as well as a starting-point of a link. The Dexter model uses an entity called a *specifier* which specifies a component and a particular anchor within that component to describe what that anchor does. It can be the source of a link, the destination of a link, both a source and destination, or neither. A link component can now be defined as a sequence of specifiers (usually of length two). The link is now external to the informational components. The information components simply contain anchors that can be used as link sources, destinations or both. Finally, there is one important rule that enforces link-consistency.

- *The component specifiers of every link specifier must resolve to extant components.*

The destination of a link must exist and if an informational component is removed, all the links whose specifiers resolve to that component's identifier must also be deleted.

In summary, the Dexter model provides a very high level definition of a link that can be used to map on to different hypertext systems implementing links in a number of different ways. The Dexter model separates the link concept from the anchor entity. This separation allows two-way links which, in turn, mean that the original documents are unaffected by link creation or deletion, but with an overhead of an external link engine.[7]

2.5 Summary

The four models of hypertext presented in this chapter (Bush, Nelson, Nielsen and Dexter) only give a flavour of hypertext and its diverse origins. The WWW has taken something from each of these and other models and hence, is a form of hypertext. Where the WWW differs from these traditional hypertext models is in the physical distribution of the information with the communication overheads this entails, and the decentralised and unregulated nature of publication. These differences become immediately apparent while actually using the WWW, when the simple abstraction of a hypertext is rarely sustained.

---

[7] The work of Gronboek et al. (Gronboek, Bouvin and Sloth, 1997) attempted to augment the present implementation of the WWW with some Dexter based hypertext services. The basic problem this work attempted to overcome was based around changing the embedded addressing of the WWW with the external links of the Dexter model. Due to technical problems the goal of the work was not achieved, but an interesting investigation into how an external link engine could be implemented for the WWW was presented. Implementation of this system, however, would need a change to the standards and architecture of the WWW and, as has been mentioned before, the world-wide commercialisation of the WWW design makes these or any other changes very difficult to achieve.

The user interface currently presents the WWW as a set of information objects that are highly connected by hyperlinks; a wealth of information that a user can effortlessly peruse, seamlessly moving from one piece of related information to another. What this dissertation will show is that in reality, rather than experiencing an instant and deterministic response when activating a hyperlink, a user must usually cope with an unpredictable delay resulting from the data transfer processes of the underlying Internet technology. Factors such as the size of the information object, the physical location of the information store, and the route the data takes to reach the user, are all potential disturbances that the WWW user interface fails to hide but does not explain; they destroy the simple abstract hypertext model. Novice users are left unable to explain satisfactorily the behaviour they witness. Only by circumventing this abstraction and acquiring knowledge of the underlying Internet technology can users start to make sense of the system behaviours they experience.

As well as the potential for an unpredictable temporal delay after activating a hyperlink, the WWW user must also cope with the volatile nature of the hyperlinks themselves. The introduction and elimination of new information objects proceeds unregulated with respect to the new object's content and connectivity. The fact that it is technically simple to add information objects and hyperlinks to the WWW, allows authors to publish new information objects easily and link them to any existing objects. Similarly, it is easy for information objects to be removed and for hyperlinks to be left pointing to non-existent information. Users are left to cope with an information space that changes continually around them and leaving them a world that is always unfamiliar.[8]

A further factor that the WWW abstraction conceals from the user is the social nature of both the information space and the underlying communicating network. Other users are concurrently reading the information provided by the WWW and transferring data using the Internet. These two distinct forms of social use could affect the usability of the WWW if revealed to the user. First, the popularity of Web sites and Web pages could provide an indicator of relevance especially if the population that was used in the calculation could, in someway, be controlled (for example, if a large number of doctors use a particular Web page this may be an indication of both the relevance and the quality of the information contained on the Web page for other doctors). Second, many different users from around the world use the Internet at the same time, and this contributes to the variable behaviour of the data transfer process. The present WWW abstraction does not provide any mechanisms for disclosing either, a form of social affordance to help the user decide on the relevance or quality of the information at the end of a hyperlink, or any information about the traffic on the network.

The solution to these problems that this dissertation proposes centres on understanding the distinction between the information space that is the WWW and the Internet technology that it uses when transferring data. This distinction emphasised at the beginning of the chapter, is fundamental to the arguments presented in the rest of the dissertation and is a guiding principle for the final implementation. Current WWW user interfaces present a blurred picture of the boundary between the WWW and the Internet, hiding some aspects of the implementation and reluctantly revealing others. It is this opaqueness of the user interface which leads the WWW user to misinterpret the system behaviour they witness and denies them the opportunity to act appropriately.

In Chapter 3 the technology behind the WWW and the Internet technology it uses is explained. The implementation of the information nodes, together with the hyperlinks that connect them and the Internet technology that transfers the data are also presented. The chapter shows that it is design issues taken at this level that are central to the creation of the abstraction barriers at the user interface.

---

[8] One solution to this problem of a changing information landscape is addressed by the Web portal (for example Yahoo, AltaVista, LineOne etc.), which provides a list of hyperlinks that reference information that has been reviewed, guaranteeing existence and providing some quality assurance.

Chapter 3

The Technology

The argument presented in Chapter 2 centred on the distinction between the abstract hypertext system of the WWW and the underlying communicating technology of the Internet. How the current WWW browsers blur this distinction and how this leads to users misinterpreting the behaviour they experience are areas of the argument covered in later chapters. Chapter 3, however, investigates the standards, protocols and languages that have been universally accepted and whose use has facilitated the rapid growth of the WWW.

The chapter starts by looking at the hypertext features of the WWW discussed in the previous chapter and shows how the system specifies them. This specification is a contributing factor to how a WWW information object is transferred from its store on a server to the user's client where it can be displayed. This data transfer process and the Internet technology it uses are considered in depth and an illustration of the complexities of the communication's temporal behaviour is presented. What the chapter reveals is that although the WWW is a collection of information nodes (called Web pages) the technology dissects each of these nodes into a number of parts. Each of these parts must be requested and transferred from various Internet hosts to the client before the node can be constructed.

In fact, the Web page does not exist as a single data object at all; it is a conceptual entity that is only realised after the user agent, using a number of different data files, has rendered it.

3.1 The hypertext of the World-Wide Web

Like all hypertexts, the WWW contains two important components: information nodes and links between the nodes. Where the WWW differs from the traditional models of hypertext is in the distributed nature of the data and the many-to-many relationship between hypertext authors and users.

The WWW is a multimedia environment so an information node can be of almost any type: an image, a video, a sound recording etc. The fact that an object is defined as a WWW information node is related to how it is addressed, requested and subsequently transferred from the WWW server it resides on to the WWW client where it is presented. The WWW does not restrict the format of an information node, but it does provide a universal language for constructing one very useful type of WWW information node. This universal language is the Hypertext Markup Language (HTML) and the information nodes that this language is able to define, are WWW hypertext pages or more simply, *Web pages*.

A WWW client receives the HTML document that defines a Web page, and renders the page for the user. Normally this is done by visually displaying the multimedia information in a window on the user's desktop machine, although the *user agent* may present the information in any format. Applications that display Web pages are often referred to as WWW browsers due to the nature of the activity that a user is perceived to perform: browsing through information.

3.2 The HTML document

A Web page is defined by an HTML document that enables an author to provide a WWW resource as a single displayable multimedia information object. Text can be defined to have a variety of sizes, fonts and colours and can be positioned at various points on the page relative to other objects. A Web page can also contain a number of embedded objects like an image or a Java applet and, in a similar way to text, these objects can be relationally positioned on the page. What is important to note, however, is that the data for an embedded object is not contained in the HTML document, but referenced from this file (Figure 3.1).

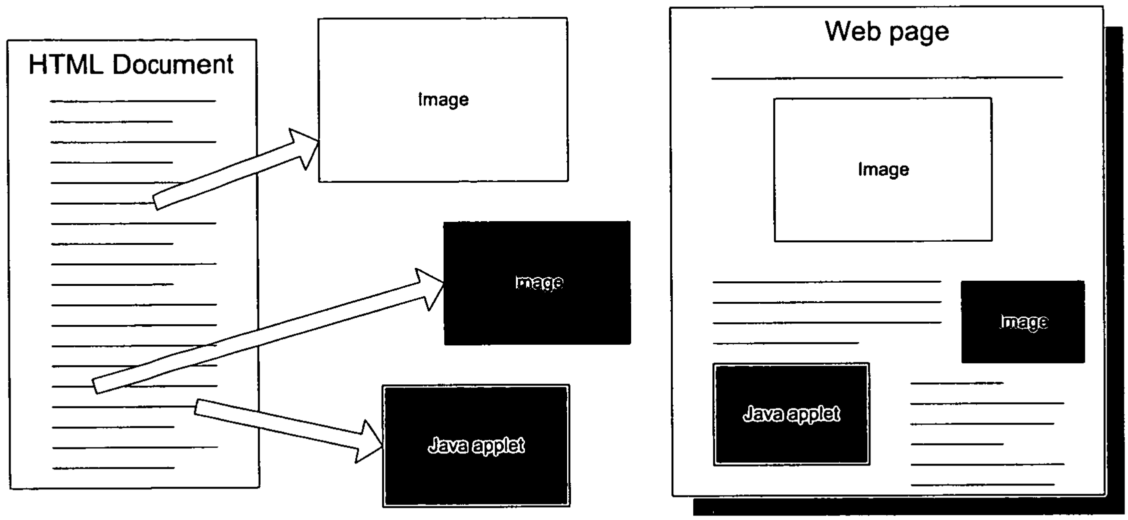


Figure 3.1: The HTML document and the Web page.

An HTML document contains an unrestricted number of lines of text, expressed in the HTML.[1] To conform to the standard, an HTML document must contain three parts:

- A line containing HTML version information
- A declarative header
- A body that contains the Web page's textual content and references to embedded objects

A valid HTML document must declare what version of the HTML it is using and so the HTML version information line states the document type definition (DTD) in use. Even for HTML 4.0 there are three different DTDs (strict, transitional and frameset) and it is important that an author states which one has been used. In Figure 3.2, line 1 specifies that this particular HTML document uses the HTML 4.0 strict DTD. The information contained in the HTML header (lines 3 to 5 in Figure 3.2) is not for rendering on the Web page but may be made available to a user through some mechanism if desired. It can contain information about the Web page, like its title, keywords for searching or other meta data not considered renderable document content.

[1] The present version and the one assumed by this work is HTML 4.0

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0//EN">
<HTML>
        <HEAD>
                <TITLE>Title of HTML document</TITLE>
        </HEAD>

        </BODY>
                <H1 align="center">Redefining the Hyperlink</H1>
        <BODY>
</HTML>
```

Figure 3.2: A Simple HTML document with its three components (the line numbers are included for reference only)

Finally, there is the body of the document that contains the contents for presenting to the user by the Web browser. The user agent controls the way that this content is presented, so the same HTML document may be visually presented as text, text and images, or alternatively, for audio user agents, it may be spoken. The fact that the presentation of the Web page is dependent on the user agent and not necessarily the HTML document author has always been an important feature of the WWW: separation of the presentation from the content. The HTML has always tried to separate presentation from content by using element types (discussed next) but a more powerful facility has been introduced by HTML 4.0 called style sheets, which specify how different types of content should be presented. Style sheets are defined either in the HTML header or as a separate remote file.

3.2.1 HTML element types

As the name suggests, the HTML is a markup language. Traditionally this means that structural, presentational and semantic information are included alongside the content data itself. The desired structural behaviour of the Web page is included in the HTML file as *element types.*

Start Tag	Content	End Tag
`<H1 align="center">`	`Redefining the Hyperlink`	`</H1>`

Figure 3.3: Heading element type defined using the H1 tag.

Element types usually take the form of a start tag, content and an end tag. (Although some element types allow the omission of end tags, some the omission of start tags and some do not need any content.) Start tags are open delimited by the “<” character and end delimited by the “>” character and contain the element name followed by a list of optional attributes separated by spaces. End tags are similarly delimited, but contain a “/” character before the tag name. Figure 3.3 shows the heading element type with tag name H1, attribute align and some textual content, ”`Redefining the Hyperlink`”. The HTML tags simply indicate the type (with some presentation information) of the text that they enclose, so in this example the text enclosed by the H1 tag must be considered by the user agent to be a centrally aligned heading. One common area for confusion is that tags are often interpreted as elements. This is not the case. The tag is just the syntax whereas the element is part of the structure of the document. An illustration of this is in the HTML 4.0 specification (W3C, 1997), which states that:

> *“Elements are not tags. Some people refer to elements as tags (e.g., “the P tag”). Remember that the element is one thing, and the tag (be it start or end tag) is another. For instance, the HEAD element is always present, even though both the start and end HEAD tags may be missing in the markup.” (W3C, 1997, p 25)*

3.2.2 Embedded objects

The HTML document does not contain the data of any embedded multimedia objects. The HTML document simply includes an element type to specify that an object has been included and a reference to the data for that object. So, for example, to embed an image into a Web page the HTML document must include an image element that is specified using the IMG tag. As one would expect, the image element has no content (since it is itself a reference to some data) and hence no end tag. It must contain the `src` attribute, which gives a reference for the location of the data of the image to be embedded and an `alt` attribute to provide an alternative text description for the image (Figure 3.4).

```
<IMG src="http://www.waporacle.com/images/picture.jpg"
        alt="waporacle title" >
```

Figure 3.4: The image HTML element.

When the HTML document is displayed as a Web page, the image element is usually replaced 'inline' by the actual image, except if further attributes are used to specify that the image should 'float' out of line (for example, the `align` attribute). If the user agent is unable to render the image, the `alt` attribute provides a textual description that can be used as a replacement. Further attributes can also be added to the element start tag to override the dimensions of the image and provide a reference to a longer description than the `alt` attribute provides.

As well as this specific element type there is also a generic embedded object element type that uses the OBJECT tag name. This element type has been introduced in HTML 4.0 to allow for an extendable media inclusion mechanism, so that as new media are introduced they can easily be added to Web pages. The OBJECT element allows the author of a HTML document to specify what the user agent needs for the presentation of the embedded object: source code, initial values, and run-time data. This means that the author of the HTML document has control over the mechanism used to render the object, be it the user agent's built in mechanisms or an external application. Figure 3.5 shows the same embedded image as Figure 3.4, but this time it is declared as a generic object. What is worth noting is that, unlike the IMG element, the OBJECT element has a type definition in the HTML document itself and so the user agent knows what type of image (in this example a jpeg encoded image) is to be rendered at this location prior to actually getting the referenced image data.

```
<OBJECT data="http://www.waporacle.com/images/picture.jpg"
        type="image/jpg">
waporacle title
</OBJECT>
```

Figure 3.5: The image HTML element

To summarise, the WWW hypertext node is rendered as a Web page by a user agent and is defined by an author using an HTML document. The HTML document separates the content and presentation of a Web page by allowing document content to be given an element type. Multimedia objects can be introduced into a Web page by referencing the location of the object data within the generic OBJECT element. A direct consequence of this is that a user agent may have a HTML document but be unable to render it immediately as a Web page since the user agent may have to retrieve all the embedded objects from wherever they are located.

3.2.3 WWW hyperlinks

As well as the information nodes of traditional hypertext models, the WWW also has the facility to allow linking between the information nodes. The WWW employs explicitly embedded unidirectional links. There is no need for an external link engine and so the overheads of link management for the system are non-existent. This simplicity comes at the cost of constraining link creation to document authors and link targets are restricted to the start of a Web page or a target element, provided by an author, in a HTML document.

To construct a simple hyperlink an author must add a link element to the HTML document. There are two types of link elements provided by HTML 4.0: the LINK element which specifies a relationship between information nodes and must appear in the head of a HTML document, and the A element which specifies a hyperlink in the body of a HTML document.

The LINK element is not what has previously been defined as a hyperlink: instead, it defines a relationship between the HTML document and another information node. A LINK element has no content and as it is in the head of the HTML document, the user agent should not render it (Figure 3.6). The relationship that it specifies may, however, be made available to the user by the user agent in some form.

```
<HEAD>
   <TITLE>Wap Fortune</TITLE>
   <LINK rel="Index" href="http://www.waporacle.com/index.htm">
   <LINK rel="Next" href="http://www.waporacle.com/wapweather.htm">
</HEAD>
```

Figure 3.6: The HTML LINK element

Although the LINK element provides a useful relationship between information objects, it does not implement the hyperlink. The defining of a hyperlink is achieved using the anchor element. The anchor element, or A element, has an attribute `href` that, if specified, provides a reference to a destination node. The body of the anchor element can then be used as an activating object and is often highlighted by user agents (by underlining text or providing a border for an image). The destination node of a hyperlink can be any WWW information node of any media type (Figure 3.7(1) shows a hyperlink referencing a jpeg encoded image).

```
1. <A href="http://www.waporacle.com/images/picture.jpg">picture</A>
2. <A name="introduction">Introduction</A>
```

3. <A href="http://www.waporacle.com/wapfortune.htm#introduction">

Figure 3.7: 1. The HTML A element 2. Destination anchor 3. Link to a destination anchor.

Using the anchor element in this way, a hypertext can be created rather like Nielsen's model (see Figure 2.4). A hyperlink can be created from an anchor node (the HTML document the anchor element appears in) to a destination node (the object referenced by the `href` attribute) and it can be activated by selecting the activating object (the content of the anchor element).

Only being able to provide a hyperlink to another WWW resource is restricting so the anchor element can also be used to provide a destination anchor within a Web page. Once a destination anchor has been specified within a HTML document, links can be made into a Web page or from within, to within a Web page. A destination anchor is specified by the anchor element using of the `name` attribute (Figure 3.7(2)). Once a destination anchor has been specified a hyperlink can be created to link to it by adding the anchor name to the end of the document reference separated by the "#" character (Figure 3.7(3)).

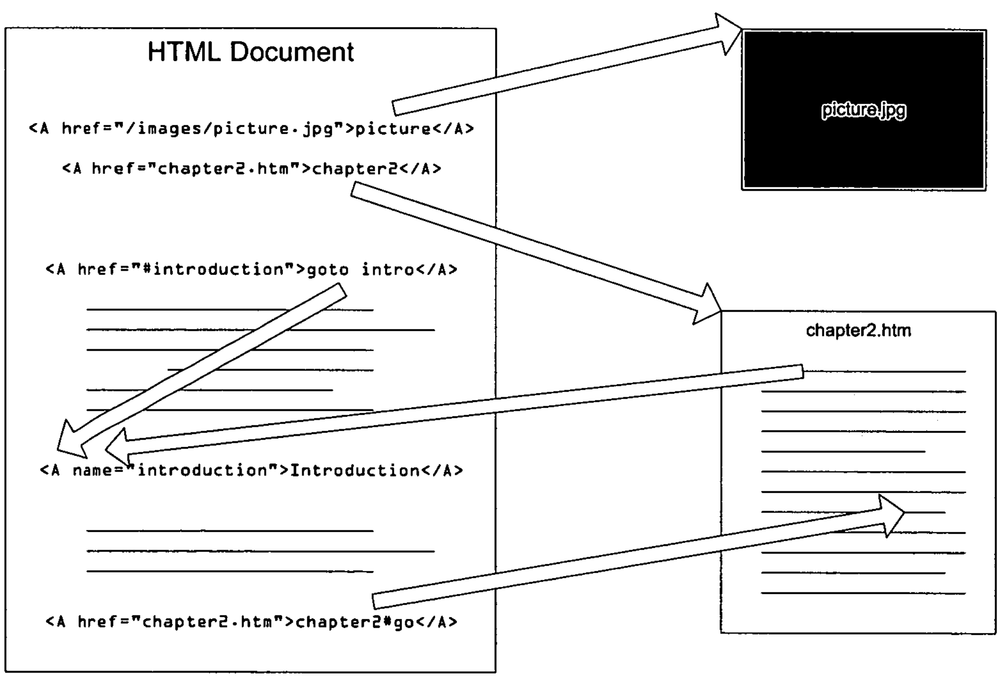


Figure 3.8: The different types of hyperlink.

The various uses of the anchor element provide the WWW with a number of different types of link, which are illustrated in Figure 3.8. Any information object can be linked to if the referencing document is provided with its unique reference. Linking to an internal location of an information object can only be achieved if the referenced object is defined by a HTML document that has provided an internal destination anchor. Linking to the internal anchor can then be accomplished if the referencing document has the referenced object's unique reference *and* the name of the internal anchor. Finally, a link between parts of the same Web page can be provided by the introduction and use of an internal destination anchor.

3.3 Communication underlying the WWW

All publicly accessible HTML documents define the information space of the WWW; the structure and presentation of each hypertext object; and the connections between each hypertext object. The HTML defines most of the information nodes and all the links between them. The final element is the rapid retrieval of an information node when a referencing hyperlink is activated. The following sections of this chapter look at the communication system that supports the request and delivering of information nodes.

3.3.1 The Universal Resource Locator[2]

The WWW is based on the idea of a boundless information space where every information object has a unique reference by which it can be retrieved. The WWW hyperlink uses the unique reference to link Web pages together, and the HTML document contains these unique references to locate the data for an embedded object. The uniform naming scheme that was developed for locating resources on the Internet is the Universal Resource Identifier (URI).

---

[2] This work uses the acronym URL to stand for the Universal Resource Locator even though the WWW consortium officially states that it stands for Uniform Resource Locator. Berners-Lee originally proposed the name Universal Document Identifier (UDI) but there was a strong reaction against what was said to be the arrogance of calling something universal. The term was therefore changed to Uniform Document Identifier and later Document was replaced with Resource and Identifier with Locator.

> *" I tried to explain at the session how important it was that the Web be seen as universal, but there was only so much time, and I decided not to waste my breath. I thought, What's in a name?" (Berners-Lee, 1999, p67)*

This apparent arrogance now seems justified and the far more appropriate term is used throughout this dissertation.

Identification of a resource can be associated with any characteristic of the resource such as its name or content. The members of the subset of URIs where identification is achieved through the resource's location are called Universal Resource Locators (URL) and it is the URL that is used for the location and retrieval of WWW resources.

A URL can be considered as a formatted string that abstractly identifies the location (*i.e.*, a computer and a position within it) of a resource. A full URL contains three main components: the scheme, host, and path.

- **Scheme** - specifies the protocol to be used to access the resource (usually the HTTP)
- **Host** - specifies the computer in which the resource is supported (usually in the form of its Internet Domain Name System (DNS) name)
- **Path** - specifies the location of the resource relative to the host using a Unix based file naming scheme

As well as these three major constituents, a URL may also present the server (specified by the host) with an operation to perform. This 'query' is concatenated to the end of the URL string and separated by the "?" character.[3] There is also an option of adding a port number (separated by the ":" character) after the host that allows the client to specify which process it wishes to communicate with on the server. If the port number is omitted it is assumed to be 80 as this is the standard secure port for HTTP and, hence, a default HTTP daemon will be listening to it and subsequently allow a connection to it. Figure 3.9 shows the BNF definition of the HTTP URL and an example URL broken up into its constituent parts.

```
http_url = "http:" "//" host [ ":" port] [abs_path]
where
        host = <A legal Internet host name or IP address>
        port = *DIGIT (80 is assumed if not given)
        abs_path = "/" [ path ] [ ";" params ] [ "?" query ]
```

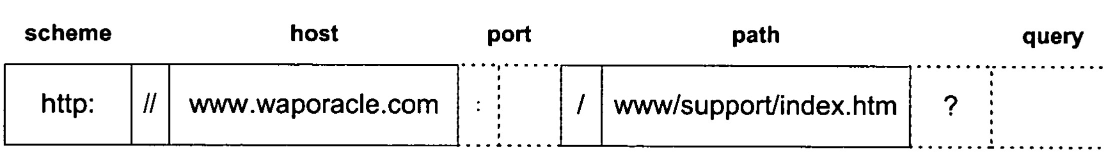


Figure 3.9: The HTTP URL.

[3] There is also the option to add a list of parameters separated by a ";" character which may be used by a process resource but this is fairly unusual and so will not be included in this or subsequent discussions concerning the URL.

This dissertation will return to the URL in Chapter 7 and look more closely at the role it plays in a user's interaction with the WWW. For now, however, the investigation of the URL is confined to its technical definition. One final important point is that the path and query are both case sensitive but the scheme and host are not.

3.3.2 The Hypertext Transfer Protocol

The central operation of the Hypertext Transfer Protocol (HTTP) is a handshake between client and WWW server. For the successful data transfer of a WWW resource the client sends a textual request that contains the URL of the required resource and the server reacts by sending a response that contains the resource. If the referenced resource is an HTML object this may contain embedded resources, such as an image, in which case there must be further HTTP requests for the resources and subsequent responses containing them. These subsequent HTTP handshakes can take place in any order and at any time (assuming that the embedded object's URL has been received). Due to the nature of these multiple handshakes, HTTP version 1.1 was introduced to allow persistent connections between client and server at the TCP level of communication (see Section 3.3.3 for a description of the TCP level of communication) and so improved performance of the WWW resource data transfer. The adoption of HTTP 1.1 is, however, far from total, so part of the HTTP handshake must confirm which version is in use.

Further modifications to the HTTP protocol are also under consideration and have been termed Next Generation HTTP (HTTP-NG) (Thomas, 1996). The basic improvement is that multiple HTTP communications can be multiplexed on to a single TCP channel (see Section 3.3.3) giving rise to a perceived concurrency in the HTTP-NG handshakes. The strong relationship between the HTTP and the WWW has also been a contributing factor to attempts to improve the performance of the HTTP-NG communication. This has resulted in the facility for a HTTP-NG response to be sent prior to a request, to allow the server to predict what the client will need next and reduce the total download time of a Web page. As stated in the previous paragraph, the uptake of HTTP 1.1 is presently modest and so whether HTTP-NG will have any impact is, as yet, unknown.

3.3.2.1 HTTP request

A HTTP request consists of one or more lines of text separated from an optional body by a carriage-return and line feed (CRLF). The only line that is *required* is the request line. The HTTP request line contains three elements: a method token that indicates what should be performed on the identified resource, the URL to identify the resource and the version of the HTTP that is being used (either HTTP 1.0 or 1.1). A HTTP server only has to support two methods GET and HEAD though there are a number of other methods that are defined (RFC 2068) which can optionally be supported. The GET method tells the server to retrieve whatever information is identified by the URL or, if the URL refers to a data-producing process, the data produced by that process. The HEAD method is implemented identically to the GET method but the server **must not** return any of the message-body in its response. The server simply responds with the HTTP header that, as is illustrated shortly, contains meta-information about the resource.

```
GET http://www.waporacle.com/www/support/index.htm HTTP/1.0
Date: Tue, 23 May 2000 08:12:31 GMT
User-Agent: Mozilla/4.0 (compatible; MSIE 4.0; Windows NT)
If-Modified-Since : Friday, 26 Dec 1999 14:27:43 GMT
```

Figure 3.10: HTTP request

After the request line, the HTTP request may contain a number of optional header lines that are categorised into three groups: general header, request header and entity header. General header lines contain information that relates to both the request and the response, but does not apply to the actual entity being transferred; for example, the date when the request was originated. Request header lines specify additional information which the server may find useful about the request and the client doing the requesting. Figure 3.10 shows two request header lines, the first being the `User-Agent` field which specifies that the request came from a Microsoft Internet Explorer version 4.0 browser running on Windows NT. After a HTTP request has been received by a WWW server (and if the request contained the User-Agent request line), the server will be in a position to extract this information about the client that could affect what information it sends back. The second request line shown in Figure 3.10, is the `If-Modified-Since` field, which tells a server to transfer the requested resource only if the resource has been modified since the date specified. Finally, a HTTP request can contain some entity header lines that specify information about the body of the request. The HTTP request does not contain a body when simply requesting a WWW resource using the GET and HEAD methods and so does not contain any entity header lines.[4]

3.3.2.2 HTTP response

The response to an HTTP GET request contains a textual *header*, a blank line (CRLF), and then the data of the referenced resource. The header consists of a status report followed by some response information that is constructed using a set of keywords. The status report contains three elements, the version of the HTTP in use, a numerical status code and a textual definition for human usage and not defined in the standard. The numerical status codes are all 3-digit integers and take the following form:

- **1xx** : Informational –request received, continuing process
- **2xx** : Success – The action was successfully received, understood and accepted
- **3xx** : Redirection – further action is needed
- **4xx** : Client Error – Bad syntax in request or server cannot fulfil request
- **5xx** : Server Error – The server has failed to fulfil an apparently valid request

So, for example, during a WWW communication a HTTP client will frequently receive some of the following status codes: 404 Not Found, 302 Moved Temporarily, 500 Internal Server Error, 204 No Content.

Figure 3.11 shows the HTTP response header to the request shown in Figure 3.10. The status report (the first line) shows that the server is using HTTP version 1.0 and the request was successful. The rest of the header contains a list of optional response information (extending the minimal information contained in the status code), information about the requested resource and general information in a similar way to the HTTP request. So, for example, the `Date` is a general information header line, the `Server` is a response information header line and `Content-Length` is a requested resource or *entity* information header line.

```
HTTP/1.0 200 OK
Server: Microsoft-IIS/4.0
Date: Tue, 23 May 2000 08:12:42 GMT
```

[4] Usually, the only time a HTTP request contains any body is when it is using the optionally supported POST method to post information to a server.

```
Content-Type: text/html
Accept-Ranges: bytes
Last-Modified: Tue, 30 Dec 1999 19:26:08 GMT
ETag: "2ff987967fbf1:467c7"
Content-Length: 3894
Age: 1216398
X-Cache: HIT from dye.dcs.ed.ac.uk
Proxy-Connection: close
```

Figure 3.11: HTTP response header.

Any HTTP version 1.1 or 1.0 response that contains a requested resource *should* contain a `Content-Type` header field to define the media type of the resource. This header item uses the multipurpose Internet mail extension (MIME) range of types to specify the type of the resource (RFC 2045-49). Although the MIME's `Content-Type` header field and media type mechanism has been designed to be extensible, there are definitions for seven top-level media types (RFC 2046). Five discrete media types are text, image, audio, video and application, and two composite types, multipart and message. This general type declares to a user-agent what the data is and a subtype specifies the format that the data takes. In this example (Figure 3.11) we have the MIME type text/html, which means that the resource is line-based text in the HTML format. Other possibilities are listed below (it should noted that the bottom two refer to MIME types that have yet to be accepted and hence start with an x-):

- **image/gif** – gif encoded image
- **image/jpeg** – jpeg encoded image
- **text/plain** – line-structured text with no implied syntax
- **application/octet-stream** – sequence of raw bytes
- **text/x-wap.wml** – line-structured text conforming to the Wireless Mark-up Language syntax
- **application/x-wap.wmlc** – complied WML

The important point to make here is that the client only knows the type of the requested resource when a successful request has been made and the HTTP header has been transferred from the server. Prior to this, the client can only know the object's type if it was embedded using the HTML OBJECT element and the client is willing to trust the `type` attribute.

3.3.2.3 The WWW cache

The HTTP communication is a simple handshake between the WWW client and the WWW server. The client and the server can be connected to the Internet at any point. They may be next to each other on a local network or far apart both logically with respect to the network and physically with respect to geography. A simple method to reduce the network traffic and the time a client has to wait for a request to be fulfilled is to store the WWW resources at a point *nearer* to the client. The intermediate place where these resources are hidden is called a *cache*. A WWW cache stores WWW resources that it has received previously and, if these resources are subsequently requested, the cached version of the resource can be supplied to the client from the cache rather than waiting for the resource to be requested and sent from the remote server. The cache must be situated between the client and the server and intercept the HTTP handshake. If the cache has the requested resource it can supply the client without communicating with the server and, if it does not have the resource, it can forward the HTTP request to the server and store the response, before forwarding it on to the client. A caching system can be, and normally is, implemented hierarchically, but the two main places where WWW caches are situated are on the WWW client called the *browser cache* and on an intermediate proxy called the *proxy cache.*

3.3.2.4 Summary of the HTTP communication

When a user activates a hyperlink situated in a Web page, the reference that is defined by the HTML document in the anchor element is used as an address to which an HTTP request can be sent. If the requested resource is not available or up-to-date in an intermediate cache, the request must be sent to the host computer where the resource is located. If this host is available for communication and the resource exists, the WWW server may respond with an HTTP response that may contain the actual resource. If the resource is a Web page, the HTTP response will contain a defining HTML document. All the embedded objects defined in the HTML document must then be requested from the WWW server where they reside. Once all the embedded objects have been successfully retrieved, the Web page can be properly presented to the user. Figure 3.12 shows the timing diagram for the request of an imaginary Web page containing two embedded objects that reside on different servers. (Note, the actual proportions of time spent on different stages of the process shown in Figure 3.12 are not representative of an actual communication but for illustration only.)

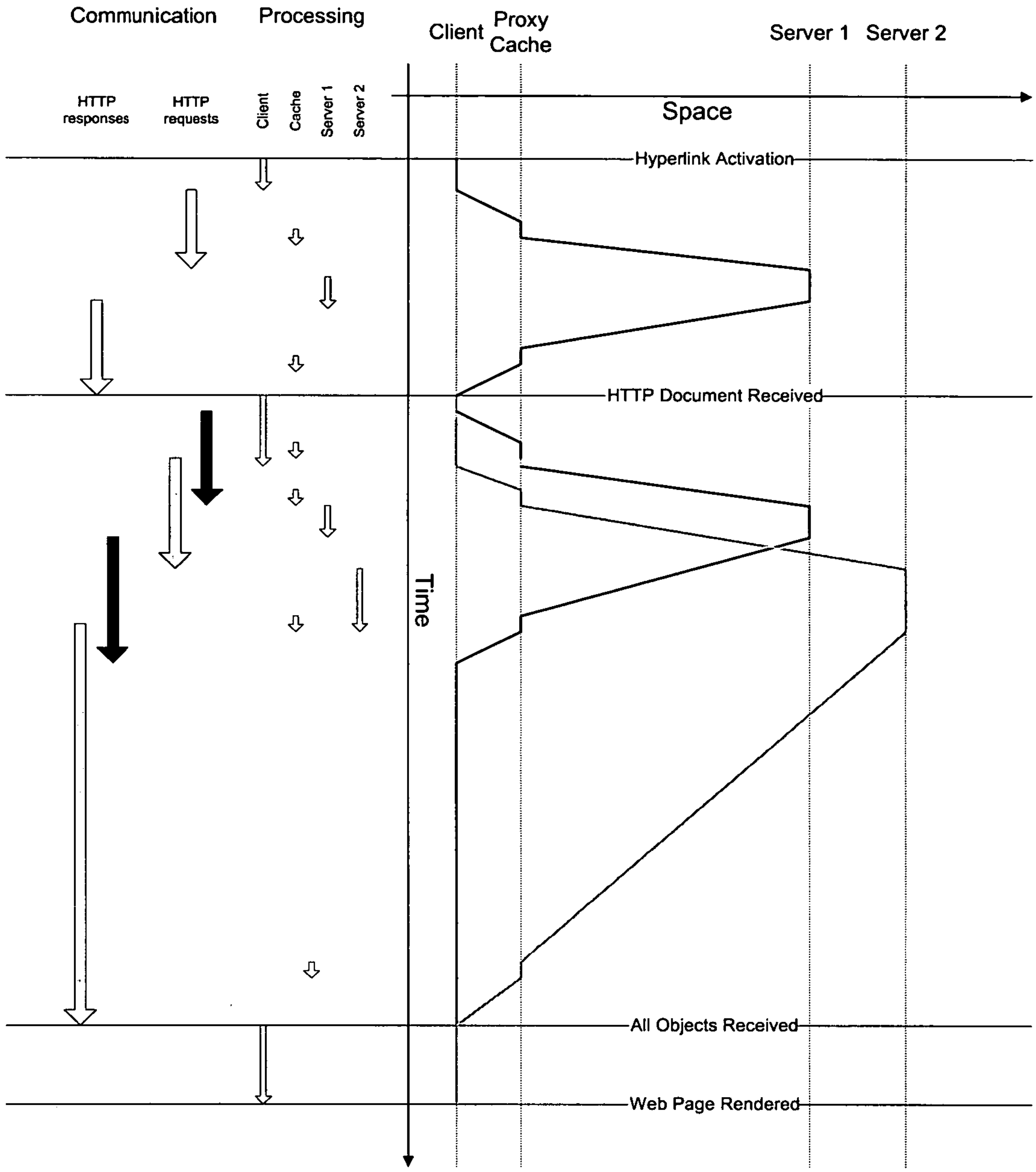


Figure 3.12: Timing diagram for hyperlink activation

The process, after activation of a hyperlink, for retrieving and displaying a Web page can be split into three important stages:

1. **HTML document retrieval** – The HTTP request for a resource and the subsequent response of an HTML document.
2. **Embedded object retrieval** – The HTTP request for each embedded object specified in the HTML document and the subsequent response of each object.

3. **Web page presentation** – The complete presentation of the Web page.

What is important about this process is that the WWW client may be able to start rendering the Web page as soon as the HTML document has been transferred. Whether the client is able to start rendering is dependent on the size and position of every embedded object being well defined in the HTML document. If this is the case, as soon as the HTML document has been received the textual parts of the Web page can be rendered, with spaces left for the embedded objects. These spaces can be filled in with an embedded object when (and if) the objects are successfully transferred. The dynamic nature of this presentation of a Web page indicates to the user that the Web page is not a single entity but rather a collection of objects that are transferred independently. In addition, this piecemeal rendering of the Web page reveals to the user some temporal properties of the communication: what will subsequently be termed temporal affordance (Chapter 4).[5]

The HTTP communications illustrated here have a strong relationship with the actual Web page they are helping to deliver. Every object has its own associated HTTP communication and so the relationship between these objects and the HTTP communication is very strong: there is a cogent mapping between each object and each communication. What has been lost is the concept of a Web page as a single information node. The simple model of the WWW as a collection of information nodes connected by hyperlinks has started to be eroded due to the behaviour that the user will experience when a Web page is requested and subsequently received as a number of temporally separated independent objects. Below the HTTP level of communication even the relationship between each object and each communication is lost.

3.3.3 The Transmission Control Protocol and the Internet Protocol

The Transmission Control Protocol (TCP) is entwined with the Internet Protocol (IP) and so the two are often referred to as a single entity, TCP/IP. Indeed, the protocol was originally conceived as one (Cerf and Kahn, 1974) and it was only after the introduction of the User Datagram Protocol (UDP), which is also entwined with IP, that people started to view the two independently. To send the HTTP request and response across the Internet the TCP is used in conjunction with the IP. A HTTP message is split up into units with the IP handling the actual delivery of each unit of data, and the TCP keeping track of all the units that make

[5] If the position and size of the embedded objects is not well defined in the HTML document the user agent must wait until after all the embedded objects have been retrieved before starting to render the

up the HTTP message. The TCP is a connection-oriented point-to-point protocol, which means that a connection is established between a single host (WWW client) to one other host (WWW server) and maintained until all the communication is complete. It is the job of the TCP to ensure that the HTTP message is split up into these units, and that they are reliably reconstructed at either end of the connection. To achieve this involves sequencing the units and requesting for the retransmission of particular units if required..

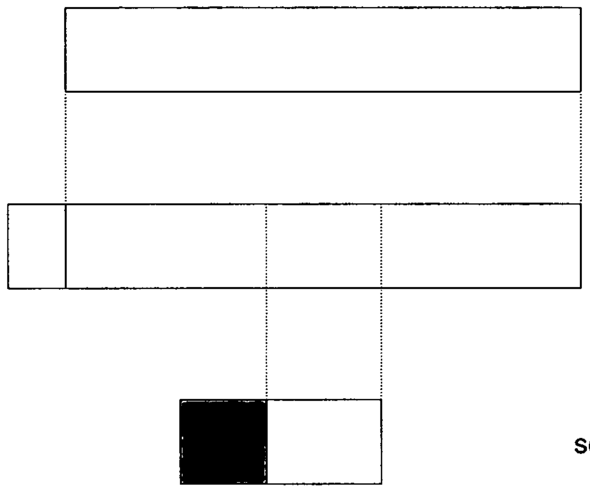


Figure 3.13: Fragmentation of HTML file (taken in part from Brebner 1997 p.351)

To transmit a HTML document a single TCP connection would be used for the HTTP request and response. Similarly, if an embedded image was referenced in a HTML document, another TCP connection would have to be established for this HTTP request and response, although the embedded image may reside on the same WWW server. As has been mentioned earlier, the HTTP 1.1 introduced the idea of a persistent connection for a whole Web page. What is actually persistent is this TCP connection, so the HTML document **and** any embedded object that resides on the same WWW server can be transferred using a single TCP connection. This eliminates temporal overhead associated with the establishment of every extra TCP connection.

The IP handles how each unit of data that makes up an HTTP message is sent from one host to another across the Internet. The IP must establish the address of the receiving and sending host. The address of the machine where the WWW server resides is contained in the HTTP request as a URL. The host part of the URL (Figure 3.13) is converted to an IP address

Web page.

using the Domain Name System (DNS) that, at any one time, converts the host part of the URL (called the domain name) into one IP address. The system for mapping domain names to IP addresses is distributed throughout the Internet in a hierarchy of authority.

Each host on the Internet has at least one IP address that uniquely identifies it from all other hosts. The IP address has traditionally[6] been a 32-bit number and is usually expressed by humans in *dotted quad notation*, which uses dots to separates the number into four eight-bit chunks each represented as a decimal number from zero to two hundred and fifty five[7].

`www.dcs.ed.ac.uk`

129.215.216.223

10000001.11010111.11011000.11011111

Figure 3.14: a) domain name b) equivalent dotted quad notation c) four, eight-bit number

Once IP has established the address of the sending and receiving hosts, these are added into each IP unit or packet. The IP is a connectionless protocol, which means that there is no established connection between both ends of the communication. Each packet is treated independently and as all packets contain both the sending and receiving host address, it is possible for them to be routed across the network differently. Each packet can arrive in a different order to the one they were sent in and, as mentioned before, it is the TCP that must reconstruct this order.

Below IP the packets of data are communicated using protocols specific to the network technology of that particular local network. There is a large number of different protocols for this data communication, but breaking an IP communication into different sub-protocols is considered beyond the scope of this description of the Internet. If, however, one of these sub-communications fails, a number of IP packets may not be successfully transmitted. This communication failure will eventually be realised by the TCP due to the sequencing of the IP packets and retransmission can be initiated. If the delivery of IP packets continue to fail the

[6] The traditional IP being version 4.0

[7] The limit on IP addresses has led to a new IP being developed which extends the IP address to 128 bits. This new protocol, IPv6, allows a theoretical maximum of over 6.65×10^{23} IP addresses (Thomas 1996) and is anticipated to allow for considerable further growth of the Internet . The uptake of IPv6 has not yet started and whether it will make any impact as an Internet protocol is as yet unknown.

TCP can, after a specified time, conclude that the total communication has failed and subsequently the download of the Web page has, either partially or totally, failed.

3.3.4 Summary of communication

A WWW server provides its data to the client in a standardised format called the hypertext mark-up language (HTML) through a standard communication protocol called the hypertext transfer protocol (HTTP). The HTTP has provision for a form of handshaking that utilises a textual header consisting of a number of lines each beginning with a standard name.

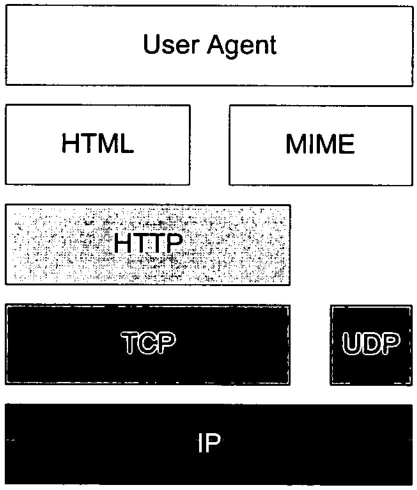


Figure 3.15: The communication 'stack' when transferring various WWW objects.

If a WWW resource specified in the HTML contains embedded objects (like an image) then there is a requirement for further HTTP handshakes to request and receive these objects. This HTTP communication between WWW client and server is achieved through the Internet's standard Transmission Control Protocol (TCP) and Internet Protocol (IP) which together, provides a *reliable* point-to-point stream. The communication model is illustrated in Figure 3.15.

3.4 Summary

The WWW consists of multimedia information nodes called Web pages. The Web page is, however, a conceptual object that exists only when the user agent renders it. The system has to manipulate a number of files before the user agent can construct the Web page in the user's browser. The most important of these files is the HTML document that defines how the text and multimedia objects should be structured when the Web page is realised. This HTML document explicitly includes the textual data and references the data for all the embedded objects. Every Web page must have a defining HTML document and the relationship between the two is further strengthened through the process of requesting a Web page. Although every object that is a constituent of the Web page is requested and sent using a separate HTTP communication, a user needs only to request the HTML document either explicitly through its URL or by the activation of a referencing hyperlink. A single user action produces a number of underlying communications that, if all are successful, results in the rendering of a Web page.

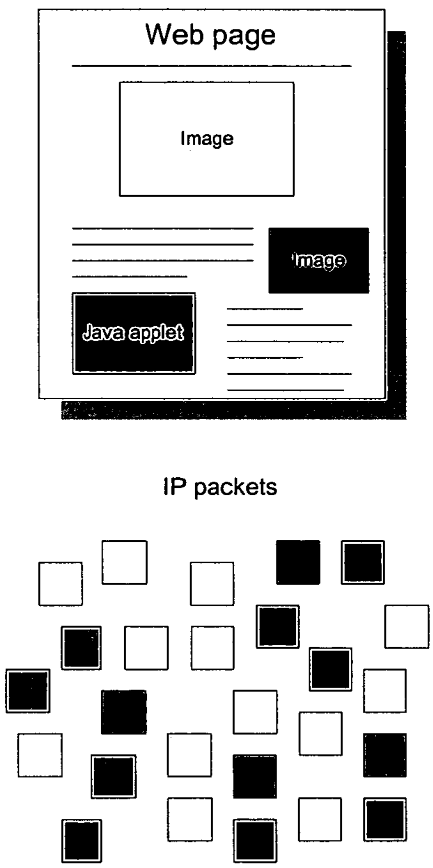


Figure 3.16: The deconstruction of the Web page during data transfer (clockwise from top left).

The simple abstraction of a Web page attempts to shield the user from the complexities of the communications required to transfer all the different type of objects from their various locations to the user's client (Figure 3.16). What is shown in later chapters of this dissertation is that this abstraction breaks down far too often, due to the communication failures and differences in the transfer time of the various objects that make up the Web page. Users are left to explain this behaviour using a model of the WWW that presents the Web page as a single entity: an indivisible hypertext information node.

In Chapter 4 the argument is moved away from the complexities of the technology towards the usability issues of the WWW. This requires a change of perspective from the design view used in this and the previous chapter, where a number of world-wide communications were viewed in their entirety and dissected, to the user view that perceives the system only through the WWW browser interface and as a result does not have the information needed to explain the unpredictable behaviour witnessed.

Chapter 4

Interacting with the WWW

The previous chapters introduced a model of the WWW based on the hyperlinks between Web pages. Underlying this abstraction are the mechanisms for transferring data using the Internet. In this chapter it is argued that the boundary between the WWW abstraction and the Internet is inflexible and not sufficiently explained through the interface, and the result is poor usability. The WWW browser presents an interface designed to help the user navigate through the WWW: the central task of the interaction being the activation of hyperlinks. The usability problems that users face can be categorised into three. First, there are the problems based on the system's dynamic behaviour not being sufficiently explained through the interface. The data transfer process is unpredictable and this may result in delay or even failure. With insufficient information concerning this process, users are unable to improvise successfully and explain the behaviour of the system in a particular situation. The second type of usability problem occurs because the hyperlink does not sufficiently explain the relationship between information contained in the hypertext. Users must completely retrieve an information object before being sure how relevant the information is to them. Discovering that the information is useless may come at a considerable *cost* due to time taken by the retrieval process. The final problem is related to the quality of Web pages. The

WWW has many users with varying needs, interests and opinions. Rating the quality of a Web page is a subjective process that the present system does not address. Recommendations based on an opinion of quality are an important area of usability as it can affect how a user navigates through a hypertext.

The WWW attempts to ignore the temporal delays of the Internet communications and only provides one simple abstraction to explain the complex relationship between information. It quickly becomes apparent while using the WWW, that the delays expose themselves through the system's behaviour and that the relevance of information is often unknown before activating a hyperlink. Browser designers have added limited facilities to try to reveal the communications of the Internet, but these facilities do not present a complete model of the communication and the user is left to make sense of misleading feedback and inexplicable behaviour. The Web page is one of the key entities for the WWW abstraction, but the behaviour users experience centres on the transfer of the HTML document and each embedded multimedia object.

Chapter 1 showed how the single abstraction employed by the London Underground Diagram restricted the traveller's ability to probe the system for other types of information like, for example, the time of travel and the exact geographic location of stations. The abstraction of Web pages connected through hyperlinks restricts the WWW user in a similar fashion. A WWW user is initially presented with a Web page, be it the result of a search, a form of *Web portal,* or a previously visited Web page. From this position the Web user can choose to activate one of the many hyperlinks embedded in the Web page or to explicitly type in a new URL. What information does the WWW provide to help the user make this decision? How can the user decide on the relevance or quality of the information referenced by a particular hyperlink? How long will a particular Web page take to appear after hyperlink activation? Would activating another hyperlink take more or less time to transfer data? What factors affect the data transfer process after hyperlink activation at this particular moment in time?

A user will not always want to know the answer to all of these questions, but in a particular situation (*i.e.*, they are in a hurry or they require a very specific piece of information) they may want to know the answer to some of these questions, or indeed other questions. Just like the underground railway traveller, a particular WWW user, undertaking a particular information retrieval task, at a particular moment in time, will want answers to a varied and diverse collection of questions. Their particular requirements are situated in the moment and related to the specific task they are undertaking. The rest of this chapter presents exactly what information the WWW interface affords the user and more specifically what information it conceals.

4.1 The WWW interface

The WWW interface performs three main tasks: it presents the information objects to the user; it provides mechanisms for navigating through the hypertext; and it communicates with the network. The WWW interface presents a hypertext to the user by rendering the current Web page and providing facilities to activate all the hyperlinks contained within it. The interface also acts as a network client, sending data to and receiving data from network servers using Internet technology. The part of the interface that interacts with the user is the *user agent* and the part that interacts with the Internet is the *WWW client*. Together they form the *WWW browser* (from now on referred to as the browser).

The browser presents the WWW information space to users by providing various navigation facilities to help them move through it. The abstraction presented to the user is a simple hypertext with its information nodes – Web pages – and the connections between nodes – hyperlink. Through the activation of a hyperlink the user can induce the rendering of a new Web page. Since the WWW is a hypertext, it has inherited the usability problem associated with *standard* hypertexts. In fact, some of the usability problems that have been observed in the use of hypertext systems are accentuated by the particular characteristics of the WWW: its immense size; the diversity of its content; its variable quality; and the decentralised nature of its construction that leads to the movement or deletion of Web pages.

The topology of the information space is one cause for usability problems with hypertext. Users often become disorientated in an information space, resulting in them eventually becoming *lost* (Ayers and Stasko, 1995). Similarly, hyperlinks may become broken if information nodes are moved or deleted, this causes a disturbance in the space that can lead to confusion or annoyance (Nielsen, 1998a). These types of usability problems are of real concern for the WWW due to its immense size, the complex connectivity of the information space and the unregulated and transient nature of hyperlinks and Web pages.

As well as usability problems arising due to the *spatial* properties of the information space, usability problems may also arise due to the actual content of Web pages. Not knowing what the content of a Web page will be before hyperlink activation is a major problem. The observed effect is that a user may activate a hyperlink only to realise that the referenced resource is useless and therefore decides to return to the referencing node (Conklin, 1987).

These two types of usability problems – spatial and informational - arise because the browser reinforces the abstraction that the WWW is a hypertext: albeit with some unique characteristics. The Web page is an information entity associated with other Web pages through hyperlinks. There are, however, a further set of usability problems not often observed in traditional hypertext systems. These stem from the WWW not fully adhering to hypertext's defining principle - the rapid retrieval of information – and this results in the breakdown of the hypertext abstraction. This set of usability problems is caused by the distributed nature of the information space, the Internet technology used to transfer data across the globe, and the temporal properties of the data transfer process.

It quickly becomes apparent when using the WWW that the simple underlying abstraction of a seamless hypertext is very rarely sustained in practice. Rather than experiencing an instant and deterministic response when activating a hyperlink, a user must usually cope with an unpredictable delay related to network performance fluctuations and failures. The information retrieval process is often far from rapid and the user experiences a delay.

System response times are recognised as a major factor in determining the usability of a system. Delays in responding to a user's action have been identified not only as a cause of annoyance, but also as a catalyst for changes in users' strategies (O'Donnell and Draper, 1995). Limitations of technology have meant that interface designers have always had to balance the increase in the computational demands of the graphical interface, with the reduction of time delays in responding to the user. The shift from stand-alone machines towards distributed systems has increased the problem of response times as network delays have started to become significant in comparison to computational delays (Dix, 1994). For example, after activating a hyperlink a user may experience a considerable delay. The user will be unaware of how long this delay may last. In fact, the delay may be caused by a communication failure and no amount of waiting will result in a new Web page being rendered. When confronted with such a delay, users are faced with a dilemma: should they continue waiting for the data transfer process to complete or should they interrupt the process and try again?

In this chapter is presented an investigation into the usability problems that face the WWW user. The problems are categorised into three domains: information, space, and time. Two of the usability problem domains are concerned with the hypertext nature of the WWW (information and space) and one is concerned with the distributed nature of the WWW and the behaviour of the network (time). This is, however, only a framework and many usability problems arise from an intersection of these domains and may be due to both network behaviour and hypertext usage. The investigation starts with an examination of the browser and its communication with the user, the actions that the browser performs and the feedback that the browser provides. Following this is a review of the literature concerned with the characterisation of the WWW and the usability problems these characteristics entail.

4.2 The browser

After the successful activation of a hyperlink, the WWW client receives the referenced resource. If this resource is a Web page, the WWW client will have the HTML document and all the embedded objects. The user agent must then unravel the presentation information contained in the HTML document from the content. Only then can the Web page be presented to the user. Once the Web page is displayed the main task of the browser is to allow the user to move to another information node by activating a hyperlink.

4.2.1 Activating the hyperlink

Current browsers support two basic mechanisms for requesting a new information resource. A hyperlink contained in a Web page can be activated by *clicking* on the hyperlink's activating object that is specified in the HTML document as the content of an anchor element. The activating object usually takes the form of some text or an image, and a browser *should* highlight these objects. In the default state of most graphical-browsers, images are highlighted with coloured borders and text is underlined and coloured. To emphasis further the existence of a hyperlink, the mouse pointer changes shape when it passes over one of these activating objects. By *clicking* the mouse button when the pointer is over an activating object, an HTTP request is made for the referenced resource and the browser subsequently renders this new resource in the browser window when (or if) it arrives.

A resource can also be requested using a navigational aid provided by the browser. In general, these navigational aids allow users to retrace their route through the information space or return to a predetermined Web page. Hyperlink activation that employs a navigational aid or an activating object to request a resource is defined as *implicit activation.* The choice of this term rises from the referenced resource's URL being implied, either through the anchor element in the HTML document, or by the browser's navigational aid. Users do not need to know the URL of the Web page they request when using implicit activation.

The second mechanism for requesting an information resource is to request it by typing in the resource's reference (in the form of its URL) into the browser's address bar. All browsers provide a facility to enter a resource's reference into the address bar. This form of hyperlink activation is described as *explicit activation* since the user must know a resource's URL. The distinction drawn between these two hyperlink activation methods is, therefore, dependent on whether the user requires direct knowledge of the requested Web page's URL, or whether the user can request the Web page by simply *clicking* a button or activating object.

4.2.2 Browser aids

Figure 4.1 shows one particular rendering of a Web page in the Microsoft Internet Explorer (IE) browser that, at the time of writing, is estimated to be the most popular in the world[1] (Figure 4.2). The part of the browser that takes up the most screen real estate is the *information window* where a Web page is presented to the user. Above the information window is the *address bar*, which displays the resource's URL. The rest of the real estate is divided up between navigational features for moving through the information space, network features for interacting with the WWW client, and communication feedback information.

---

[1] During the period of the research reported here and the work reviewed in this chapter, the popularity of browsers has changed significantly. The experiments undertaken for this dissertation have used the Netscape browser as well as IE. Some experiments reviewed in this chapter used a derivative of the now obsolete Mosaic browser (for example, Xmosaic). The general appearances of all these browsers are very similar and little has changed in the actual default interface. However, wherever possible, the exact flavour of the browser referred to in the text will always be specified.

4.2.2.1 Navigational features

The navigational features shown in Figure 4.1 are the *Back*, *Forward* and *Home* buttons. The *Back* and *Forward* buttons are available on nearly all current WWW browsers and they allow WWW users to trace and retrace their history of Web pages. Tracing and retracing recreates the temporal context of a Web page, which is an important factor in reducing disorientation (Tauscher and Greenberg, 1997). Similarly, the *Home* button provides a simple aid to request a single WWW resource that the user has previously defined. These three navigation features provide the main facilities to move around the hypertext. Some browsers also provide a more direct access to a user's history through a *History* button or menu option that will display a textual list of the Web pages visited. The user can select any Web page from the list and the browser will request it.

4.2.2.2 Network features

The facilities that allow a user to interact with the network are limited to the *Reload* and *Stop* buttons. The *Stop* button provides the facility to interrupt the data transfer process and stop all further communication. Similarly, the *Reload* (or Refresh) button also interrupts the data transfer process, if one is taking place, and then requests the present Web page again.

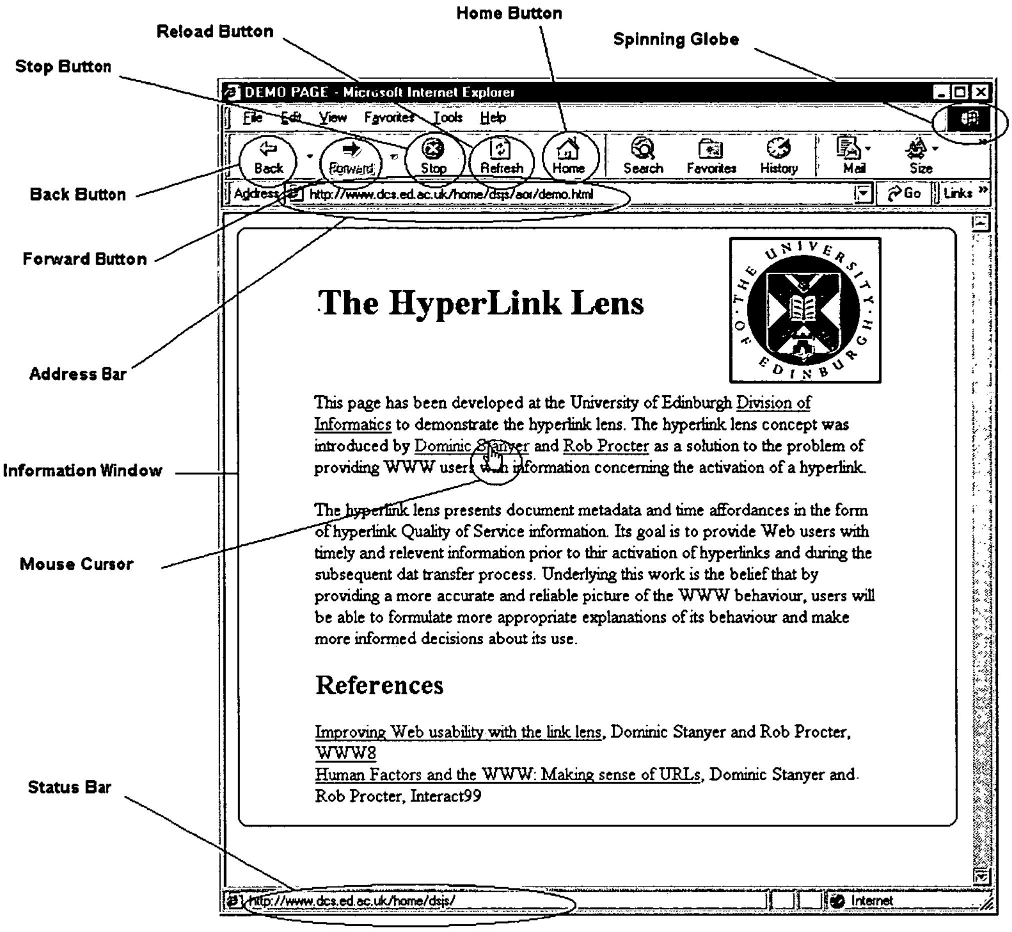


Figure 4.1: The Microsoft Internet Explorer WWW browser

4.2.2.3 Feedback features

The browser also provides some limited feedback on the data transfer process. The *status bar*, situated at the bottom of the browser, provides some limited information about the data transfer process (for example, the particular HTTP request it is performing) and some estimates about how long a particular download will take. There is also the *spinning globe*, situated in the top right of the browser, which animates if, and only if, the data transfer process is taking place. Although it is generally referred to as the spinning globe, this feature can take the form of many different objects like shooting stars or an animated letter. Nevertheless, all the major browsers provide a form of animated feedback.

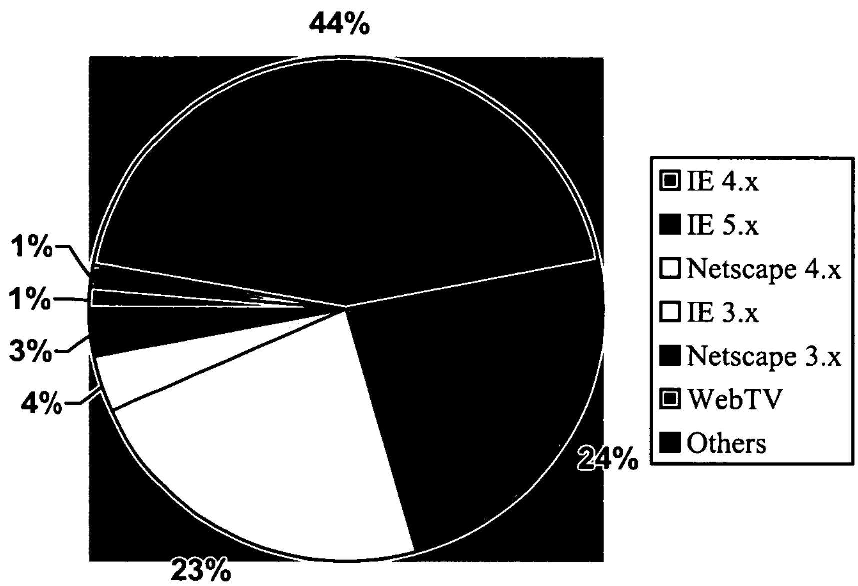


Figure 4.2: Thursday August 4th 1999 StatMarket analysis of browser usage http://www.statmarket.com/

To recapitulate, the browser provides a number of facilities to interact with the hypertext and others to interact with the Internet. It also provides some limited feedback concerning the data transfer process.

- **Hypertext navigation features** - *Back*, *Forward*, *History* and *Home* buttons
- **Internet interaction features** - *Stop* and *Reload* buttons
- **Communication feedback** - *Spinning globe* and *status bar*

In the following section studies that have been designed to investigated how each of these features is used are examined.

4.2.3 Studies of browser usage

The decentralised nature of the WWW makes measuring and gathering characterisations of WWW usage difficult. Data can be collected from a number of different locations in the system: the browser; the server; an intermediate proxy or gateway. Analysing data collected at the browser presents two major drawbacks. The first is that the browser software has to be modified to allow the particular user actions under study to be logged. This is not always feasible because the complexity of the source code and the commercial interests of the browser manufactures prevent access to the source code. Using a proxy to intercept information requests and responses, or monitoring operating system level events, are both possible techniques that circumvent the need to alter browser software. However, unlike the modification of the browser source code these alternative methods cannot capture all user events and this affects the scope and accuracy of any results obtained. Changing the browser source code, so that user actions can be collected, although not often feasible, is desirable when studying browser usage.

The second drawback to browser-side studies is that individual users' interactions have to be studied and this entails a high experimental overhead. Unlike server-side logs of HTTP requests and responses that have the potential to assess a large and diverse range of users, browser-side studies are in more danger of creating a bias towards a particular user segment. These problems are perhaps the reason why this type of study is quite rare. Despite these potential problems, two studies discussed below have managed to gain access to browser source code and, after modification of the code, have undertaken browser-side studies of WWW users. These studies present what interaction events users actually employ and provide an insight into how users actually use the browser.

Catledge and Pitkow performed the first browser-side characterisation of the WWW in 1994 (Catledge and Pitkow, 1995). The study was conducted on 107 subjects and used a fully instrumented version of the Xmosaic browser (estimated to be the most popular browser at that time). It recorded events using a set of task representation guidelines. The general findings of the study were that two events dominated browser usage: using a hyperlink's activating object to request a new Web page (51.9% of events) and using the *Back* button to request the previous Web page (40.6% of events).

In 1995, Tauscher and Greenberg conducted a browser-based study using a similar experimental methodology to Catledge and Pitkow (Tauscher and Greenberg, 1997). Tauscher *et al.* employed 28 university students to use the same fully instrumented version of Xmosaic but introduced some additional metrics. The extensive analysis of these results, when compared with a re-evaluation of the Catledge and Pitkow data[2], uncovered the same general findings. Hyperlink activation accounted for 50.1% of events and the use of the Back button for 30.2% of events.

[2] Tauscher's conclusion that the WWW is a recurrent system is examined later in section 4.6.2.

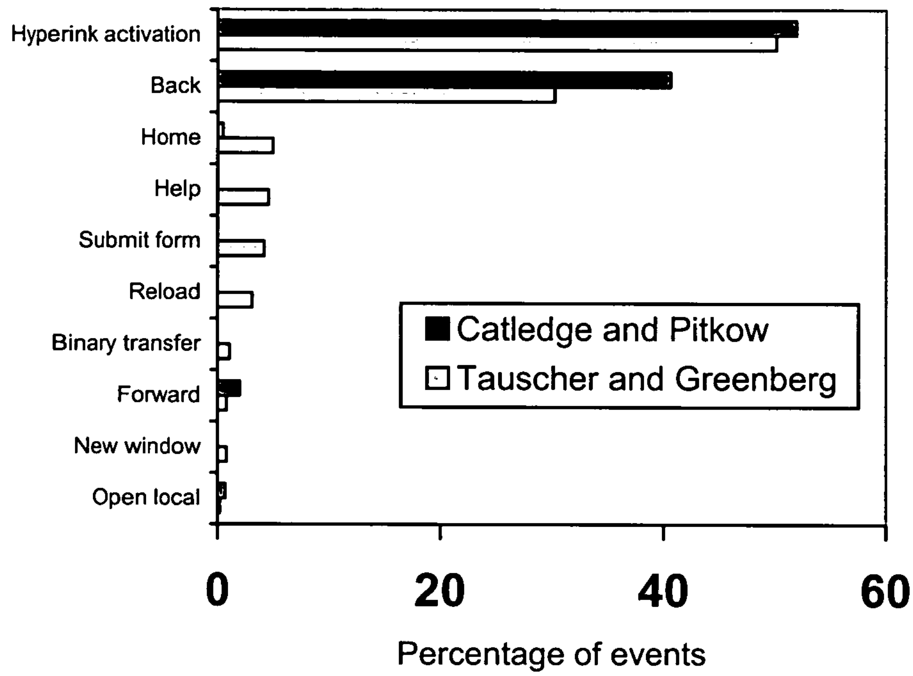


Figure 4.3: Comparison of the two browser-side studies of browser events

The results from these two studies show the predominant use of the hyperlink's activating object and the *Back* button (Figure 4.3). In fact, the hypertext navigation features of the browser, coupled with the hyperlinks in Web pages, dominate the interaction events employed by users. Catledge and Pitkow suggested that the minimal use of the *Reload* and *Stop* browser features showed that users were "*insensitive to retrieval delays*". Both studies used logs of user events to produce their statistics. There was no actual observation of the user using the WWW. The finding that the user is insensitive to the retrieval delay is based solely on the statistical use of the *Reload* button. The investigations undertaken as part of this dissertation employed observational studies that clearly show users, not only struggling to cope with delay, but also having a greater awareness of potential delay before hyperlink activation. The conclusion drawn by Catledge and Pitkow is therefore refuted

4.2.4 Communication feedback

Helping users navigate through the WWW is the primary task of the browser. Activation of the hyperlink and using the *Back* and *Forward* buttons are interaction elements that help users to achieve this navigation. The browser also provides limited feedback about the actual Internet communications that underlie it. This communication feedback is presented in the status bar and by the animation of the spinning globe. The exact nature of the information presented in the status bar is dependent on how much of the optional HTTP request and response header attributes are in use and the type of resource being transferred. This results in an unpredictable presentation of the communication. To help characterise this temporal feedback, the following paragraphs explain a framework which describes temporal interaction: Conn's complete *time affordance* (Conn, 1995).

Property	**Definition**
Acceptance	Acknowledgement that the task has been correctly specified and is executable.
Scope	The overall size of the task and the corresponding time the task is expected to take barring difficulties.
Initiation	How to initiate the task and, once initiated, clear indication that the task has successfully started.
Progress	Clear feedback as execution of the task proceeds, what additional steps have been completed within the scope of the overall task, and the rate at which the overall task is approaching completion.
Heartbeat	Constantly available visual indication that the task is still alive.
Exception	A task that is alive has encountered errors or circumstances that may require user intervention.
Completion	Clear indication of termination of the task and the status of termination.
Remainder	Indication of how much of the task remains and how much time is left before completion.

Table 4.1: Conn's complete time affordance.

Conn extended the concept of affordance proposed by Norman (1988) and presented a taxonomy of eight properties which he argued constituted a full set of delay information that users must have in order to understand and respond appropriately to delays. He called the presentation of these eight properties *a complete time affordance* (Table 4.1). In order for a user to fully understand the nature of the temporal delays that occur during the data transfer of WWW pages, the browser must satisfy all eight of Conn's requirements. Despite some claims by Microsoft (IE4 White paper), no browsers yet provide a complete time affordance.

> *"Internet Explorer 4.0 has improved user feedback, so users always know what the browser is doing. Sound effects help cue the user to when the page has been*

requested, transmitted, and received. An improved status bar details how long the process will take." (IE4 White Paper, p 7)

Property	**Browser fulfilment**	**Reason**
Acceptance	Unsatisfactory	A type of acceptance is fulfilled once hyperlink activation has occurred, the user is presented with another change in mouse pointer and status bar, and the globe starts revolving. Knowledge that the reference is correctly specified is not available, since reference may be incorrect.
Scope	Unsatisfactory	The task in transferring a Web page and status bar information is concerned with different HTTP requests and responses
Initiation	Satisfactory	Fulfilled as a hyperlink's activating object is highlighted, both visually in the browser's information window and dynamically when the mouse pointer changes and the referenced resource's URL is displayed in the status bar.
Progress	Unpredictable	The Web page itself provides some form of progress affordance, as parts of it may start to be rendered before the total data transfer of other embedded objects is complete. This is, however, an unpredictable form of progress that is dependent on how the HTML document specified the Web page. The status bar also provides some form of progress affordance, though it is unpredictable and related to HTTP requested objects.
Heartbeat	Satisfactory	Fulfilled due to globe continually revolving throughout data transfer.
Exception	Unsatisfactory	The only exception is when the status bar reports that the communication has stalled or when the communication times out.
Completion	Satisfactory	When all objects that make up a Web page have been transferred, the progress bar states that the process is complete and the globe stops revolving.
Remainder	Unsatisfactory	The status bar does sometimes provide a remainder indicator but it is unpredictable and is concerned with embedded objects rather than the Web page as a whole.

Table 4.2: Fulfilment of time affordance by WWW browsers

Table 4.2 presents a summary of which requirements the browser presently satisfies. The problem faced by the browser in satisfying the eight requirements is that it presents to the user an abstraction of the hypertext with Web pages and hyperlinks. The temporal properties of the communication are, however, dependent on the actual objects that make up the Web page: the HTML document and the embedded objects.

To present a complete time affordance the browser would be required to break down the abstraction of the Web page into its constituents and show the progress of each of these parts during the communication. The result is a conflict between presenting the WWW as a hypertext with the navigational aids required to navigate through it, and providing a temporal affordance that shows the make up of the Web page as a collection of objects of different types (perhaps) stored in different locations.

The browser has no alternative other than to use the hypertext abstraction and the result is an unsatisfactory time affordance when Web pages are downloaded. Just like the London Underground Diagram presented in Chapter 1, the WWW provides only one static abstraction of a hypertext and does not have the flexibility to allow the presentation of a more appropriate abstraction to explain the data transfer process in full. The result is that users are left to explain and react to the behaviour they experience during the data transfer of a Web page using inappropriate information. Their capacity to act appropriately and effectively is therefore compromised. For example, if an HTML document does not explicitly specify the location of all the embedded objects in a Web page, no information will be rendered until all the data for all the objects has been transferred. A failure in the data transfer of any of these embedded objects will result in nothing appearing in the information window of the user's browser. The user has to explain the communication disruption of the Web page unaware that only one object from a single location has failed to transfer. If this object is not vital, stopping the data transfer process may be desirable, since this will force the browser to render the Web page with the objects already obtained. Alternatively, realising that the missing object is stored on a different server may provide the information that the user needs to improvise in this situation and make a judgement about whether to wait or to interrupt the data transfer process. The interaction between the user and the WWW is restricted during the data transfer process because the abstraction supplied by the browser does not allow the user to prod and probe the system in order to improvise if circumstances require that they should.

4.3 A framework for WWW usage

So far, in this chapter the browser and the various actions available to the user in navigating around the information space have been introduced. A short review of some statistical work has provided an overview of what actions users actually employ and how the hypertext abstraction prevents the browser from providing a complete time affordance. The remainder of the chapter looks in detail at work that has characterised the WWW and at work that has investigated some of the usability problems users face. To present this considerable amount of work a framework is proposed that can be used to categorise both the characteristics of the WWW, and the usability problems that result. This framework was introduced at the beginning of the chapter but some background to its adoption is provided in the next section. A framework for hypertext usage is presented and shown to be inappropriate as it is unable to explain the underlying processes. A simple communication framework is therefore introduced which is shown to fulfil the requirements.

4.3.1 Spatial, semantic and social navigation

Dourish and Chalmers (1994) have suggested that interface designers were overusing *spatial* navigation mechanisms. Since the real world is organised spatially it is natural for users to use spatial analogies when *'moving'* through information. Dourish and Chalmers recognised the importance of this mechanism but argued that by combining it with two other navigation mechanisms a better overall interaction could be achieved.

The second navigational mechanism was termed *semantic* navigation. If a system can group information objects together according to similarity, then there exists a semantic relationship between them and this relationship can be exploited to facilitate semantic navigation. Dourish and Chalmers used a hypertext navigational aid (like the *Back* button) as an example of the semantic navigational mechanism being mapped onto a spatial arrangement.

> *"In hypertext systems, for instance, the primary form of navigation is semantic – in particular, associated with the (domain –dependent) semantic properties of various links. It's not uncommon to add a spatial representation to these systems, although typically these model not so much properties of the*

information as properties of a user's 'route' through it." *(Dourish and Chalmers, 1994, p 1)*

As well as the hidden semantic navigational mechanisms, Dourish and Chalmers also introduced a third navigation mechanism: *social* navigation. If a system has some form of collaborative activity, a form of social navigation may arise. Social navigation implies that the strategy a user employs to navigate through the information is influenced by the activity or recommendation of other users.

"Here, then, there is a model in which the opportunity to explore information is based not on either location or content, but rather on recommendation and social factors." (Dourish and Chalmers, 1994, p 2)

One way that social navigation appears in the WWW, which is not explicitly collaborative, is in the form of Web pages that label hyperlinks as *"interesting links"*. Another is the content of an author's *homepage,* which by definition contains hyperlinks that the author has implicitly recommended. The fact that a hyperlink is situated in a Web page means that a user experiences a form of implied recommendation from the referencing Web page. Similarly, the popularity of a Web page or Web site can form a basis for recommendation. These are forms of social navigation.

The Dourish and Chalmers' framework provides a categorisation for navigational aids. A hypertext designer can use the framework to implement new and innovative features for navigating the hypertext. The framework is, therefore, relevant to the hypertext: it is fixed within the abstraction of Web pages and hyperlinks. Spatial navigation relates to the spatial properties of the hypertext, how information nodes are connected. Semantic navigation refers to the semantics of the hypertext, the content of the information nodes. Social navigation applies to the views of the hypertext readers and authors. The Dourish and Chalmers framework is based in the hypertext abstraction.

The framework tries to capture information concerning how relevant an information node may be to a hypertext user. Before activating a hyperlink, users can be presented with a number of relevance abstractions to help them assess whether the referenced information node will be useful. The information node may be spatially near – the author provided a hyperlink to associate the current Web page with the referenced Web page so they must have something in common. The hyperlink may contain a semantic property tying the Web pages together – the referenced Web page is, for example, the *index* to a whole set of resources on this topic. Another party can recommend the information node – it was relevant to this user so it may be relevant to you. In a particular situation users may be able to make use of one or more of these relevance abstractions to help them assess how useful the referenced Web page will be to them, before actually downloading the object. The Dourish and Chalmers framework provides interface designers with a set of possible approaches to provide users with abstractions on the relevance of two information nodes.

The framework required by the work presented in this dissertation must be able to cope not only with the hypertext abstraction and the relevance of information, but also the Internet communication. The Dourish and Chalmers' framework does not provide facilities to explain the dynamics of the WWW. It only describes the content and relationships of the hypertext. The framework that is required must be able to explain the temporal properties of the WWW and the disturbances and failures of the data transfer process that often manifest themselves as delays after hyperlink activation. A framework to describe the communication between two computers may better fit these requirements.

4.3.2 Information, space and time

Brebner identified three separate domains which together totally describe a communication between computers (Brebner, 1997). By retrospectively describing the communication and taking the perspective of an external observer, both the static and dynamic makeup of the communication is known and can be presented within the framework. The static component of the communication is described using the *information* and *space* domains of the Brebner model, and the dynamic component is described using the *time* domain.

- **Information** – type of information communicated.
- **Space** – the computers involved in the communication and the channel used to connect them.
- **Time** – when the communication happens and how the communication proceeded over time.

These three domains are considered by Brebner to be conceptually orthogonal and so the relationships between the domains describe the nature of the communication. Changing an informational component by, for example, compressing the message, can affect the description of the communication in another domain, for example, the total time to communicate the message.

The final concept central to this model is the abstraction of a communicating *channel*. A channel connects communicating computers together and allows information to flow in either direction between them. Although an actual communication between two computers (like the request for a Web page) may send messages through a number of intermediate computers and wires, using a number of different protocols, the whole communication can be described using a single channel. At a particular time a communication that uses this *virtual* channel can be totally described using the three domains. What is more, the information and space domains describe the static makeup of a communication and so they exist independently of a communication.

In theory, any communication can be described using this framework. The framework provides a means to describe the WWW in terms of both the Internet communications and the hypertext abstraction. The information domain can be used to describe properties of the hypertext abstraction through to the IP packet: the Web page; the HTTP requests and the responses of its constituents - HTML document and embedded objects; the TCP connection; and the IP packet. It can also be used to describe the actual content of the Web page and the semantic properties of the hypertext. Similarly, the space domain subsumes both the spatial properties of the hypertext and the topology of the Internet, including the physical location of each object's WWW server. Finally, the time domain allows a method for describing the dynamics of the behaviour that users experience during data transfer[3].

There now follows a description of various studies of the WWW presented within the three-domain framework. What the framework allows is a detailed investigation into issues concerning both hypertext and Internet technology. In the information domain, studies that investigated the constituents of a Web page - the number of images and hyperlinks they contained – are presented alongside studies that proposed relationships between a Web page's *popularity* and its size. In the space domain, the Web site – a set of Web pages authored by a single organisation – is introduced, adding a new dimension to the hypertext's spatial properties. In the time domain, delays in system responses are shown to be

[3] The only area that cannot be described in this three-domain framework is the social aspect of the hypertext that was highlighted in the Dourish and Chalmers model.

contributing factors to user's annoyance and the perceived quality of the downloaded Web page.

Before starting a presentation of the literature one final point needs to be made. Some issues can be considered from more than one domain. For example, the use of the *Back* button and the usability issues it raises can be considered as being related to both space – navigating through the hypertext – and time – it is based on a user's *history* and provides a temporal context. In order to avoid the confusion, whenever an issue could be described in an alternative domain it is clearly stated.

4.4 Information

A classic usability problem identified in hypertext literature (and mentioned previously in this chapter) is that of users leaving a referencing node only to immediately return after determining the uselessness of the referenced node (Conklin, 1987). Nielsen termed this usability problem *"going down the garden path"* (Nielsen, 1998b). The problem stems from the hypertext user being insufficiently aware of the content of the referenced node before hyperlink activation. In general, a user can only assess the content of the referenced node using the information provided by the referencing node: the activating object itself and possibly some text that leads to the activating object. In the WWW, it is left to the Web page author to provide information to help a user assess the contents of the referenced node. Although hypertext systems have implemented overviews aimed at providing useful information to aid this assessment of content prior to hyperlink activation (called a *'gloss'* (Zellweger *et al.*, 1998) or *'scent'* (Furnas, 1997)) the WWW is some way behind (although some work in this area is presented later in Chapter 8).

In general, printed materials have undergone a form of editorial filtering. In the case of scientific journals and papers, the work has most likely passed through peer review. The very fact that the work appears in a journal or book provides a cue to its quality and possibly the style of writing. What is more, a well-respected author or a particular publisher may provide further information on the quality of a piece of work. For example, newspapers come in two distinct styles, broadsheet and tabloid, and by using this simple visual cue, a reader is able to make some form of judgement on the style of the newspaper's content before reading the paper. Web pages on the other hand, can be published by anyone about anything, leading to works that may not only be devoid of e.g., academic value but may even be misleading and erroneous. Of course, the WWW also provides the opportunity for good quality works to be promulgated outside the normal channels and an author who would otherwise have been denied the opportunity to publish may just as easily enrich a subject. Providing cues to the quality of information and the presentation style before actually downloading it is therefore very difficult. The WWW hides the content of the referenced node behind the hyperlink and users are left unable to determine anything about the style and quality of its content before download.

The following sections try to characterise the Web page - its size and the makeup of its content – as a first step in determining if any of its characteristics may be useful cues to the style and quality of its content. Providing this information to the user will provide a cue to quality and help to solve the problem of assessing the referenced resource before hyperlink activation, thus preventing the annoyance and frustration that arises from unnecessary downloads.

4.4.1 The 'average' Web page

A Web page is defined by an HTML document that contains references to all the embedded objects. This distributed property means that describing a Web page as a single entity becomes difficult and research has avoided this problem by producing statistics on the HTML document rather than the Web page. It is not the Web page that is transferred by the Internet, but rather the HTML document and the embedded objects. A number of studies have surveyed Internet traffic, in particular HTTP data transfers, to determine the average size of WWW related data. These studies have generally been involved with networking issues, such as the construction of caches, and have often failed to interpret the data with respect to the Web page. Analysing separate HTTP responses means that studies distinguish HTML documents from any type of embedded object and do not capture the Web page as a single entity. Despite this general trend towards collecting data about HTML documents rather than Web pages, some interesting findings can be drawn from the data.

Woodruff *et al.* (Woodruff, Aoki, Brewer, Gauthier and Rowe, 1996) used a data set comprising 2.6 million HTML documents to extract some characteristics of the WWW. They reported that within the data set, the mean size of HTML documents without the markup was 4.4 Kbytes and the median was 2.0 Kbytes. Of particular interest was the finding that about 90% of their HTML documents contained at least one anchor element[4].

Bray (1996) conducted a similar analysis using a data set comprising 1.5 million *"textual documents"*. Bray reported that the mean size of the textual documents was 6.5 Kbytes and the median was 2.0 Kbytes (the standard deviation being cited as 31.6 Kbytes). Although constructed in a slightly different way to the Woodruff *et al.* study, Bray concluded that just under 75% of these textual documents contained at least one hyperlink and 50% of these documents contained at least one image. The results from these and other studies are summarised in Table 4.3.

	mean	median	mean (without mark-up)	median (without mark-up)	at least one image	at least one hyperlink
Woodruff et al. 1996	-	-	4.40k	2.00k	~ 60%	~ 90%
Bray 1996	6.52k	2.02k	-	-	~ 50%	~ 75%
Gwertzman et al. 1996	4.79k	-	-	-	-	-
Lawrence et al. 1999	18.70k	3.90k	7.30k	0.98k	-	-

Table 4.3: Size and content of HTML document.

4.4.2 Summary

The hyperlink does not provide the user with much information about the referenced Web page. To assess the informational content of a Web page a user must download it. This may result in unnecessary data transfer and annoyance for the user. In the real world, objects offer visual cues to users that allow assessment before any operation of the object: the hyperlink could do the same. Between 50-60% of Web pages contain at least one image and between 75-90% of Web pages contain at least one hyperlink. Estimations of the mean size of HTML documents range between 5 and 19 Kbytes. The WWW is continually changing and the data used to present this characterisation should be treated with caution. The Web

[4] Similarly, about 60% of their HTML documents contained at least one occurrence of the IMG tag, and hence contained an embedded image. The mean number of IMG tag occurrences per HTML document was just under four.

page may not yet be stable in terms of the characteristics that are highlighted and so the validity of these conclusions must be used with care.

4.5 Space

In Chapter 2, it was argued that the WWW could be viewed as a directed graph whose vertices are Web pages and whose edges are the hyperlinks between these pages. As users navigate through this graph of information, they are in *WWW space*. Similarly, the Internet was introduced as a tree like structure of 50 thousand networks containing 16 million WWW servers. As data are transferred from one of these servers to a client, the data are travelling in *Internet space*. Taken together, these two concepts encapsulate the spatial issues of the WWW.

4.5.1 WWW connectivity

The work of Lawrence and Giles (1999) estimated that there were 800 million Web pages contained in the WWW. The HTML documents that define these Web pages were investigated in the information section of this chapter and Table 4.3 illustrated some work that dissected these HTML documents and uncovered statistical averages for the chance of an HTML page containing at least one hyperlink. Woodruff *et al.*, also provided the statistic that out of the 2.6 million HTML documents examined, the mean number of anchor elements per document was fifteen (Woodruff, Aoki, Brewer, Gauthier and Rowe, 1996). These results suggest that the WWW is not only large but also a *highly* connected information space. The exact nature of the connectivity is, however, not clear.

In order to understand the topology of the WWW in more detail, one piece of research approached the problem by constructing a mathematical model and calculating the WWW's diameter (Barabasi, Albert and Jeong, 1999; Albert, Jeong and Barabasi, 1999). Barabasi *et al.*'s model attempted to capture the scale-free characteristics of the WWW by using a decaying power law to model the introduction of new information objects and links. The decaying power law distribution[5] was determined from an experiment that used a '*robot*' to

[5] The power law distribution of the WWW's connectivity used in the Barabasi *et al.* model was independently observed by another group (Huberman and Adamic 1999). Huberman and Adamic predicted that the distribution of the number of Web pages in a Web site followed a power law distribution. They confirmed this result experimentally by analysing two search engine *crawls* of the

recursively follow hyperlinks in HTML documents. The '*unexpected*' result (Barabasi, Albert and Jeong, 1999) was an overall system that, despite the total freedom of the owner of an information object to link to any number of other objects, obeyed the scaling laws of a highly self-organised system. The authors did go on to suggest that their model only captured the rich dynamic structure of the WWW in a minimalist way, but they nevertheless concluded that two randomly chosen WWW pages are on average only nineteen links away from each other[6]. This is an extraordinary result, which indicates that despite its huge size of eight hundred million nodes, the WWW is a highly connected graph with an average diameter of only nineteen. Giles states this property of the WWW more colloquially:

> *"Despite the web's staggeringly large size, any two randomly chosen sites are, on average, no more than 19 clicks apart. This figure gives us an estimate for the diameter of the hazy ball that these sites and links form." (Giles, 1999)*

4.5.2 The Web site

Now that an understanding of the size and connectivity of the WWW has been gained, it is not surprising that a well-documented problem with the WWW (and hypertext in general) is for the user to become "*lost in cyberspace*" (Nielsen, 1998; Mukherjea and Foley, 1995). Navigational aids such as the *Home*, *Back* and *Forward* buttons reduce the risk of the user getting lost (Ayers and Stasko, 1995) although they may introduce their own problems to the user (section 4.5.3). Another method to reduce disorientation in an information space has been the overview diagram or navigational view. By presenting a *map* of the information space, the user is allowed to see where they are, what other information is available and, how to access the other information. Due to its size and high connectivity, visualising the whole WWW would be very difficult. Work has, however, been undertaken to visualise small parts of the WWW that are owned by an individual or single institution and often referred to as *Web sites* (Munzer and Burchard, 1995; Mukherjea and Foley, 1995; Wills, 1999). What these visualisations have in common is that they are all centred on the hyperlinks between Web pages (Figure 4.4).

WWW. The first crawl covered about a quarter of a million Web sites and the second covered about half a million. The probability of drawing at random from these sites, a site with a given number of Web pages was plotted to prove the power law distribution.

6 $\langle l \rangle = 0.35 + 2.06\log(N)$, where l is the smallest number of links one needs to follow to navigate from one page to another, and N is the number of vertices in the graph (8×10^8). Using this $\langle l_{www} \rangle = 18.59$.

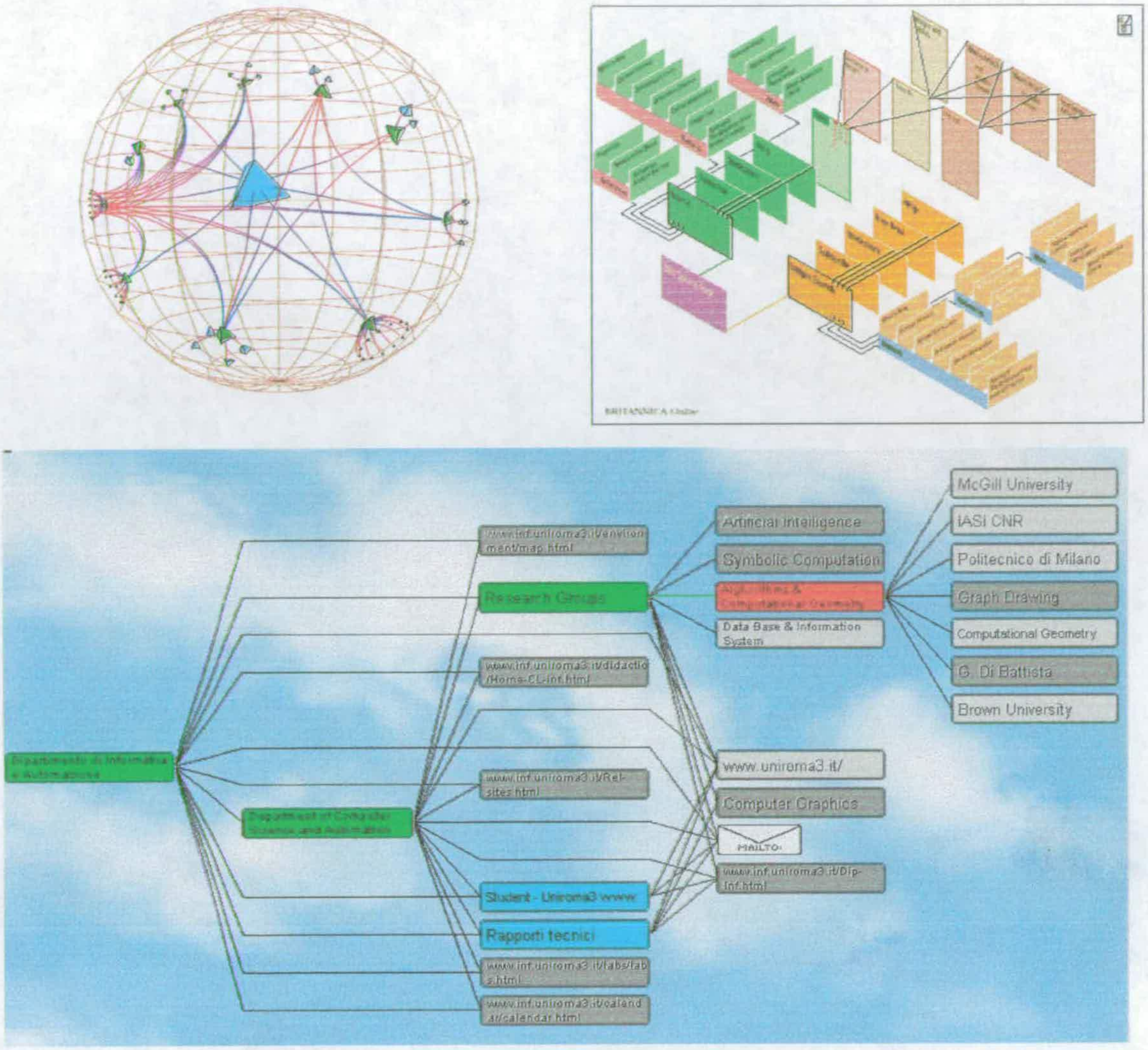


Figure 4.4: Some examples of Web site visualisations; (clockwise from top left) Munzer's visualisation of the link structure of a Web site, using three-dimensional hyperbolic space (Munzer and Burchard, 1995). Dynamic Diagram Inc's (`http://www.dynamicdiagrams.com/`) perspective view of Britannia On-line, Ptolomaeus – The Web cartographer (`http://www.dia.uniroma3.it/~ptolemy/`)

As well as the Web site containing a set of Web pages authored by a single organisation, the set is often stored on the same server. This has implications for the data transfer process when activating hyperlinks within a Web site. There is a high chance that a user will experience a similar temporal behaviour during the data transfer processes when communicating with the same server (all else being equal). The Web site is therefore a very important concept that redefines WWW space by clustering Web pages together. This in turn produces a spatial dichotomy for the hyperlink.

- **Internal** hyperlink – The referencing Web page is located on the same Web site as the referenced resource.
- **External** hyperlink – The referencing Web page is located on a different Web site to the referenced resource.

4.5.3 Spatial navigation aids

In a previous section of this chapter, the findings of two studies showed the high usage of the *Back* button. Although the *Back* button is, in a sense, related to a temporal property of the hypertext – the user's Web page history – it is also related to a hypertext's spatial properties as it retraces a user's route through space. Despite its high usage (claimed to be due to its visual economy by Greenberg and Cockburn, (1999)), studies have shown that the use of the *Back* button does result in a number of usability problems.

Cockburn and Jones (1996) suggested that users have very naive models of how the *Back* button actually works and attempted to prove this with a simple experiment. Current browsers use a stack-based algorithm to implement the *Back* and *Forward* buttons' behaviour. A user can therefore move up and down a *single* hierarchical path captured by the stack but cannot use it to revisit pages that branch off this path. This results in behaviour highlighted by the following scenario:

1. A hyperlink contained in Web page, **A**, is traversed to referenced Web page **B**,
2. The *Back* button is used to go back to Web page **A**,
3. Another hyperlink contained in Web page **A** is traversed to referenced Web Page **C**.

From the present Web page, **C**, can the user return to Web page **B** by just using the *Back* button? This was the critical component of a simple navigational task that 11 computer professionals undertook as part of the Cockburn and Jones study. The result was that eight of the 11 subjects predicted incorrectly that they could indeed, return to Web page **B** using just the *Back* button. They were unaware that Web page **B** had been popped off the stack. Work has been undertaken to develop new algorithms for the *Back* button using different strategies, though all of these strategies present different potential usability problems (Greenberg and Cockburn, 1999).

4.5.4 Linkrot

The final usability problem presented in the spatial domain arises due to the particular hypertext model used by the WWW. Chapter 2 and 3 of this dissertation described the hypertext model employed by the WWW and identified that one of the consequences of this model is that Web pages are variant in time. The objects that make up a Web page can be deleted from a server by their author at any time. The effect that this produces is that hyperlinks, which previously referenced this Web page, will return a HTTP 404 error (Not found) when activated. Similarly, a Web page may be moved to a location either on the same server or onto another server. This movement may mean that the resource's URL changes, leaving hyperlinks that previously referenced it either pointing to nothing or pointing to another information resource. A hyperlink that no longer references a resource is a *dead link* and one that points to a Web page that it was not intended to is simply a wrong link. The Web pages that contain dead links or wrong links have undergone the process of *linkro*t (Nielsen, 1998b).

Linkrot is caused by the movement or deletion of a Web page. These are clearly spatial issues of the hypertext as it is the WWW space that is being changed. Similarly, the effect that linkrot causes is also spatial, as hyperlinks no longer reference a Web page: a connection is broken. Linkrot is, however, a process that occurs to a referencing Web page some time after it has been published. As the name suggests, linkrot is a slow degradation of WWW space and as such, it is related to temporal properties of the hypertext. Gwertzman and Seltzer (1996) collected data from a Boston University Web server to determine the average lifespan of WWW resources in order to optimise a WWW cache. The period for the study was 186 days and Gwertzman and Seltzer found that on average, a HTML document changed every 50 days (the median age of a HTML document being 146 days). They also investigated how images changed over time and concluded that images were less likely to change, having an average lifespan of 85 to 100 days. These results suggest that Web pages do change frequently and linkrot, therefore, takes place on a regular basis.

It is the Web page author's responsibility to ensure that the Web page does not succumb to linkrot. Nielsen suggests that the overall quality of a user's experience is strongly influenced by issues such as dead links. He advises Web page authors to run a hyperlink validator at regular intervals to detect if linkrot has occurred. To ensure that hyperlinks are not subject to rot, Nielsen also provides two rules.

- Keep all old pages on the server forever (unless they are truly misleading and are replaced by an update)
- If a page is moved, leave a redirect behind.

Sullivan conducted a survey to uncover the nature of linkrot by estimating the percentage of Web pages that contained dead links (Sullivan, 1999). The method for creating the sample pool of Web pages was rather unusual. Three word-lists each containing 45 randomly chosen nouns, adjectives and adverbs were used to construct three-word sets. These word triplets were then used as query terms in the Alta Vista (`http://www.altavista.com/`) search engine. After some *'scrubbing'* the results from the search engine where then used as a Web page sample. This sample was pruned to remove all Web pages that contained less than four hyperlinks, which left a 200 pages sample from an original 2250 search results. Sullivan used this sample to define two terms, linkrot *incidence* and linkrot *prevalence.*

- **Linkrot incidence** - the number of dead links as a percentage of the total number of hyperlinks
- **Linkrot prevalence** – the number of Web pages containing at least one dead link

	Survey 1	Survey 2	**Survey 3**
Date	August 1997	May 1998	**May 1999**
Linkrot incidence	3.0%	5.9%	**5.7%**
Linkrot prevalence	18.0%	23.0%	**28.5%**

Table 4.4: Sullivan's linkrot surveys.

The results presented in Table 4.4 show that linkrot prevalence increased steadily over time whereas linkrot incidence tended to be a stable value of just under 6%. Sullivan undertook some simple distribution analysis on dead links and found that the vast majority of surveyed Web pages (71.5%) contained no dead links and two Web pages (1%) contained nothing but dead links. The result of the distribution analysis indicated that 8.7% of the Web pages that contained at least one dead link, presented a user with a probability of over 50% of activating another dead link from the same Web page.

4.5.5 Summary

WWW space has 800 million highly connected Web pages. It is not surprising that users may become lost in this space. Browsers provide a limited set of features, like the *Back* button, to help users to navigate through the space and introduce some temporal context. Unfortunately, the implementation of these features introduce their own usability problems as the route that users take through WWW space is a complex tree structure that cannot be easily retraced. To impose some order onto WWW space and help navigation, a collection of Web pages authored by the same organisation and often stored on the same server are collectively referred to as a Web site. The Web site aggregates WWW space down into manageable chunks that can be regulated by the publishing organisation. The results of this aggregation are some attempts to provide overviews of the Web site space that centre on the connectivity of the information. Finally, WWW space is not static. Web pages are constantly being created, deleted and moved. The upshot of this dynamic behaviour is that hyperlinks can become broken. This process is termed linkrot and its prevalence, worryingly, is on the increase.

4.6 Time

The results of the GVU's WWW user surveys (GVU) reported that the usability problem causing the most concern to users is the download time of Web pages[7]. The actual process of data transfer was presented in Chapter 3 and this chapter showed that the browser struggles to present a complete time affordance for this process due to the abstraction of the Web page. The usability issues that arise from the data transfer process are an important theme throughout this dissertation and some background to this area is presented later in this chapter. In addition to the data transfer process, the next section investigates how a user's Web page history affects their present browsing strategy.

---

[7] Between the 4th and 9th survey respondent citing download as the most important issue have ranged between 64.8% (9th survey) to 80.9% (5th survey).

4.6.1 WWW history navigation mechanisms

The navigational mechanisms that use Web page *history* are similar to facilities offered by earlier small-scale hypertext systems. However, the unique characteristics of the WWW, the most apparent of which being its immense size, have been an incentive for research into new mechanisms (Ayers and Stasko, 1995). If a Web page has been visited previously most browsers inform the user of this by changing the colour or style of the activating object of any hyperlink that references this resource (regardless of whether this particular hyperlink had been used to implicitly activate the request). A user can set preferences to determine for how long a browser should remember this *"already-visited"* information (Tauscher and Greenberg, 1997) and this could mean that the browser has inter-session memory. This is the most basic history mechanism provided by most browsers but there are others that are also based around the list of previously visited Web pages like the *Back* button.

4.6.2 Revisiting Web pages

Tauscher and Greenberg (1997) calculated the probability of any Web page visit being a repeat of a previous visit and concluded that the activity of WWW browsing could be characterised as a recurrent system. A recurrent system is characterised by users predominately repeating activities they have invoked before, while still adding new actions (Greenberg, 1993). Tauscher and Greenberg calculated the *recurrence rate*[8] to be 58% (std=9%) for the 23 subjects of their study and 61% (std=9%) for 55 subjects from the donated results of another study. These results imply that about 40% of all Web pages received by the WWW client were new, with respect to the user's personal history, and around 60% of all Web pages received had been visited previously.

Tauscher and Greenberg extended their study by examining sequences of a user's Web page visitation history to determine how the chance of a revisit declines the '*further*' from the last visit the user is. They calculated a frequency distribution, called the *recurrence distribution,* which plotted the number of times the next Web page visited was a revisit of a Web page in a history list of increasing length. The probability of the new Web page being in that user's history, of course, increased as the history was lengthened to the theoretical maximum of 58% (this being the recurrence rate for the Tauscher and Greenberg study). The data

[8] The recurrence rate R = (total URLs visited – different URLs visited)/total URLs visited as a percentage

revealed that there was an extreme recency in the distribution: there was a 39% chance that the next Web page visited could be matched to the last **six** Web pages in a user's history, a 43% chance for the last ten Web pages, 48% chance for the last twenty and 52% for the last fifty (Figure 4.5).

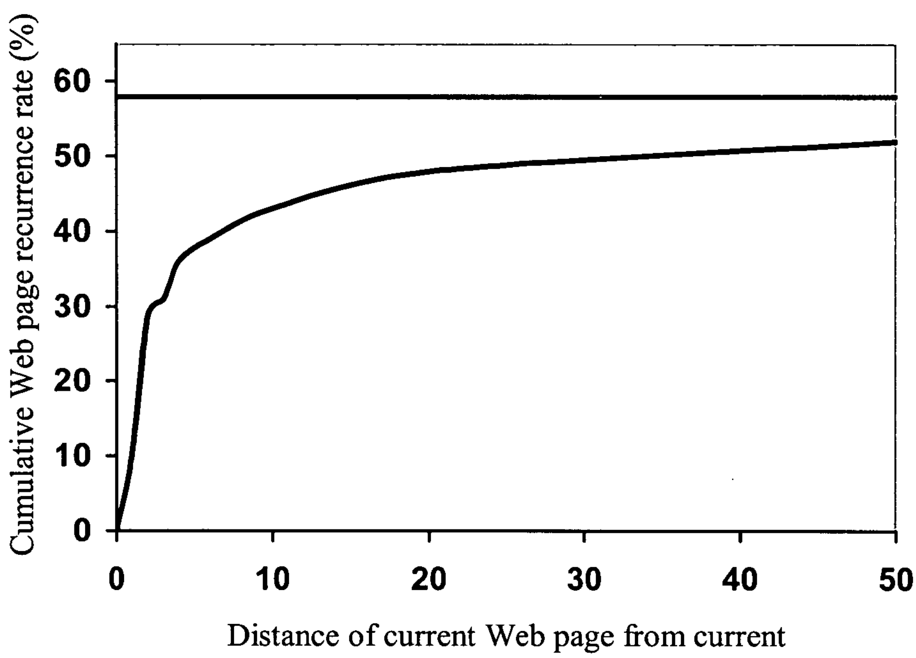


Figure 4.5: Recurrence rate as a running sum of distance

The result that Web page revisitation is so regular a navigational occurrence and that there exists a strong recency in its distribution, is both a result of, and the reason for, WWW browsers containing facilities to view and navigate through a user's Web page history.

4.6.3 Time and the data transfer process

Treurniet, Hearty, and Planas (1985) examined a user's reaction to a simulated teletext system with varied retrieval times of pages. Teletext pages are cyclically broadcast one at a time and the system response time is directly related to the number of pages in the requested cycle. They concluded that a user's '*annoyance*' grew linearly with respect to the square root of the delay in finding the page. This work was subsequently extended (Planas and Treurniet, 1988) to see how different types of feedback would affect this '*annoyance*' rating of users. Planas *et al.* again artificially varied the delay and introduced three types of feedback: an immediate response to confirm that a page request had been made (*immediate system feedback)*; a continuous indication of progress (*continuous system feedback)*; and a message predicting the duration of the delay (*time feedback)*. They concluded that annoyance grew *more slowly* with continuous system feedback, but knowledge of the duration of the impending delay (time feedback) had no extra effect.

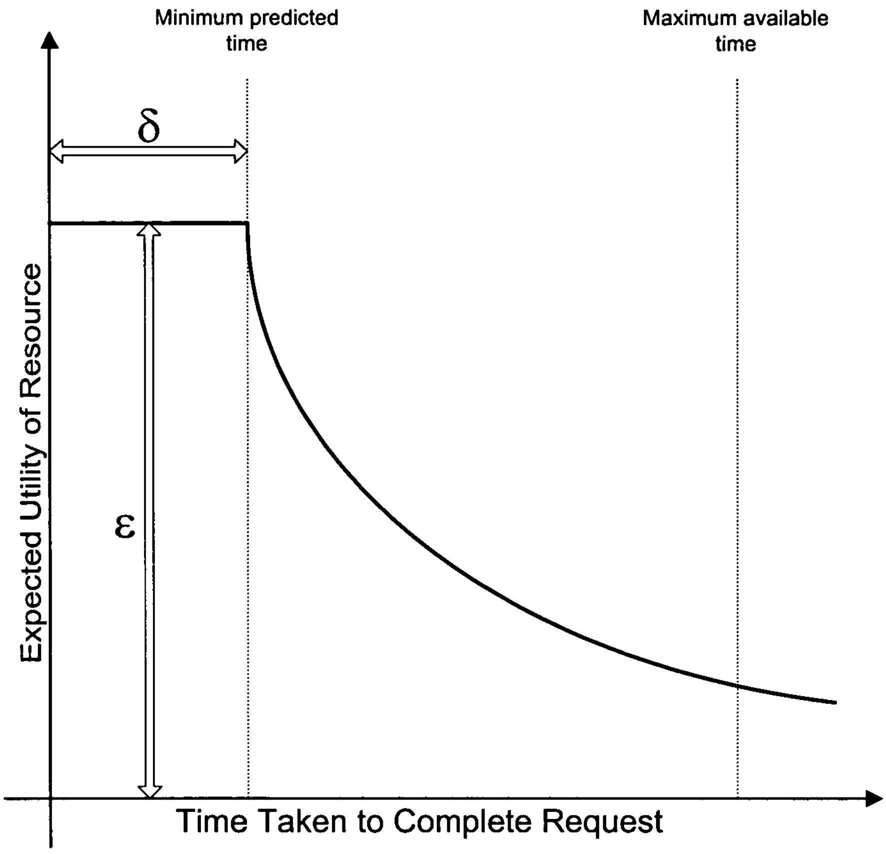


Figure 4.6: Johnson's convex utility curve

Studies have shown not only that users' annoyance increases with respect to retrieval delays (though, as we have seen, it may be constrained by ameliorated feedback), but also that users' perceptions of document quality is related to the length of download times. For example, users' ratings of *how interesting* a document is have been shown to decrease for longer download times (Sears, Jacko and Borella, 1997; Ramsay, Barabasi and Preece, 1998). Johnson and Kavanagah (1996), however, have suggested that users' tolerance of download delays can be influenced by their expectations of document quality. They introduced two terms to highlight the areas of a distributed information system that are the causes of usability problems. *Electronic gridlock* is a term that refers to the situation where networks are unable to cope with the amount of information being requested and *information saturation* refers to the phenomena of users being unable to cope with the information offered to them.

Johnson extended this work (Johnson, 1997) by proposing a theory to describe how a retrieval delay impacts on the value that a particular user associates with the requested information. He applied the concept of marginal utility, derived from the domain of microeconomics, to explain the relationship between the value that a user believes a requested resource may contain and the time that a user is prepared to wait for it. There is a *utility* or capacity for the remote resource to satisfy the user's desire or needs and a relationship of *marginal utility* to define the additional benefits to be gained (or lost) from consuming an extra unit of the goods. In this case, the marginal utility is the additional value of the information to the user that may be obtained by investing a unit of additional time in acquiring it.

Johnson suggests that the cost of downloading a remote resource can be characterised by a convex utility curve as shown in Figure 4.6. Initially, there is an anticipated minimum time within which the user wishes to complete the transfer (δ). Prior to this point, the utility of any information retrieved will be high (ε). After the minimum predicted time, the anticipated value of the information declines. Eventually there comes a point where the user exceeds a maximum time and interrupts the data transfer. The user can no longer 'afford' to wait for the data to be retrieved and hence the utility is very low.

What is interesting is that Johnson believes this cost curve can be manipulated by giving users information about the referenced resource. If users are presented with information concerning the content of the referenced resource prior to hyperlink activation they are in a better position to predict the quality of the resource. The result is that they are better able to appreciate their starting utility for the resource: in other words, how much they desire it. Similarly, the more temporal information provided concerning the data transfer, the more likely a user is to determine whether it is "*worth his while*" downloading it. This extends the period of high relative utility before the minimum expected retrieval period is reached (shown as δ in Figure 4.6).

Although this is a very interesting theory that provides a good argument for describing the temporal properties of the WWW more fully, Johnson does point out that the theory lacks confirmation through experimental results. Furthermore, an individual's attitude towards risk must be assessed before utility can be correctly modelled. Some users may activate many hyperlinks and subsequently abandon them after a short time and others may choose a more rigorous approach by completing more data transfers but consequently activate fewer hyperlinks.

4.7 Summary

This chapter has reviewed some of the work that characterises the WWW and the usability problems that manifest themselves due to the browser presenting too abstract a model of the actual system. The usability problems and characteristics have been categorised into a three-domain framework – information, space, and time - that is used to encapsulate the hypertext abstractions of the WWW alongside the dynamics of the Internet communications. From this perspective, it is clear that the browser was designed to navigate through the WWW using the abstractions of the Web page and the hyperlink and that usability problems encountered appear to revolve around these two abstractions.

First, the behaviour a user experiences after hyperlink activation is a result of the underlying communication of the objects that make up a Web page. The fact that the Web page is not a single entity becomes clear during the data transfer process, and the abstract model of the WWW being a collection of connected Web pages can become lost. Small exceptions have been make by the browser designer to present some information concerning the underlying Internet technology, but the hypothesis put forward in this chapter is that they are insufficient and unpredictable.

The second set of usability problems is related to the concept of relevance between two Web pages. The hyperlink provides a very basic form of relevance abstraction that appears to be deficient. The cost in lost time of retrieving a Web page and subsequently finding it is not relevant is high, and this leads to frustration and annoyance. The problem is that the concept of relevance is situated, since a user in one particular circumstance will find a Web page interesting, but in another circumstance useless. The browser must afford more information about the content of the referenced Web page before hyperlink activation, so that users are able to assess its worth without having to commit to a potentially costly data transfer process.

The next chapter looks at some of the investigations undertaken to show that these problems are real and that the abstractions are the cause of the usability deficiencies. The aim of the investigations is to highlight the areas where users struggle to explain the behaviour they witness, to elicit the real strategies they employ while trying to cope, and show that by presenting a redefined model of behaviour to the user through an enhanced browser, the usability of the WWW can be improved.

Chapter 5

Mental Models

It has been argued in this dissertation that the interaction of the user and the WWW is constrained by the abstractions of the Web page and the hyperlink. The interface conceals the dynamics of the underlying system but the data transfer process is still revealed through the behaviour a user experiences. Insufficient information concerning the dynamic behaviour restricts the improvisations that users can undertake to help them cope with the behaviour of the system. In addition, the hyperlink abstraction presents only one concept of relevance of information and as the cost of verifying relevance is potentially high, users become annoyed and frustrated by unnecessary and unproductive hyperlink activation. This chapter starts to present the results of the investigation undertaken to confirm that the abstractions create a barrier to interaction and also to reveal how users understand and cope with the behaviour they experience, despite the abstractions. The chapter uses a formal approach to interface design to introduce the notion of users' mental models – the internal models that users form of the artefact they interact with. Using this approach, the work undertaken investigates how accounts of these models can be extracted to inform the interface designer and the findings of a preliminary study that extracted accounts of WWW users' mental models are presented.

5.1 Uncovering a user's *mental model*

The term *mental model* has been used extensively in HCI research (for example: Nielsen, 1994; Tauber, 1990; Foss and Deridder, 1987). Different studies observe and reveal different aspects of the interaction between a user and a system. Analyses of users' models are often used as a means of presenting results but can refer to fundamentally different things. Defining exactly what the term mental model means has become an area of some debate. The main area of confusion appears to arise from the use of three different types of model: the model that a user possesses of a system; the model of this model that the experimental psychologist uncovers; and the model that interface designers try to engineer into their systems to help the user (Lansdale and Ormerod, 1995).

The terminology employed in this dissertation is broadly consistent with Donald Norman's original work in the area of mental models (Norman, 1983). When users interact within an environment, they form internal mental models the artefacts they interact with. The users gain experience of the system, acquire knowledge about the system's behaviour and, construct theories about the causality of their actions and the system's responses. The users' mental models provide a means of explanation for observed behaviour and a predictive basis to determine the results of actions. Users will constantly adapt and change their mental model of a system as a result of the behaviours they experience. That is not to say that users' mental models will always evolve consistently and will eventually be complete, they probably will not. The mental model is volatile, with details often being forgotten. The mental model may be inconsistent and non-deterministic with a user often employing a different strategy or providing a different explanation when confronted with the same situation. A mental model often contains idiosyncrasies and superstitions that are revealed as a pattern of unnecessary actions that take physical effort but save cognitive effort.

The work presented in this dissertation has adopted an approach to human computer interaction that was inspired by the ideas on situated action proposed by Suchman (1987). Some care needs to be taken in explaining how mental models can be used in conjunction with the situated action approach, as there is a potential conflict between the two. The improvised nature of action implies that users do not act following a predetermined system of set rules: actions are situated in the local circumstances of the moment. This dissertation does *not*, therefore, consider the mental model to be a formalisation of any set of rules to govern action. Users' mental models are a basic understanding that users employ to guide their actions: a set of heuristics that users try to use to manage a particular situation. The mental model is therefore used in conjunction with the situated information that continually reveals itself during interaction. Users improvise and react in a particular situation using information provided by the interface, which in turn is interpreted using their understanding of the system as a whole; their mental model of the system.

In order to elicit an individual user's mental model psychological experimentation and observation is necessary. The results of such an experiment will present an analysis of the mental model of the user under observation: a model of the mental model[1]. This model of the mental model *cannot* be considered as a process diagram that can somehow be *run* to produce a model of the user. A psychological experiment will only reveal inconsistent volatile and fractious accounts of the cognitive processes involved in using a system.

> *"...as scientists who are interested in studying people's mental models, we must develop appropriate experimental methods and discard our hopes of finding neat, elegant mental models, but instead learn to understand the messy, sloppy, incomplete, and indistinct structures that people actually have." (Norman, 1983)*

Throughout this dissertation, the term mental model refers to the actual model in the user's mind. The term mental model does not refer to the accounts of these models that are uncovered and presented as results of the experiments. It has been argued that the design of the WWW, with its onus on low cognitive load when navigating, short response times and seamless presentation of associated ideas, is not how users consider the system. The work addressed in this, and subsequent chapters, demonstrates this by attempting to uncover accounts of users' mental models, showing that they include ideas and concepts employed in order to explain the dynamic behaviour of the system: the delays they experience and the possibility of failure. The dynamic behaviour may have been abstracted over by the designer's conceptual framework, but it is still experienced when using the browser.

[1] It is these models of mental models that have caused some confusion in the literature. Norman (1983) describes such a model as a researcher's conceptual model of the user's mental model. The use of the words conceptual model are unfortunate as others have employed the term conceptual model to mean one of a number of models that make up a user's overall mental model (for example Augustine and Covert, 1991). Similarly Landsdale and Ormerod (1995) use the term conceptual representation to mean mental models that are system overviews.

5.1.1 Example: extracting mental models of chess players

To illustrate how research has extracted accounts of mental models from users, a short example is presented from a different type of interactive system. In this example, the system under analysis is the game of chess. The question that these studies attempt to answer is how do chess players (users) understand the game of chess through the chess board interface. The approach that these psychological studies illustrate, is how to extract accounts of the chess players' mental models. It is not only the techniques that they employ that will guide the investigations, but also the observation that different experiments can produce different accounts that when combined show interesting and enlightening aspects of the users' mental models.

In 1965 the psychologist and chess master, Adrian de Groot demonstrated what distinguished the chess master from the chess novice by investigating what he termed, the player's *mental models* of the game (de Groot, 1965). Players with differing chess ability were asked to continue a partially completed game of chess[2] and to justify their decisions whilst they played. The conclusions drawn from de Groot's experiment, suggested that there was very little to distinguish the thought processes of the lesser players from the masters and grandmasters. The most consistent difference was that masters and grandmasters only explored strong moves, whereas the weaker players spent time analysing what would turn out to be bad moves. What surprised de Groot and had a significant impact on subsequent computer chess programs (Hartston and Wason, 1983) was that the number of moves that subjects *looked ahead* had no bearing on ability.

The inability of these experiments to highlight any real differences between players' abilities, led de Groot to construct a new experiment that turned out to be the prototype for a large amount of subsequent research in chess thinking. The experiment consisted of exposing subjects to a middle game position for a number of seconds (from 2 to 10) and then, after a period of consolidation (30 seconds), asking the subjects to reconstruct the position on an empty board whilst giving an introspective commentary. Only four subjects were used, one from each of the following categorises of chess ability: grandmaster, master, expert, and a "*representative from the weaker class of player*". What this experiment uncovered was that when players reconstruct a position their accuracy varies as a positive

[2] Called a middle game position.

function of chess skill; the rating that de Groot employed to score the position reconstruction gave the grandmasters the highest scores, the masters next and so on.

The conclusion that de Groot drew was that chess skill could be captured in how a position is mentally understood in the first few seconds of seeing it, and not in a deliberate consideration of how the game might continue. In addition, using the subjects' verbalisations de Groot uncovered that positions were perceived as complexes. The master and the grandmaster perceived the distribution of chess pieces in chunks (Hofstadter, 1979). They produced mental images of the board that exploited the relationships between pieces and they achieved this recognition quickly, whereas the novice simply tried to remember actual piece positions. This was highlighted by the fact that the master's mistakes in reconstructing the position often involved misplacing whole groups of pieces leaving the position strategically similar. In contrast, the novice would simply misplace pieces that were perceived as being less important due to memory loss. The real proof of this theory was that when subjects were exposed to positions where pieces were randomly assigned to squares[3] there was no difference in the reconstructing abilities of all the subjects. The relationships between the pieces were lost and so all the subjects had to rely only on memory[4].

5.1.2 Verbal protocol analysis

The de Groot experiments used a technique called verbal protocol analysis to extract the accounts of users' mental models that helped lead to his findings. The experimental subjects were asked to verbalise their thought processes whilst playing or during a particular experiment. This type of verbal protocol analysis is used to explore the cognitive processes of problem solving and is employed by asking experimental participants to *think-aloud* while working on the problems. It has subsequently been suggested that this technique is unnatural for participants, and the data collected often lacks substance due to the incompleteness of a participant's introspection (Wright and Monk, 1989). As a consequence of some researchers' misgivings for this original think-aloud technique, a variety of different verbal protocol analysis methods have been introduced and subsequently used by psychologists.

---

[3] Not a real chess position at all.

[4] Chase and Simon (1973) copied the de Groot position reconstruction experiment but also measured the number of glances the players took of the board. They confirmed de Groot's findings and added that, whereas grandmasters could construct the position of real middle game positions in one or two glances, when presented with a random position they took as many as twenty glances.

In an attempt to consolidate this field of research, Wright and Monk (1989) reviewed the various methods employed by HCI researchers and suggested a framework into which different verbal protocol analysis methods could be categorised. The Wright and Monk framework distinguished between three distinct types of method.

- Constructive interaction.
- Co-operative evaluation.
- Re-enactment.

Constructive interaction[5] is a technique that collects data using participants' interaction with each other. Two users work together on a problem and the verbal protocol analysis is performed on their conversation. This method is an attempt to increase the '*naturalness*' of the participants' utterances and has a basis in ethnomethodology as it is the participants' social interaction with each other that is observed and recorded.

A similar method to constructive interaction is a question and answer technique where a participant is encouraged to think of themselves as a co-evaluator of the system under study and to work on equal terms with the experimenter (or system designer). While the participant is actually using the system, the conversation between the two evaluators revolves around the experimenter being deliberately obtuse by not answering questions directly and instead, directing questions to the participant concerning their understanding of the actions offered by the system and the observable feedback. This technique is termed co-operative evaluation (Wright and Monk, 1989; 1991).

Both constructive interaction and co-operative evaluation are concurrent verbal protocols. The verbal utterances of the participants and experimenter are produced at the same time as the system is actually being used. The fact that a participant must talk whilst using the system necessarily transforms the working practice of the participant. Although constructive interaction tries to minimise this disruption to the observed setting by introducing another *natural* social actor, the environment is, nevertheless, being artificially manipulated.

The final group of techniques categorised by Wright and Monk remove the concurrency between the use of the system and the verbalisation of the participant by retrospectively interviewing the participant. An experimenter and participant re-construct some portion of a

---

[5] The term constructive interaction was actually coined by Miyake (1986).

recorded trial using either system logs or audio-visual recordings, and the participant explains their actions with respect to their desired goals. This final technique is understandably termed re-enactment.

Verbal protocols are employed to extract an account of the user's understanding of a system: the relationship between the user's actions and their predicted system response. During any particular experiment, however, the use of a verbal protocol analysis will only extract a single account of this understanding or mental model. In another situation, the same user may describe their understanding differently. This is the nature of the mental model and the understanding that users employ when using a system. There also exists another problem when using methods that directly ask users to explain their actions. Norman terms this problem the *demand structure* of a situation.

> *"If you ask people why or how they have done something, they are apt to feel compelled to give a reason, even if they did not have one prior to your question. They will tell you what they believe you want to hear. Having then generated a reason for you, they may then believe it themselves, even though it was simply generated on the spot as an answer." (Norman, 1983)*

Extracting accounts of mental models is a difficult process. The accounts extracted through verbal protocol analysis can only be considered as a model of the understanding present in a user's mental model. Care must be taken not to confuse these accounts of mental models for the actual mental models that users possess.

5.1.3 Returning to the chess player

Verbal protocol analysis is only one technique used to extract accounts of users' mental models. Another technique involves experimental subjects attempting to draw a picture of their mental representation, which they have previously verbalised. One such study that used this method was also performed on a chess master in an attempt to uncover an account of a mental model. To confirm a hypothesis that chess masters conceptualise the game and do not therefore need to visualise the board or pieces in any visual detail like colour or form, Binet (1894) performed an experiment that required a chess master to draw a picture of his mental representation of a particular chess position. The picture and the chess position are shown in Figure 5.1. Binet suggested that this picture does indeed show that the chess master's mental model does not rely on visual detail. The squares have no colour and only a few are shown. The pieces themselves are not present on the drawing, only the lines of force that they control and the squares they potentially influence[6].

Although the picture shown in Figure 5.1 is simply one account that would make up Binet's conceptual model of the chess master's mental model, it is enlightening. Although Binet did not try the same experiment on the novice chess player, he postulated that the difference would be marked. The work of de Groot would imply that the novice player would conceptualise the position with coloured squares and pieces and would miss some of the more subtle relationships between the pieces. Whether one can say that the picture is really how the chess master actually perceives a position is not important. It simply provides another account of the player's understanding of the system highlighting some understanding and abstracting over others.

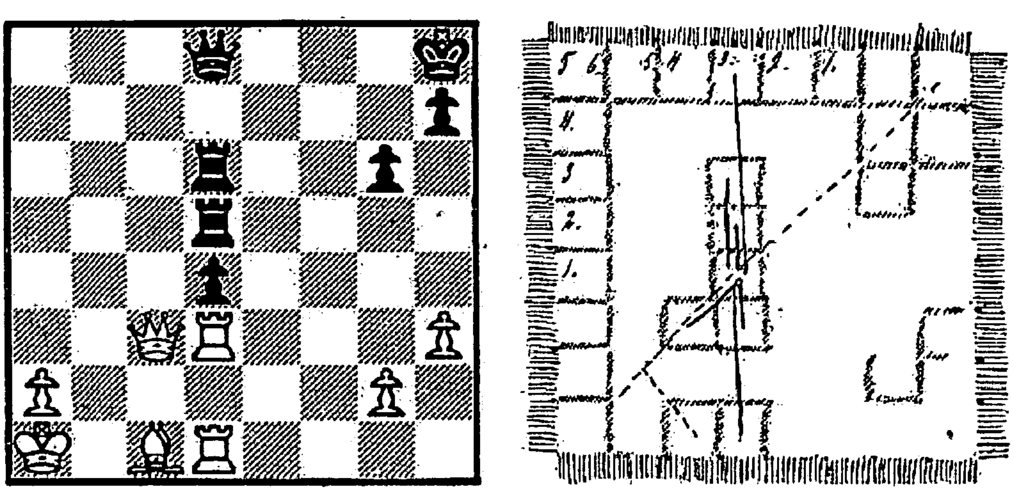


Figure 5.1: Pictorial account of chess master's mental model of chess position shown on the left. (Binet, 1894)

To summarise, this presentation of the work undertaken to extract accounts of chess players' mental models has included a number of techniques. Taken together, a picture of a chess player's understanding of the game can be formed and in particular the differences between the novice player and the experienced (and gifted) grandmaster highlighted. Binet showed that the chess master does not use the visual properties of the chessboard to conceptualise a

6 The numbers indicate the number of moves that the King would need to make to catch the white pass pawn on square a2 (bottom left corner).

position. De Groot confirmed this and suggested that the difference between the good and bad chess player was evident in their ability to reconstruct a position. The good player understood the relationships between pieces and could quickly remember chunks of information whereas the bad player simply tried to remember the pieces and their positions on the board.

What this highlights is the difference between the expert and novice users' mental models, and their understanding of the game of chess. All the communication between the system and the user takes place through the system image, which in this case is the chess board. To make the game of chess easier for the novice, the system image should present a better representation of the game. The understanding gained from de Groot and Binets' experiments imply that a better interface would present the relationships between pieces, the squares they influence, and the lines of force they produce rather than the pieces. If this new interface could be produced, any player could construct a mental model of the position accurately and consistently. Any player could be coaxed into seeing the important chunks and avoid the unnecessary analysis of a bad move. The game of chess is intellectually challenging because it is a complex process to decipher the large number of important relationships from the visual image of the board position. To make chess a good game the chessboard and the pieces must be a *bad* interface.

Producing a new interface for chess, although very interesting, is outside the remit of this dissertation. What it illustrates, however, is the process of using accounts of users' mental models to inform the change of the system image and increase usability. Chess is a highly consistent game with strict rules and deterministic outcomes and it is presented to the user using a static interface. The WWW is dynamic, inconsistent, and unpredictable. Despite this marked difference, a similar design approach can be taken with the WWW, which is to extract accounts of users' mental models and use them to build an understanding of the problems and difficulties that users face. This is the initial aim of the investigation into WWW use that follows. This chapter presents the accounts of users' mental models using a visual approach in accordance with the work of Binet, and the following two chapters employ the verbal protocol techniques in the manner of the work of de Groot.

5.2 Visual representations of the WWW

The objective of the work now presented in this study was to extract accounts of users' mental models of the WWW. The study involved asking a set of subjects to draw a picture to explain their understanding of the WWW. By contrasting the different visualisations produced by subjects with varying degrees of expertise, the study aimed to show what concepts subjects acquired of the WWW as they gained experience. The form and value of the data collected using this visual approach were, however, completely unknown. It was even suggested that the task could be impossible to complete for some subjects, since a visual explanation might be non-existent or very unnatural to describe.

5.2.1 Subjects

Subjects were unpaid volunteers, recruited from under-graduate and post-graduate students at the University of Edinburgh. Subjects were asked to classify their computer expertise and in particular, their experience of using the WWW by selecting the category that best represented them: expert, intermediate, novice, or virgin. Thirteen subjects were recruited for this study with a mixed range of expertise: four experts, three intermediates, five novices, and one virgin user.

5.2.2 Data collection

At the end of a browsing session, each subject was asked the following question:

> *"Could you draw a picture of how you perceive the World-Wide Web. Considering the present Web page as a starting point, could you draw a picture of its relationship with other Web pages and what you consider to be the important differences between them."*

Although not requested to do so, subjects often provided a concurrent verbalisation of their drawings in explanation, and although these utterances were not formally recorded, they are used in the presentation of the findings. Only two subjects were unable to draw anything, claiming not to understand the question. Both of these were novice users.

5.2.3 Results

Figure 5.2 shows a set of four drawings produced by three novice subjects and one who had never used the WWW before the experiment (Drawing IV). The similarity between the drawings is striking: each picture presents a graph with Web pages represented as the vertices and hyperlinks as the edges. From their very limited experience, these novice users had already been able to understand the nature of hypertext. They show an understanding of the Web page as a single entity and the concept of the hyperlink connecting Web pages together.

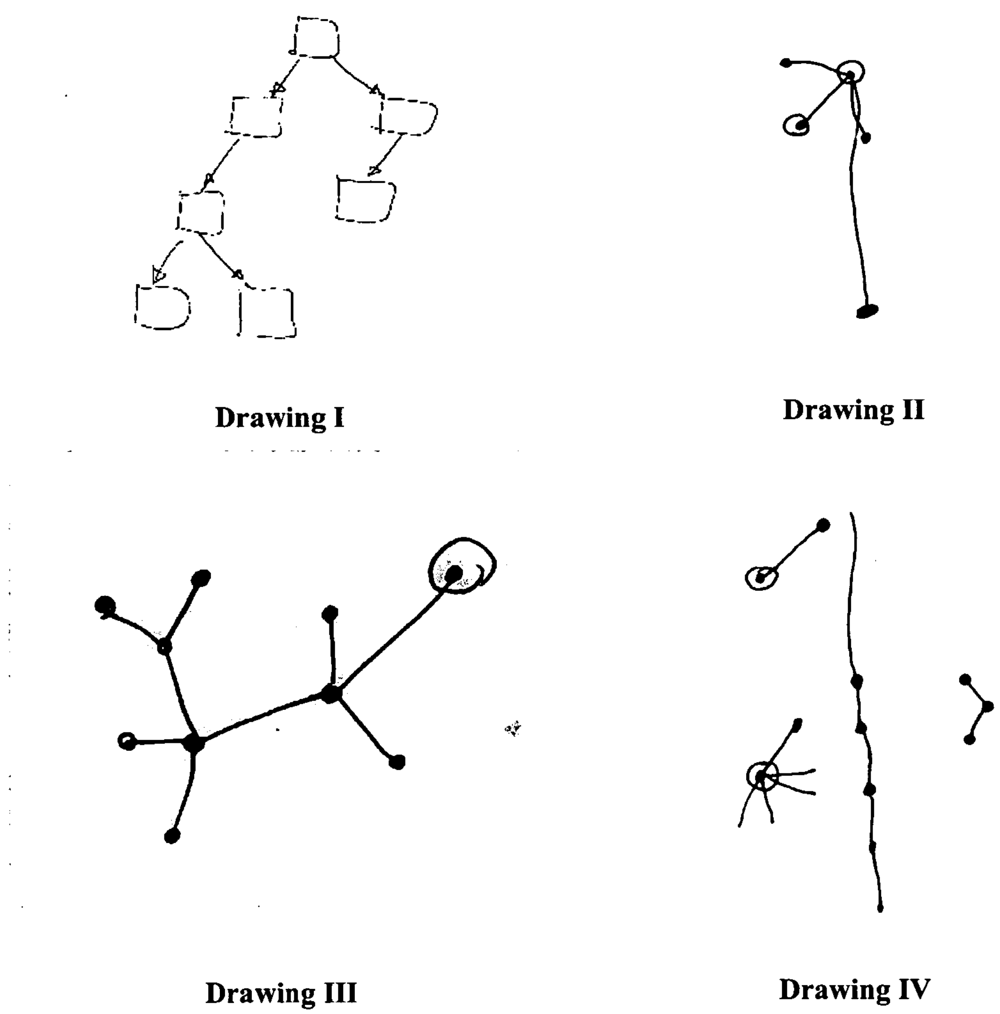


Figure 5.2: Three pictures of the WWW produced by novice subjects, and one produced by a first-time WWW user (Drawing IV).

Drawing I extends the simple graph model by employing directed edges between vertices, correctly suggesting that the hyperlink is unidirectional. The regular tree-like structure of the graph is also worth noting as it suggests uniformity. In fact, the subject explained that although the WWW was large, each time they actually browsed they were only interested in a very small number of Web pages that were connected together. This demonstrates a basic understanding of relevance introduced by the hyperlink. If the user could not get to a Web page by activating a hyperlink, it was not relevant to their present browsing session.

Drawing II, III and IV are all very similar showing the simple graph with vertices as Web pages and edges as hyperlinks. All three of the drawings have circles around one or more vertices, and these denoted the movement of the user through the hypertext – *i.e.*, "*I am here, and then I move to here*". This was another common feature of the subjects' pictures and explanations; they all described browsing as a process of *moving through* the hypertext. The implications of describing hyperlink activation as moving the user to the Web page are that any delays experienced after hyperlink activation are often attributed to problem with the user's browser or machine. There is no understanding of data being transferred from a remote computer to the browser and so the blame for a disturbance or failure must lie with the user.

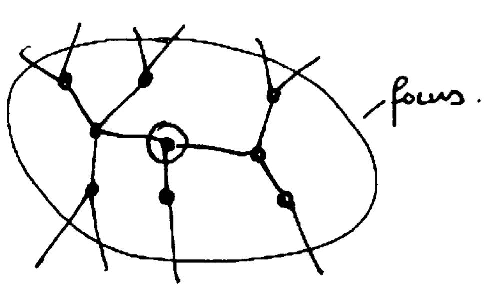
An example of this misunderstanding of the data transfer process is when users explain that the Web page is taking a long time to appear because *they* have a slow computer: the implication is that with a faster computer – *i.e.*, faster processor or more memory - the disturbance would not happen. Although there is actually some truth in this understanding - the user's computer must render the Web page and execute any processes embedded in the Web page, so a faster computer may decrease the total time between hyperlink activation and rendering - in general it is not the local processing that is the major factor causing a delay: it is the data transfer across the network. In subsequent chapters, there are a number of examples of novice users trying to explain a delay using the argument that *their* computer is slow. These explanations all centre on the user possessing a mental model that interprets hyperlink activation as the user moving through the WWW space, rather than the data being sent to the user.

One slight difference between the drawings that is worth further explanation is the existence of islands within the graph of drawing IV. These islands suggest a form of grouping that could be concerned with either WWW space – some areas of the Web are not connected to others – or information – showing relevance, like Web sites. When asked about this feature of the drawing the subject commented that in their browsing session they had used a search engine and systematically gone through the list of results. Each island represented moving to a different result[7].

To summarise, the novice subjects' accounts all present the WWW as a hypertext in a noticeably similar fashion. Web pages are single entities connected by hyperlinks. In addition, the novice subjects explained hyperlink activation as them moving through the hypertext from the present Web page to the new Web page. Novice subjects understand WWW space.

5.2.3.1 Intermediate subjects

Although classified as having intermediate experience, the subject whose drawing is shown in Figure 5.3, has a very similar understanding as the novice users. The hypertext model is again clearly shown and the only element that has been added is a grouping of vertices and a concept entitled *focus*. The subject stated that the focus was a particular set of Web pages relevant to the present topic. Each Web page was easily accessible using either a hyperlink of a navigational aid like the *Back* button. Although this drawing includes a circle to denote the present Web page, the subject stated that the graph moved under this circle rather than the circle moving around the graph.



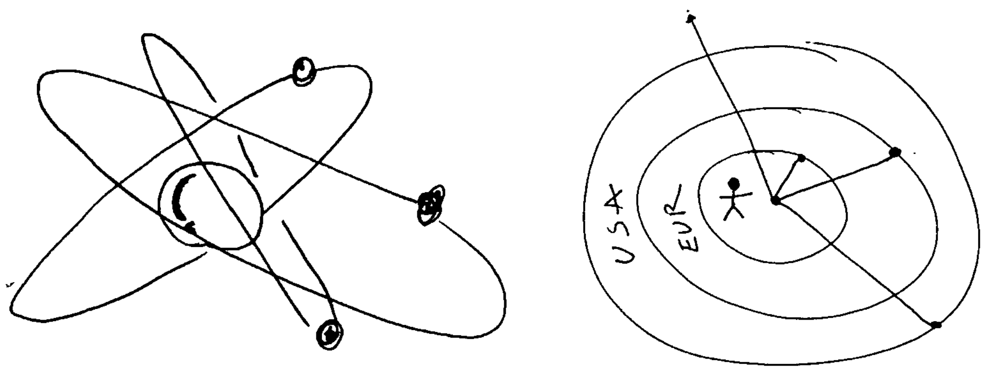
Drawing V

Figure 5.3: Intermediate subject

[7] Another difference is that drawing II has one edge longer than the others. When asked to explain the significance of this the subject stated that it had no meaning.

This is the first major difference between the intermediate subject and the novices: the Web pages come to the user rather than the user moving around the Web pages. Other than this significant point, however, drawing V is very similar to the novice drawings.

Figure 5.4 shows two drawings produced by subjects who rated their experience in the intermediate category. The drawings from these more experienced subjects start to diverge from the simple graph representations produced by the novice subjects. The features that the intermediate subjects introduced into their drawings that were missing from the novice subjects were the existence of variable download time and the concept of Web pages coming to the user (rather than users moving through the hypertext).



Drawing VI **Drawing VII**

Figure 5.4: Pictures of the WWW from two intermediate subjects.

Drawing VI is rather unusual in comparison to **all** the other drawings as it employs a three dimensional solar system metaphor to describe the WWW. Planets represent Web pages and the Sun represents the user. The concept of download time is captured by how far the Web page is from the user – Web pages have orbits with different distances from the Sun suggesting a different download time. In addition, the Web pages are constantly moving, orbiting the user, showing that there is continual change to the download time. Not only does this drawing capture the unpredictability in download times between different Web pages, but it also captures the inconsistency of the download time for a single Web page.

Drawing VII, although less imaginative, shows a similar enhancement to the account in comparison to the novice users. The simple graph used in the novice subjects' drawings is superimposed over a new spatial dimension. The subject that drew this picture implied that different Web pages (the vertices on the graph) were located in different countries and this had a direct effect on the download latency. Drawing VII shows the continent of Europe (labelled EUR) and the United States (USA) as concentric areas around the subject (the stick figure in the centre). A Web page in the USA is further away from the subject than a Web page in Europe, and so it will take longer to download. The length of the edge therefore represents the download time of the Web page – vertices further away from the user take longer to download.

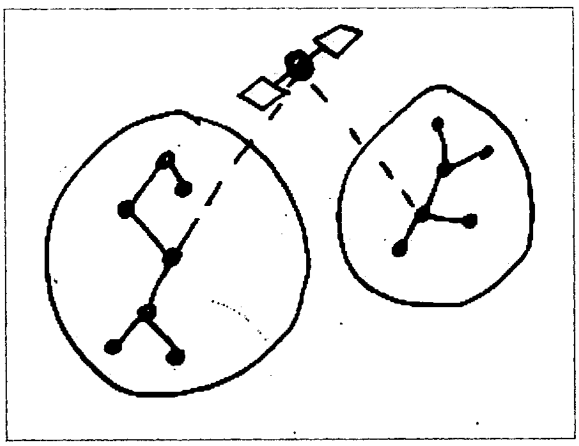
The intermediate subjects no longer view the WWW as the simple graph of the hypertext model. The dynamic behaviour they have constantly experienced while using the WWW has meant that their mental models have had to cope with a new temporal dimension. This has forced the subjects to introduce the new concept of time into their understanding. The most noticeable changes between the novice subjects and intermediate subjects are:

- the concept of Web pages coming to the user
- the unpredictable difference between Web page download times
- the inconsistent download time of the same Web page
- and, (in the case of drawing VII) the physical location of Web page directly affecting download time.

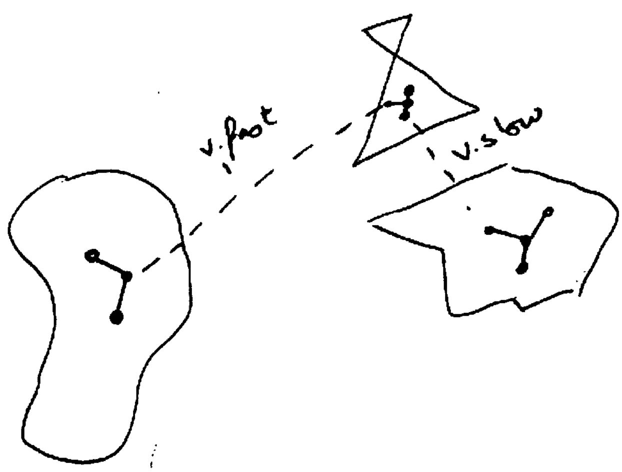
5.2.3.2 Expert subjects

Figures 5.5 and 5.6 show the drawings of three expert subjects. The idea of a graph showing the connected Web pages is not totally lost, but even more so than the intermediate subjects, these drawings show an appreciation of physical location of data. All three of these drawings show geographic land masses centred on the United Kingdom. The download time of a Web page is affected by the continent or country where its data is stored. The physical location is, however, not the only factor that affects download time and the expert subjects introduce the idea of different international connections having different data transfer properties.

Drawing VIII centres on the idea of a special connection between two landmasses. The subject stated that the connection between the island on the left (the USA) and the island on the right (UK) was slow in comparison to internal connections. If the subject knew what country the Web page was located in before hyperlink activation, the subject claimed to be able to predict how long it would take to download using this simple model. Downloading a Web page from America is very slow because the connection between the countries is slow. This is represented on drawing VIII by the satellite image connecting the landmasses. Although the subject knew that the data transfer between the countries did not actually use a satellite connection, the subject stated that this was how he *'visualised'* the connection. Despite the drawing displaying temporal and spatial elements, it has not totally lost the hypertext model. The hypertext model still exists in the form of simple graphs within each landmass, but they are supplemented by properties related to physical space and data transfer.



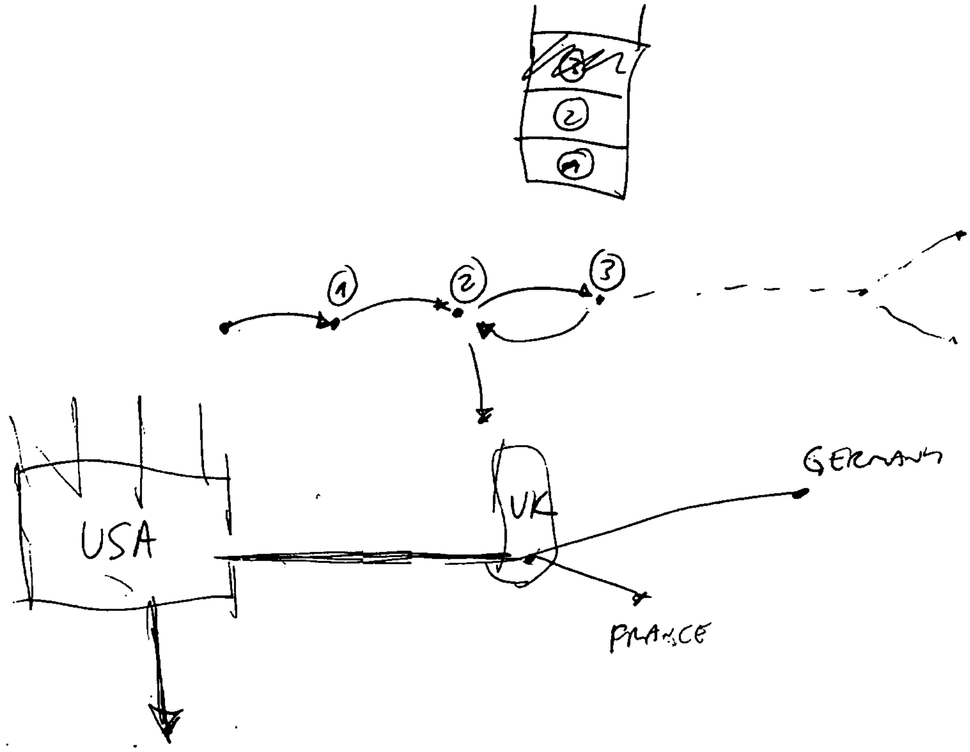
Drawing VIII



Drawing IX

Figure 5.5: Drawings from two expert subjects

In a very similar fashion to drawing VIII, drawing IX introduces the same concept of physical geography but also includes Europe (or perhaps just France). What is interesting is that the important temporal issue that this subject highlights is that the connection between the UK and the USA is actually fast in comparison to the connection between the UK and Europe. Again, the hypertext model is not totally lost as simple graphs appear inside each landmass. The advance in understanding that these drawings display over the intermediate subjects is that download times are not simply proportional to physical distance they are also, in some way, related to properties of the connection to these countries: *i.e.*, the properties of the Internet.



Drawing X

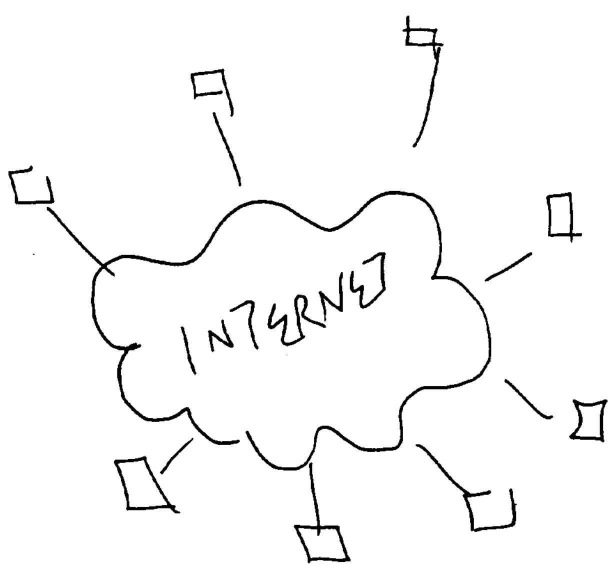
Figure 5.6: Drawing from an expert subject

Although drawing X presents the same country based geographic dimension as the other expert subjects, it separates this model from the hypertext model. The numbers and directed graph that appear on the top of the drawing represent the movement between Web pages using the navigation aids. This subject understands the stack based nature of the *Back* and *Forward* buttons and shows the popping of the stack (the crossed out number 3 at the top of the drawing) and the movement back through the graph. The interesting point is the separation of the two models. There is no attempt to link the hypertext graph with the geographic map of connections. There is a model of hypertext, useful in some circumstances, and a geographic model showing where data is located, useful in another circumstance.

5.2.3.3 Results summary

What these drawings start to reveal is how users' understanding of the WWW changes as they gain experience of using the WWW. Initially users understand the WWW as Web pages connected together by hyperlinks, but as they continue to witness the temporal behaviour of the system they gain experience and start to try to explain the underlying behaviour of the Internet. A realisation that there is a physical location associated with Web pages and that this may affect the retrieval latency of the Web page, clearly starts to become a key ingredient of a user's model of the system.

For completeness, Figure 5.7 shows one more drawing from an expert subject. This is the only drawing that does not fit so elegantly into this presentation as it does not show the same dependence on geographical location as the others. This drawing was produced by an expert subject who outlined the workings of Internet technology in some detail.



Drawing XI

Figure 5.7: Pictures of the WWW by an expert subject

The real difference between this account and what others described is that this is the only account that shows any understanding that the Web page is made up of different objects. That is not to say that the other subjects did not know this, only that their accounts presented the Web page as a single entity. It was the geographical location of the Web page that the other subjects were using in their explanations and not the geographical locations of the objects that made up the Web page. This is a very significant observation, since even expert users who know how the technology works still conceptualised the Web page as a single entity. Although the subjects attempted to circumvent the hyperlink abstraction and the seamless and rapid movement through the hypertext, with the exception of the subject who drew drawing XI, all the subjects either could not, or did not need to, circumvent the Web page abstraction. From the novice through to the expert, all accounts used the Web page abstraction.

The reason why the acceptance of the Web page abstraction is such a fundamental point, is that the behaviour of the data transfer process is dependent on the data objects that make up the Web page. As was shown in Chapter 3, each object is potentially stored in a different location and transferred across the network separately. To present the full time affordance for hyperlink activation, breaking down the Web page abstraction is essential.

5.3 Summary

This chapter introduced the design principle that the interface attempts to present the designer's conceptual model of the system to the user. If this conceptual model is appropriate to describe the functionality and behaviour of the system and the interface successfully presents this model to the user, the user will form their own mental model in accordance with the conceptual model and be able easily to use the system. There are two areas, therefore, where usability can be compromised: the interface not successfully presenting the designer's conceptual model to the user, and the conceptual model not explaining the system behaviour appropriately.

It appears that the designers of browsers have employed the simple hypertext model as the WWW's conceptual model. The accounts of users' mental models that have been presented in this chapter show that the model of hypertext has been presented and accepted by users through the browser. The more experienced that users become, however, the more their understanding must explain the dynamic properties of the data transfer process that they witness continually. This primarily manifests itself with a shift from the idea of navigating through a hypertext seamlessly, by moving between associated information, to the concept of data being transferred to the user and the delay that this may entail.

The models employed by the more experienced users to explain the data transfer process appear to centre on the geographical location of the data. This suggests that the browser affords users this geographical information, despite the browser being designed to present the hypertext model. It would appear that users can easily elicit this spatial information in comparison to the temporal properties of the Internet. Having uncovered the geographical location of the Web page data, the user is then able to use this as a guide for determining the temporal delay they experience. Making an accurate prediction about the actual temporal behaviour of the Internet would appear to be outside the understanding that users possess, but by using knowledge of the data's spatial location – that apparently is easy to elicit – and a simple set of heuristics describing the relationship between space and time, a form of understanding can be generated about the temporal properties of the data transfer process. Users extract spatial information and then use this as a heuristic for explaining delays, in preference to predicting the actual data transfer behaviour of the Internet. Subsequent work will show how users actually extract this spatial information and why it is easier to elicit than an accurate delay prediction.

The results of this simple investigation lead to the conclusion that usability is compromised due to the inappropriate conceptual model of the WWW. The simple abstraction of a hypertext is often not adequate when explaining the behaviour of the WWW. The course of action that must be followed to remedy this problem is to redefine the conceptual model of the WWW and construct a new interface to present it. The next chapter sets out a further course of study to reveal more accounts of users' understandings: to uncover the actual problems and the situations that users find themselves in during WWW use. These studies use more rigorous methods than the simple drawing technique employed so far, but their findings confirm the general principle set out here. The hypertext model does not explain the behaviour of the WWW sufficiently. Users often find themselves in situations where they must first, explain the undesirable dynamic behaviour they witness and second, employ a course of action to curtail any disturbance.

Chapter 6

The WWW User

Chapter 5 introduced the reader to the concept of the user's mental model and presented a short investigation that produced pictorial accounts of how users understand the information, space and time dimensions of the WWW. Although the findings of this investigation were enlightening, they only provided general overviews of the understanding that users possessed. In this chapter, and subsequently in Chapter 7, the reader's attention is focussed onto the more acute and situated empirical investigations that were undertaken to reveal and justify this thesis further. In order to uncover the subtleties of human-computer interaction, a range of qualitative and quantitative experiments was conducted: both situated fieldwork and controlled laboratory experiments. This mixture of experimentation led to the experimental perspective exercised becoming an important cause for concern. It is dealt with next where a short discussion is presented to explain and clarify the position. This is followed by the presentation of the method and the results of the second investigation, which aimed to generate an initial focus for the research; to reveal the salient issues of the interaction between the WWW and its users so that subsequent investigations could become progressively more directed. This more directed study is described in Chapter 7.

6.1 Prior set of concerns

The core concern of HCI research has been the design of the interface, rather than the design of experiments to study interaction with the interface. The research concerned with experimental design has been left in the hands of psychologists and cognitive scientists. The realisation that research into HCI needs to be inter-disciplinary is nothing new, but as recently as 1998, Gray and Salzman voiced concern about the validity of the experimental methods employed by HCI researchers. They examined the design of five experiments that had had an important influence on HCI thought and practice and proposed that, due to problems with the experimental methods used, serious questions could be asked about the validity of the results (Gray and Salzman, 1998).

The problems that Gray and Salzman revealed concerned the *validity* of experiments. Does the experiment measure what it claims to measure, and are the conclusions drawn from a manipulation of these claims valid? Are the observations of the experiment a direct consequence of the proposed cause, and can the claims be generalised across settings and persons? The particular needs of the HCI researcher, mean that the experiments used to reveal aspects of the interaction must be designed to address concerns regarding causation and the validity of the experimental results.

6.1.1 Causation

What an experiment is designed to uncover is the relationship between variables. Does A cause B, or does another variable C causes B, or does C cause A and B? In the domain of HCI, for example, how can an experiment prove that a certain affordance presented by the interface causes a user to interact with the system more easily than if the cause, the affordance, is not present? What is it that the affordance has actually affected? Does the experiment discover whether it is the affordance that has improved the usability of the system, or is there another unknown cause that the experiment has not isolated? Even if one can isolate the relationship between cause and effect, it may not be a simple matter to uncover the causal relationship. The causal relationship itself may have properties that need to be determined.

- If an effect is observed if, and only if, a cause is present, then the cause is **necessary and sufficient**.
- If an effect is sometimes observed when a cause is present, but never present when the cause is absent, then the cause is **necessary but not sufficient**.
- If the effect is always observed when the cause is present, but sometimes observed when the cause is absent, then the cause is **sufficient but not necessary**.

Returning to the human-computer interaction example, the presence of an affordance in one instance may have an observed effect on some aspect of how easy the system is to use, but in another instance, with everything else being *equal*, it may not. Is the causal relationship between the affordance and the ease of use necessary but not sufficient, or is the experiment's control on the equality of two situations flawed and is another variable causing the observed effect? Rather than simply verifying an experimenter's hypothesis, the experiment can, at best, only rule it out if it is proven false (Cook and Campbell, 1979). In reality, this means that the validitation of a hypothesis is at best an approximation of its truth or falsity.

6.1.2 Experimental validity

To increase the confidence in a hypothesis, the experimental method must be valid. The validity of the experimental method can be divided into two categories.

- **Internal validity** - the effects that are observed are caused by the variables under observation and not another unknown variable.
- **External validity** - the cause can be generalised across settings, different types of people and, at different times.

These two classifications can be divided up further into more specific areas of experimental validity (the interested reader is directed to Cook and Campbell, 1979, and Gray and Salzman, 1998). For the purposes of this discussion, however, these two general types of validity are enough to highlight the general areas of concern.

It is already clear that the WWW user is affected by many variables delivered through the WWW browser and from the user's previous domain knowledge. Isolating the particular variables and proving they cause an observed effect on usability, involves a lot of control of the environment. If this control is not appropriate, the internal validity of the experiment is potentially an area of concern. The existence of a controlled environment will, however, have an impact on the external validity of the experiment. If a cause can be isolated and shown to affect the user, the results can only be generalised through this controlled setting. There is a conflict between isolating causes and effects by controlling the conditions, and observing the natural environment.

> *"The more open the system, the more fallible will be causal inferences"*
> *(Cook and Campbell, 1979, p 35)*

Here then is the dilemma. Either control the environment and shield the participants from the very causes that shape their interaction with the system, or study the participant in their natural environment, but relinquish any control over the variables one wishes to uncover.

6.1.3 Case study approach

A case study is a research method used to investigate both a causal relationship and the context within which it is occurring. A case study needs to be used if the causes and effects cannot be distinguished from the context, or if the context itself is believed to contain important variables (Yin, 1993). These contextual or environmental variables may be numerous and so different that no single data capture method may be able to elicit them. A case study is, therefore, a method used to study an environment where not all the variables can be observed at once. Despite this lack of control over the variables, the case study collects data from a multitude of sources and uses this data as evidence to explain a restricted set of effects.

The case study is an approach that can be characterised as fulfilling the following statements. First, the case study defines specific questions or hypothesis *before* the period of observational study. The observational study then objectively collects empirical data, and from this data alone hypotheses are validated and conclusions are derived[1]. Second, a case study carries out directed observational studies, targeting on variables that are perceived as relevant and influential to the proposed hypotheses.

---

[1] This is termed Logical Positivism and is a philosophical school of thought that traditionally formed the foundation of the natural sciences.

6.1.4 Ethnography

Ethnography is a process of sustained participant observation in the *social* setting under study. The basic idea is that the observer joins other participants and works for a prolonged period as a normal social member of the group. The observer immerses himself in the society to experience the social interaction first-hand from a participant's point of view. An ethnographer must bring no assumptions to the study and refrain from controlling the environment in any way (Jirotka and Goguen, 1994). An ethnographic approach is more than just a style of quantitative fieldwork; it also implies a method for presenting the findings. The focus of ethnography is on uncovering participants' experiences rather than simply reporting their observable actions.

In contrast to the case study approach, the ethnographic approach does not allow the researcher to enter the field with predetermined questions and hypotheses. The ethnographic approach relies on the ethnographer experiencing the world from a participant's perspective and producing a detailed, or *thick*[2], description of the environmental interactions. The ethnography uncovers the natural strategies people employ to make sense of the situation they find themselves in and attempts to elicit the models that people use to understand the behaviour of the people and things they interact with. Finally, and again in contrast to the case study approach, ethnography encourages unstructured and long-term observational study. The natural actions and strategies must be allowed to surface as they happen, and participants must not be forced to perform. The result of following these characteristics is that ethnographies tend to produce voluminous and discursive descriptions that are necessarily interpretative, since the ethnographer must determine the significance of actions without the aid of large amounts of statistical data.

Traditional ethnography as practised by social anthropologists has often been considered as having too much of a culture bias, being too holistic and too time-consuming for studying information systems. Immersing an ethnographer in the field for years at a time might help in uncovering the subtleties of foreign cultures, but in computer science (and in particular the WWW) systems change so rapidly that this approach will not only be investigating a constantly changing setting, but may eventually produce results that are no longer relevant to system design. Analysis of the ethnographic method shows, however, that in reality an ethnographic approach often involves periods of analysis between or during actual observation (Emerson *et al.*, 1995). The use of ethnography in studying information systems has even produced a proposal that ethnographic studies should be employed in quantity rather than produced in quality (Sharrock, 1997). In practice, this means that ethnography can be used to focus in on important issues quickly, to uncover common strategies and to form the basis for further, more directed, experimentation. In social anthropology this is often termed *microethnography* (Wolcott, 1995) and implies that a culture may be reflected in a collection of particular behaviours in particular settings and it is not therefore necessary to attempt to understand and describe the whole cultural system.

6.1.5 Experimental perspective

This short discussion has presented a cross section of experimental methods. The targeted laboratory study that isolates causes and observes effects to reveal causal relationship; the more open case study approach where causal relationships are uncovered from an understanding that one can only record a small set of the possible causes; and the ethnographic approach where observable actions are used to explain experiences rather than to prove hypothesis. The question that now needs to be addressed is, where do the investigations undertaken for this dissertation fit into this spectrum of methods?

A simple distinction between the investigations undertaken in this dissertation has already been made. The investigation presented in this chapter was designed to highlight areas of concern, to uncover general pictures of the interaction and, to create an initial focus to direct further experiments. The second investigation, which is presented in Chapter 7, is the result of this focus. This second investigation used an approach that was directed towards the particular areas of interest, the salient issues of the interaction, in this case the URL. This investigation was conducted in a highly controlled setting where interaction was inhibited to try to examine the existence and properties of the causal relationships present. This is not to say, however, that the experiments were performed in a metaphorical vacuum. They can, however, be considered as tending towards laboratory experiments, they had *a priori* hypothesis, observation was directed onto specific variables, and they intended to uncover causal relationships.

---

[2] The term "thick description" was first used by Geertz (1973) and has since been used throughout ethnographic literature.

Monk and Watt (1999; 2000) have used a similar approach - mixing *field* studies with laboratory experiments – whilst investigating the collaborative nature of telemedical consultation. Ethnographic studies, undertaken by HCI researchers, have identified overhearing as an important part of coordinating work; for example, in the study of a London Underground control room (Heath and Luff, 1992). Using this finding, Monk and Watt conducted a number of field studies into telemedical consultation and concluded that overhearing the communications of other people in this work place was an important tacit action that contributed to the overall success of the work. Having identified this particular issue from the situated field studies, Monk and Watt proceeded to isolate this phenomenon and undertook laboratory experiments to investigate in detail the effects of overhearing on electronically mediated collaborative work. Monk and Watt uncovered a salient issue of the interaction from a situated field study, and then proceeded to investigate it in the laboratory. The same approach is taken by the work presented in this dissertation: a situated investigation to uncover the important issues of the interaction, followed by controlled laboratory experiments to investigate some of these issues more closely.

The aims and objectives of the focusing investigation were qualitative and this research had to be based on situated fieldwork. The question was, therefore, whether this investigation should be considered as using an ethnographic approach or, a case study approach.

> *"...ethnography is more than a slogan or a label. It is a commitment to provide a 'cultural account' for the actions of an individual or group."*
> *(Van Maanen, 1995, p 24)*

There are a number of factors about the focusing investigation that place it very much in the ethnographic camp. The researcher entered each study with no hypothesis to prove or particular variables to observe. Every effort was made to let the participants perform their WWW browsing in a natural and uncontrolled manner and most importantly, the field notes that were collected offered the researcher's interpretation of the interaction and not just empirical observations of actions. Indeed, in the presentation of the results for this experiment many experiences transcended the actual observations. The researcher gained many insights and produced many field notes in other environments throughout the lifetime of this research. The researcher was regularly immersed in the environment, interacting with the WWW and studying the interaction of others.

In complete contrast, however, there are a number of issues that mean the approach can be considered a case study. Empirical data was collected concerning the information system and despite all the efforts to observe participants in a natural environment, participants were actually asked to browse the WWW in a laboratory and in front of audio-visual recording equipment.

It has been argued that the choice between ethnography and the case study approach is polarised, and fieldwork can only be considered as one or the other (Yin, 1993). It is clear from the previous paragraphs that this is not an argument subscribed to in this dissertation. The experimental perspective that this work uses sits uncomfortably between the two approaches drawing something from each. Every attempt was made to give this experiment a degree of "*ethnographicness*" (Wolcott, 1995) but to claim that it was a pure ethnography would be misleading mainly because the presentation of many of the results is centred on the evidence uncovered in the short and controlled studies presented in this chapter. Having established the experimental perspective, the reader's attention is now directed towards the first investigation.

6.2 Aims and objectives

The prime objective of the focusing investigation was to elicit participants' explanations for the dynamic behaviour they experienced while using the WWW. The aim was to distinguish between what information users extracted from the interface to support their explanations, and what information the user assumed or brought to the session as domain knowledge or derived from the environment. The main method used to extract a participant's explanation for an event was direct interviewing but this was blended with an acute observation of the WWW user in action. The hope was to reveal participants' tacit actions, behaviours and practices and with the use of audio-visual recording equipment collect and analyse data that would subsequently form a medium for consultation between the researcher and the experimental participants.

As well as the participants' overall mental pictures and the specifics of actual explanations for events, the study also aimed to collect empirical data on the Web pages that participants requested and viewed. Automated logging was used to collect data on the make-up of all the Web pages requested during the study and the analysis of this data formed the basis for a characterisation of the WWW. This characterisation was not only a contribution to the field in its own right, but also a basis for validating the generality of the study.

The main tool to extract the users' explanations was the use of verbal protocol techniques. A pilot study was therefore undertaken to decide whether to use a concurrent, or a re-enactment verbal protocol. Half of the pilot study participants were observed and then interviewed using a re-enactment technique, and the other half were observed and concurrently interviewed using a constructive interaction technique. As well as this decision on methodology, the pilot study also uncovered practical problems with visually recording the participants and software problems with logging their browsing sessions. An explanation of the revised apparatus and its use to record a participant visually and log their browser sessions is given in Section 6.2.3.

6.2.1 Pilot study

What the pilot study uncovered was that although the use of a concurrent verbal protocol technique involved less work for the experimenter, the results were not wholly satisfactory. The constant verbalisation between the experimenter and the participant produced a very controlled environment and participants often asked for help or simply claimed ignorance and refused to continue when faced with a problem. The participants constantly justified their actions before performing them. Rather than change their justifications when confronted with unusual or unpredicted behaviour that contradicted them, they would attempt to transform the events to account for the behaviour. This behaviour of the participants is very similar to what Norman calls the *demand structure* of a situation that was introduced in Chapter 5 (Norman, 1983).

In contrast, the participants who were left alone whilst browsing and then interviewed using the re-enactment verbal protocol technique, were often observed performing intricate sets of actions to circumvent a problem. Their explanations for these actions were often self-mocking, and highlighted a problem with their own mental models that they subsequently rationalised and adapted. The use of re-enactment not only allowed for a more natural browsing environment but it also produced far less defensive interviewees more willing to explain their misunderstandings. A re-enactment verbal protocol was, therefore, used in the subsequent investigation.

6.2.2 Participants

Participants were unpaid volunteers, recruited from under-graduate and post-graduate students at the University of Edinburgh. Nineteen participants were recruited for the study, six of whom undertook the pilot study to determine the best verbal protocol method. The empirical data presented in the results does not include any data from the six participants who were observed as part of the pilot study.

Recruitment from a student population produced a very limited age range from 17 to 29 years of age with a gender ratio of about 2:3 female to male. Each participants' classification of their general computer expertise and in particular experience of using the WWW was recorded. Out of the thirteen valid participants, two participants claimed never to have used the WWW, four claimed to be novice users, three intermediate users and four expert users. The distribution of participants' nationality was nine British, two Irish, one Mexican, and one Greek. Only two had a first language that was not English, although all the participants could speak and read English to a high standard.

Participant	Gender	Age	Expertise
A	F	21	Novice
B	M	28	Expert
C	M	27	Expert
D	M	25	Expert
E	F	22	Virgin
F	M	27	Intermediate
G	M	22	Intermediate
H	F	29	Virgin
I	F	21	Novice
J	F	21	Novice
K	M	28	Expert
L	M	17	Novice
M	M	21	Intermediate

Table 6.1: Participants' gender, age and expertise.

6.2.3 Apparatus

The apparatus used in the experiment can be considered in two parts. First, there was the data logging software used to record participant's browsing sessions. Secondly, there was the apparatus used to record the browsing session onto videocassette so that it could be analysed in depth.

6.2.3.1 Data logging

Participants were invited to use the Netscape 3 browser on a PC running the Windows95 operating system. The browser was set-up to request Web pages through a proxy server running on a machine connected to the local network[3]. The proxy server was used to intercept HTTP requests and responses and log a summary of the Web pages that were subsequently sent to the participant's browser. The log recorded the Web page's URL and the HTML document's title and a count of the embedded objects.

```
URL           : http://mirage.snu.ac.kr/index.html
HTML Title    : CODESIGN AND PARALLEL PROCESSING LAB.
No_frames     : 0
No_images     : 7
No_links      : 19
No_applets    : 0
No_others     : 0
Download time : 23223
```

Figure 6.1: Extract from a log file

The proxy server also recorded the time taken from the HTTP request for the HTML document being sent from the proxy to the last embedded object being fully received by the proxy; the total download time of the Web page[4]. The time recorded for the download was slightly less than the delay experienced by the browser user, since there was an additional two communications between the browser's client and the proxy. The first communication involved sending the first HTTP request to the proxy and the second communication was

[3] The proxy server was a Java program that was run on a SUN sparc5 using the Java Virtual machine.
[4] The Web page was under study here and not the separate HTTP communications. A Web page is defined by an HTML document and may include a number of embedded objects. All these objects may be communicated concurrently but the download time of the Web page is simply the overall time from the first HTTP request to the last HTTP response. One problem confronted was distinguishing between different Web pages when frames were used. Frames allow what this work would define as a single Web page to be defined by a number of different HTML documents. The total time for the download of a Web page, therefore, had to include the download time for all the frame's HTML documents.

receiving the last piece of the final embedded object from the proxy. The temporal cost of these two communications was ignored, since it was, in general, insignificant compared to the total temporal cost and in practice non-trivial to calculate during the experiment.

6.2.3.2 Audio-visual recording

The actions of a participant were observed and recorded using a variety of equipment. The output to the monitor from the computer running the browser was converted to a PAL signal for recording on to a videocassette. This signal was *mixed* with the image from a video camera that was used to capture the physical actions of the user. A microphone was used to record the utterances of the user as well as any external environmental sounds, and this was mixed with the sounds produced by the computer. This mixed signal was then recorded on to the videocassette's stereo sound track. The result was a single videocassette of the WWW browsing session capturing the physical actions of the user, the output to the monitor, the output to the speaker and the surrounding environmental sounds. The audio-visual recording set-up is shown in Figure 6.2 and a frame of an actual recording is shown in Figure 6.3.

6.2.4 Data collection

The data collected during each study falls within one of three different categorises. First, there is the data concerned with the participant, their personal details and experience with using the WWW. The experimenter recorded this by hand, before the browsing session had started. Second, there is the data collected during the browsing session. This form of data includes the automatic log files and the audio-visual recordings. Finally, there is the audio recording of the re-enactment interview performed after the browsing session was complete. Each study concerning a single participant can, therefore, be divided into the following four phases with data being collected at stages 1,2 and 4.

1. Instructions to participants
2. Observation and logging
3. Review of data
4. Re-enactment interview

6.2.4.1 Instructions to participants

All the participants were asked to sit in front of the browser and were then given the following instructions orally.

1) The Netscape 3 WWW browser was running on the machine in front of them and could now be used to browse the WWW.
2) After any amount of time, the participant was free to stop their browsing session for any reason.
3) The room was being filmed and all sounds recorded but the participant was free to turn off the audio-visual recording equipment at any time (and shown how to do this).
4) Data was being collected about the content of Web pages they requested and the time of the subsequent download.
5) All the data collected was confidential and would only be accessible to the experimenter.
6) The participant's identity would be kept confidential if requested.
7) If the participant had any questions at any point during the session, they could ask the experimenter or anyone else who happened to be in the room.

After these instructions had been given, the participant was asked to state their experiences of using the WWW. When they had finished explaining their experience they were asked to categorise themselves into one of the following types of user: virgin, novice, intermediate, or expert. The participant was then left to start their browsing session.

6.2.4.2 Re-enactment interview

Once the participant had finished their browsing session, the audio-visual data was analysed and interesting sections of the video noted. After this analysis, the participant was asked to return to the laboratory and, in conjunction with the experimenter, review selected sections of the video.

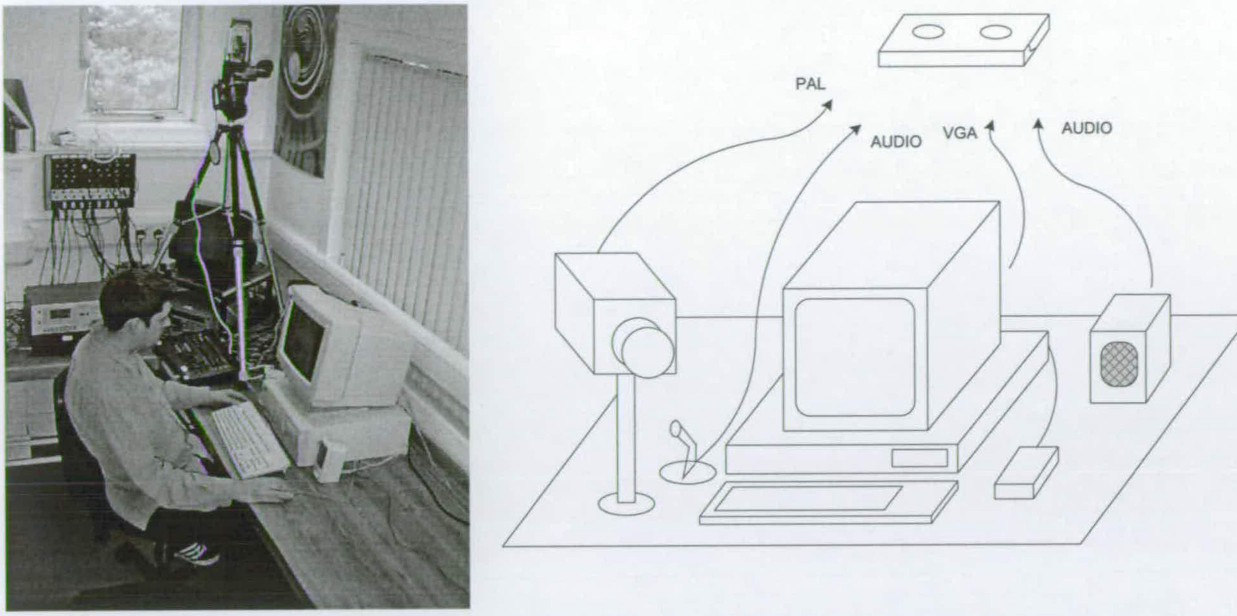


Figure 6.2: The audio-visual recording apparatus.

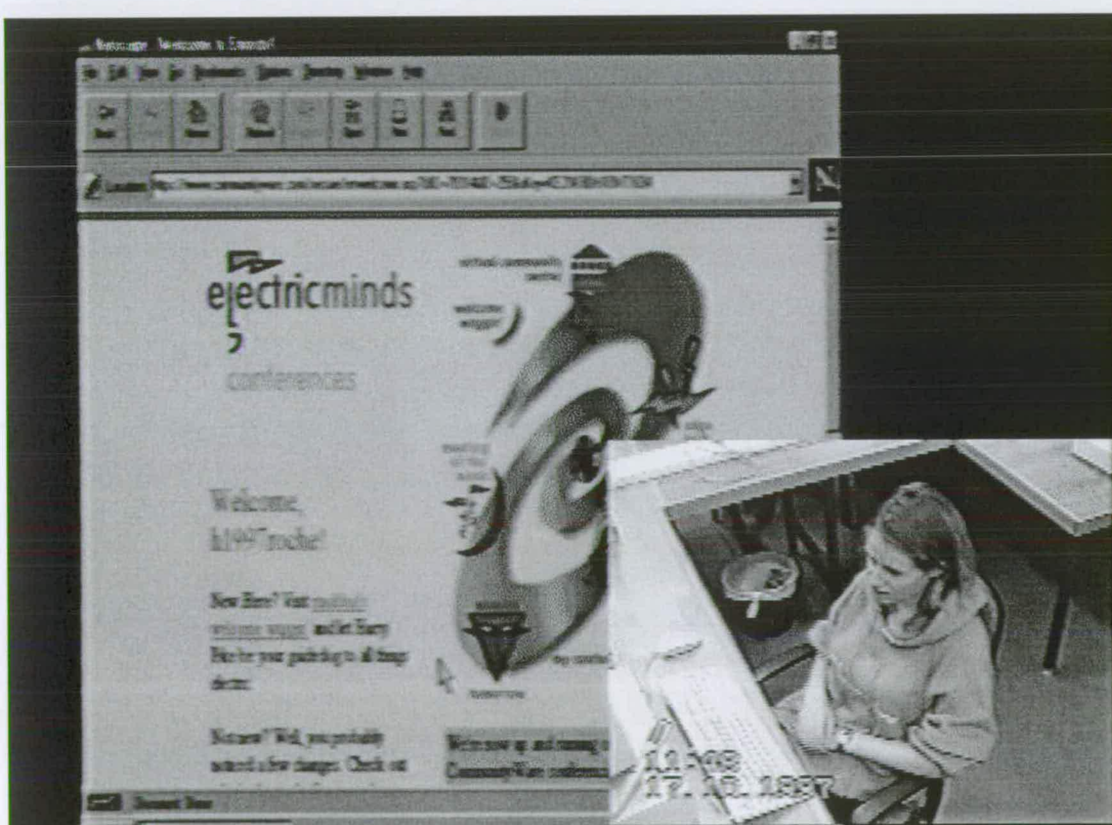


Figure 6.3: A frame from the audio-visual data.

An interview between the experimenter and the participant retrospectively interpreted the actions of the user and the user's understanding of the system's response. This is, of course, an example of a re-enactment verbal protocol technique. Each interview was recorded on audiotape and extracts from these recordings are presented later in this chapter. The notation used to present the verbal utterances is adapted from Atkinson (1984).

6.3 Results and analysis

The analysis of the data collected throughout this investigation divides into two distinct sections. First, there is the analysis of the quantitative data collected in the log files, which confirms some of the WWW characterisation presented in Chapter 4 and adds further statistical data to this body of knowledge. Second, there is a presentation of the qualitative study, which aims to report on the nature of the interaction and present a detailed picture of users' understanding.

Table 6.2 presents a summary of the data collected in the log files during the browsing sessions of the thirteen main participants. The analysis of the log file data is presented within the three-domain framework – information, space, and time – and provides both new findings to characterise browsing sessions and also, repeats the analysis of some of the studies presented in Chapter 4.

Total HTML documents	691
Different HTML documents	415
Images	3757
Embedded objects (excluding images)	7
Web sites (different domains)	129

Table 6.2: Summary of the investigation's log files.

The experimental perspective of the investigation focussed on the observation of the nature of interaction rather than the collection of large amounts of statistical data. For this reason, the empirical analysis of the data collected in the log files uses a limited amount of data in comparison to other studies and care must be taken when making comparisons.

6.3.1 Information

The log files recorded the number of embedded objects and the hyperlinks contained in each Web page. During this investigation, the mean number of images contained in the Web pages downloaded by participants was 5.3 images, with 67.2% of these Web pages containing at least one image. Similarly, the mean number of hyperlinks that these Web pages contained was 24.6 hyperlinks, with 80.8% of all these Web pages containing at least one hyperlink. Table 6.3 shows these findings and compares them to some of the studies by other researchers previously presented in Chapter 4.

Although there are some differences between the findings of each study the only significant contradiction is in the mean number of images contained in a Web page. Moore *et al.* (2000) state that the mean number of images contained in a Web page is 14.4 whereas Woodruff *et al.* (1996) and the investigation carried out in this dissertation state the figure to be 4 and 5.3 respectively. The reason for such a significant difference may be due to the Moore *et al.* investigation being undertaken more recently. There may exist a trend of including more images in a Web page and the Moore *et al.* result is simply a reflection of this. More studies have, of course, to be undertaken to justify such a conclusion.

	at least one hyperlink	at least one image	mean number of hyperlinks	mean number of images
This study	**80.8%**	**67.2%**	**24.6**	**5.3**
Woodruff *et al.* 1996	~90%	~60%	15	4
Bray 1996	~75%	~50%	-	-
Moore *et al.* 2000	-	-	28.6	14.4

Table 6.3: Web page characteristics in comparison to other studies

Overall, however, the statistical findings from this study presented in Table 6.3 are in general agreement with the other studies and this shows that the set of Web pages downloaded and read during this investigation appear to conform to *normal* Web page characteristics. Conclusions drawn from their use and dynamics can therefore be generalised across Web pages with some confidence[5].

[5] The actual size of embedded objects and HTML documents was not recorded in the log files. In hindsight, this can now be seen as a mistake since it would have provided a further variable to

6.3.2 Space

To help in the analysis and presentation of the data collected about the spatial movement through the WWW, a simple graph was constructed for each participant's session whose vertices represented *Web sites* visited and whose edges represented their connectivity with respect to the observed session. If an edge exists between nodes A and B, then at one point during the session, the participant requested and subsequently received data from Web site B when previously having viewed data from Web site A. The participant's session was then described as a walk[6] through this simple graph. Stanyer (1996) suggested that analysis of walks in information graphs could be used to determine automatically whether users were disorientated and, therefore, lost in a hypertext. One of the major usability problems already highlighted in this dissertation is the requesting of a Web page that turns out to be useless: what Nielsen (1998a) termed *"going down the garden path"*. The simple use of this graphical technique here explores usability problems like this.

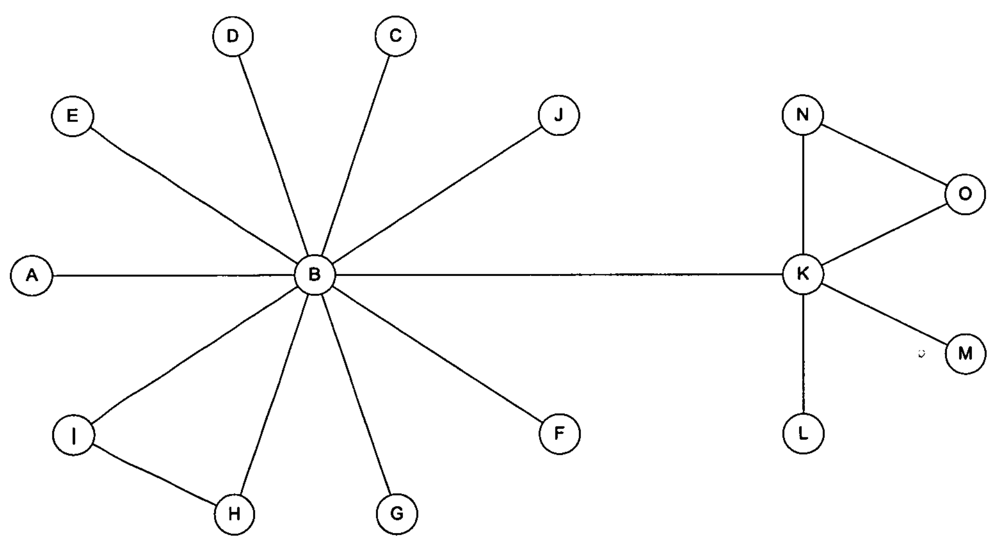


Figure 6.4: The walk of participant F between Web sites

---

compare with the other characterisations of the Web page and hence would have been another way to ensure external validity.

[6] A walk is a finite sequence of vertices where every consecutive pair of vertices is connected by an edge.

Figure 6.4 shows the *movement* of a participant through the hypertext using a walk through the simple graph. It is clear that vertex B is central to this particular browsing session as it is returned to no less than ten times. In fact, participant F used a search engine to display a list of hyperlinks that *matched* a search string and was supposed to be relevant to participant F. The results of this search were displayed on a Web page delivered from Web site B and this explains its constant use: participant F started at the top of the search list and worked his way down activating each hyperlink in turn.

The graph and walk reveal that it was only after eight different Web sites that participant F found relevant information at vertex F and started browsing properly. This example confirms the usability problem of users downloading what turns out to be irrelevant Web pages. What this simple analysis cannot do is explain why users make these *mistakes*. This is left to the ethnographic analysis.

6.3.3 Time

The studies on WWW usage that were presented in Chapter 4 predominantly used either server-side or client-side logs to determine the characteristics of the Web page and the events that users performed. These studies did not observe the user actually using the browser. One area where this event logging technique caused problems for the researchers was in determining how long a session was; *i.e.,* for how long a user actually used a browser. Cunha *et al.* (1995) proposed that a session could be defined as a single execution of the browser. This, of course, implies that users *start* a browser before actually browsing and subsequently *close* the browser on completion of their session.

The first problem with such an assumption is that users do not close a browser after using it and the browser process may stay *idle* for a long time before being used again. Should this be considered as two sessions or one? Catledge and Pitkow (1995) stated that these were independent sessions and so devised an artificial *timeout* for a session. Having calculated the mean time between events to be 9.3 minutes, Catledge and Pitkow used a simple statistical model to decide that if there was a delay of more than 25 minutes between events, each event should be considered as being part of different sessions.

The second problem with the Cunha *et al.* assumption is that users have often been observed using multiple browsers. This technique is employed to compensate for slow data transfer times and has been termed *browser multi-threading* (McManus, 1997), or as an ad-hoc example of *temporal disturbance management* (Parker, 1997). This behaviour is investigated more fully later in this chapter (section 6.4.3). Nevertheless, if a user employs multiple browsers the Cunha *et al.* assumption would mean that they are in fact undertaking a number of simultaneous browsing sessions: this is not the case.

Determining for how long users actually employ the browser has therefore become an area of some debate. The fact that the investigation presented here used observational techniques means that it was relatively easy to determine when a session started and when a session ended. The result was that the mean session time was determined to be 32 minutes 51 seconds. This result is shown in Table 6.4.

6.3.3.1. Data transfer delays

The log file recorded the length of time the Web page – the HTML document and all the embedded objects – was in the state of download. This length of time was the time it took to transfer totally all the parts of the Web page, or the time between hyperlink activation and a user-defined interruption. Analysis of this data uncovered that the mean time for the transfer of a Web page throughout the investigation was 16 seconds and, perhaps more interestingly, that the percentage of the session taken up by the data transfer process had a mean of 46%[7], but could be as high as 72%. These findings are presented in full in Table 6.4.

	Session time (min)	Data transfer time per page (seconds)	Total time transferring data (seconds)	Data transfer time as a % of session time
Mean	33	16	912	46%
Sd	6.4	6	302	14.1
Max	46	187	1130	72%
Min	21	0	427	23%

Table 6.4: Summary of the duration and delays during a browsing session

[7] This does not mean that the user had to be idle for 46% of their session, since Web pages can be partially rendered before data transfer is complete, but it does show what an important part of the browser usage the data transfer process is.

One important factor that users often suggested as being important in their verbal accounts was that the time of day affected the download time. For completeness, Figure 6.5 shows the time of day at which each participant undertook their browsing session and it also shows the percentage delay they encountered as the greyed out area of the boxes. A correlation between time of day and total percentage delay is not immediately apparent from the data.

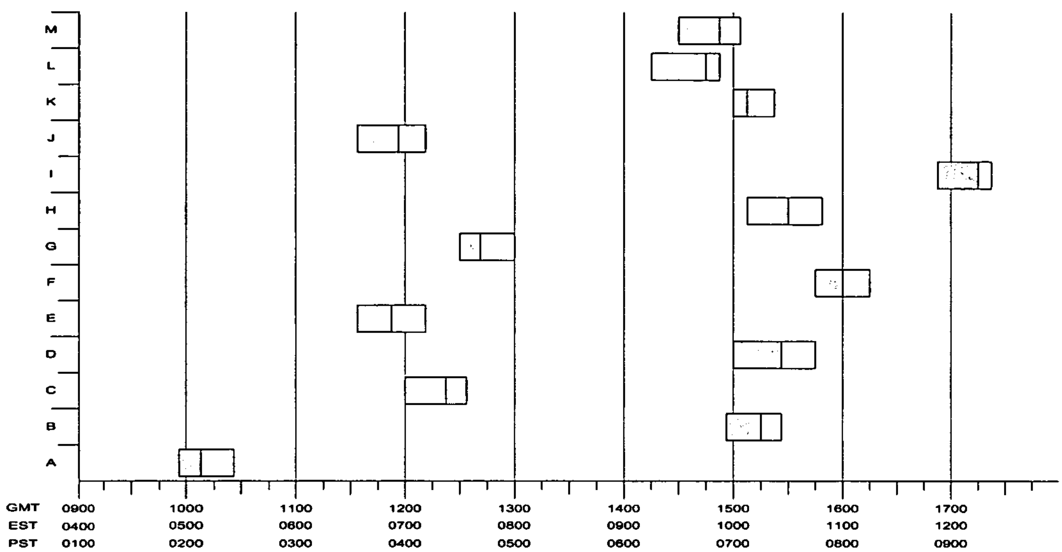


Figure 6.5: Distribution of sessions throughout the day (grey areas indicate the total data transfer delay that each user experienced)

6.3.3.2. Recurrence rate

The mean number of requests for Web pages during a participant's session was 61 Web pages. The standard deviation around this mean was 25.9 suggesting that the difference between the numbers of Web pages that participants looked at was marked. The number of requests is, however, not equal to the number of *different* Web pages that a participant viewed during their session: it is only equal to or greater than the number of *different* Web pages. The mean number of *different* Web pages viewed by participants was 36 Web pages. These results are summarised in Table 6.5.

Another characteristic of WWW usage that was presented in Chapter 4 was the concept of the WWW being a recurrent system. Tauscher and Greenberg (1997) introduced the concept of a *recurrence rate*, and showed that around 60% of all Web pages received by a user had been visited previously. Employing the same method in this analysis revealed that the recurrence rate was only 41.6% (std = 11.4%). The original definition of recurrent systems stated that the system must posses the characteristic that users *predominately* repeat activities they have invoked before, while still adding new actions to the repertoire from the many that are possible (Greenberg, 1993). The analysis of the data collected from this investigation shows the WWW failing Greenberg's definition of a recurrent system, although it may be considered as a borderline case.

	Total Web pages visited	Different Web pages visited
Mean	61	36
Sd	25.9	13.2
Max	113	63
Min	24	14

Table 6.5: Web pages visited

The implications of this result are that the users observed during this investigation were less likely to request a Web page seen previously than the users recorded by Tauscher and Greenberg. It is not clear what to conclude from the difference in recurrence rate. In the worst case, this finding may imply that the actions of the participants were in some way transformed by the environment and the existence of the observation apparatus: the users were inhibited by the recording equipment and therefore were less likely to perform repeated actions (since these may have been interpreted as admissions of a mistake). This would jeopardise the external validity of the investigation and be a cause of great concern. The difference in the recurrence rates proposed by Tauscher and Greenberg and the one determined by this investigation is not, however, great. A more appropriate conclusion may be that further studies, which analyse the recurrence rate of the WWW, will simply place its value between the two discussed here, and the actions of the participants under investigation have not been transformed in any way[8].

6.3.4 Summary

Quantitative analysis of the log files was used to confirm some of the WWW characteristics and usability problems presented in Chapter 4. By comparing the Web pages downloaded and perused during this investigation with the studies of Web pages presented previously, it can be concluded that this investigation appeared to be using *normal* Web pages. The one area where the findings reported in Chapter 4 were in contrast to the findings of this investigation was in the recurrence rate of Web pages and whether the WWW is a recurrent system. Despite this small difference, the Web pages used and the actions of the users are not sufficiently different to undermine the external validity of the results presented here, and in the subsequent ethnographic analysis.

A very significant result of the log file analysis that was not a confirmation of previous studies was that on average, the browser was transferring data for just under half of the total session time. The percentage of session time devoted to the data transfer process actually ranged from 23% to 72% of the total session. The data transfer process appears to consume a very significant proportion of the whole browsing experience. The mean time for a Web page download was found to be as much as sixteen seconds, although this may be misleading as it is skewed by some very long delays indeed. Nevertheless, the data transfer process does take up a considerable amount of every participant's browsing session and their sensitivity to it must contribute to the usability of the system as a whole.

6.4 Qualitative analysis

The results presented in the following section are concerned with the experiences of the WWW user, the beliefs they possessed and the strategies they employed. The aim of collecting these results was to report on the nature of the interaction, to interpret the observed action's significance to the users' experience. The justification for these reports on the nature of the interaction is mainly derived from the transcripts of the participants' re-enactment interviews and the field notes of the researcher. The reports also contain the direct beliefs of the researcher and, as is the tradition in ethnography, they are justified not with a large quantity of in-depth observations but rather by subtle interpretations of events

6.4.1 The general model of WWW interaction.

Browsing the WWW revolves around the user activating hyperlinks. As has previously been defined in Chapter 4, a user can activate a hyperlink *implicitly* by clicking on the activating object of the hyperlink; *explicitly,* by typing a URL into the address bar; or the user can use a

[8] The argument can, of course, be reversed and it may be the actions of the participants involved in the

navigational feature of the browser like the Back or Home buttons. A browsing session consists of a user performing a sequence of these different methods of hyperlink activation.

In response to each hyperlink activation, the browser attempts to render the resource referenced by the hyperlink. The browser's reaction is to start the data transfer process for the referenced resource; the download of the HTML document and every embedded object required to render the Web page. An associated data transfer process follows each hyperlink activation. The data transfer process is complete when all the data has been successfully transferred to the client and the resource is rendered. The dialogue between the user and the WWW cycles between the user initiated action, hyperlink activation, and the system's response, the data transfer process (Figure 6.6).

The data transfer process may, however, not complete successfully. A failure can occur in both the fabric of the WWW or in the communication processes of the Internet. The result is either a communication *timeout* or an HTTP error message being sent to the client and subsequently displayed. A communication timeout is usually set very high (10 minutes is the default) and so users have the option to interrupt the data transfer process by performing an *interrupt* action. Different WWW browsers offer a number of different actions to interrupt the data transfer process. The Netscape 3 browser offers just two.

- **Stop Button** – this stops the communication of any unsent HTTP requests and stops the receiving of any further outstanding requested data. It then forces the browser to try to render the Web page as best it can, using the limited data already received.
- **Reload button** – this stops the browser sending any more requests, discards the data that it has already received, and then re-sends the initial HTTP request for the HTML document.

Throughout a browsing session, the user has to make two different types of major decisions. The first is to assess whether a hyperlink should be activated, and the second is to decide whether a particular data transfer process should be interrupted. The decision whether to interrupt the data transfer process is looked at in detail in Section 6.4.2, but basically a user has to decide whether the behaviour they are experiencing in the form of a delay is the result of a network disturbance or simply the *normal* behaviour of the communication. After deciding to interrupt, the user is then faced with a further decision: whether to repeat the

---

Tauscher and Greenberg study that have in some way been transformed.

request for the Web page or to abandon the request altogether. The resourcefulness of users at uncovering information to inform the making of this decision is what is of interest: the improvised nature of their actions to cope with a system behaviour that, although common, is always situated in the circumstance of the moment – the time of day, the contents of the Web page, the location of the data *etcetera*.

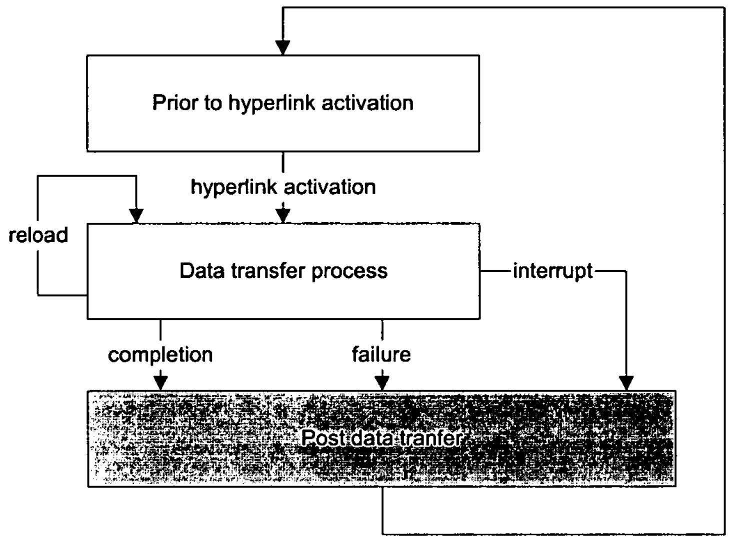


Figure 6.6: The hyperlink activation cycle

The qualitative results that follow centre on the decisions to activate a hyperlink and to interrupt the data transfer process; what information the system revealed, what knowledge users brought with them to the session and what beliefs users possessed to help them make these decisions.

6.4.1.1 Hyperlink activation

Participants showed a great desire to uncover information about the referenced resource before activating the hyperlink. The simple question that participants constantly expressed during their interviews was whether it was worth activating the hyperlink. In the language of Johnson, was the perceived utility of the resource high enough to justify the potential cost of waiting for it to download (Johnson, 1997)? Participants based their predictions about a referenced object's utility using a combination of resources:

- The Web page in which the hyperlink was situated
- The hyperlink's activating object
- The URL of the defining HTML document
- Previous knowledge or recommendation

The last of these methods, previous knowledge and recommendation, is the least contentious method and it is easy to empathise with such reasoning (Extract 6.1). If a participant had been to a Web page or site previously, they believed they knew what information would be presented to them on activating the referencing hyperlink. The dynamic nature of the WWW meant that this information object might well have changed since they last retrieved the information, but participants seemed either unaware of this or were prepared to take such a risk.

Subject C: I know the Tesco site, it has a load of recipes so I may be able to find a pizza base there.

Subject K: I knew it was going to be slow because I'd been there before.

Subject E: I have to go here, my lecturer told me this place was really good.

Extract 6.1: Recommendation and previous knowledge

In general, however, participants have no previous knowledge of the referenced resource and so must employ other methods to assess the resource. These methods are far subtler and deserve some attention. A hyperlink is situated in a Web page and this Web page provided a means for participants to assess the potential value of the referenced resource. Principally there is the hyperlink's activating object that usually provides a text explaining some aspect of the referenced resource. The actual textual activating object may itself be included in a surrounding text providing more detailed context for the referenced resource. How useful this text is at describing the referenced resource is, of course, dependent on the author of the particular Web page in which the hyperlink is situated. These textual descriptions would in most cases just provide a summary about the referenced resource's informational content but sometimes, technical information like the resource's type and or, the data's size would be provided (for example, a 45 Kbytes gif encoded image).

Interviewer: How did you know that by activating that hyperlink you would get information on pizza bases?

Subject C:	Well I didn't from the link, it just said food links.

Extract 6.2: Participant C using a textual activating anchor to assess the referenced resource

In some instances the referencing Web page would provide information about the probable quality of the communication and the probable time the download would take, depending on the type of connection being used. These were, however, the exceptions and in most cases users were left to assess the relevance of the referenced resource and the quality of service they could expect from a short sentence or single word. The diversity and richness of participants' reasoning about the referenced resource prior to activation of a hyperlink revealed that users employed another very important component to help them make their assessments.

6.4.1.2 The URL of the HTML document

As previously described in Chapter 3, a hypertext Web page must have a defining HTML document and this document, like all WWW resources, will have a URL. In technical terms, the relationship between this URL and the Web page as a whole is strong, since a browser will download all the resources of the Web page having been explicitly given just this URL. Participants generally perceived this URL to reference the Web page as a whole. Even though many of the expert users knew the technology behind the referencing of resources, the language that they used always implied that the Web page was a single entity and it had a single reference; the defining HTML document's URL.

This perception of the very strong relationship between the URL of a HTML document and the Web page should not cause any surprise. The WWW browser places great emphasis on the defining HTML document's URL by prominently displaying it in the address bar when the Web page is rendered. Indeed, while assessing whether to activate a hyperlink, the only feedback a user receives from the browser is the display of a URL when the mouse pointer *hovers* over a hyperlink's anchor[9]. This naturally reinforces the association between this URL and the Web page as a whole. The result is that a user associates one URL with one Web page despite the fact that there is actually a URL for every different object in the Web page.

[9] With some browsers, there is also the possibility of a textual label appearing when a mouse cursor hovers over an activating object. This feature is introduced and fully explained later in the dissertation.

What is surprising is the extensive and continued use of the URL to assess the informational content and the quality of the communication of the whole referenced resource. The more experienced the WWW users were, the more confidant they were at justifying their understanding using the URL of the defining HTML document. Even when the WWW user had in-depth knowledge of how the technology worked they still justified their understanding using a model that had a one-to-one relationship between this URL and the whole Web page entity (see Extract 6.3). The use of the URL to assess referenced resources was not exclusive to expert users. During the observational investigation all the participants except three novices (two of whom had never used the WWW before) used the URL of the defining HTML document as evidence to justify an assessment of the referenced resource's value to them. The relationship between a particular URL and a WWW resource is an area that, therefore, needs special attention and that is why it is highlighted as an issue for further experimentation.

Subject C: I thought it would have lots of text because the Web address told me the Web page was from an academic institution. They normally have really boring pages, not too many pictures and things. Also the address is long, lots of slashes so it is quite specific.
Interviewer: How did you know it was an academic institution?
C: the dot ac dot uk, you know it was Manchester University
I: Manchester?
C: yes man ac uk. I thought it would be quick as well as we are on the academic network.
I: But how did you know all every object on the Web page would come from the same place?
C: What do you mean?
I: Well you are only looking at the URL for an HTML document, and all the images and
C: Oh well yes I suppose so but I always think of a Web page as coming from one place, you know Manchester, it probably all did, I wouldn't know though, I suppose it all came from the same place. Maybe not, that is why it was so slow?

Extract 6.3: Participant C using the URL to assess hyperlink activation.

The belief that Web pages were single entities referenced by a single URL was extremely strong and dominated every participant's explanation. This abstraction is what the WWW is based on and it is strengthened by the hyperlink itself. The fact that WWW users accept this abstraction so completely is a factor in making the WWW usable. In the static, pre hyperlink activation phase of WWW use this abstraction was never in question, the abstraction does, however, sometimes start to fray and this seems most apparent in the data transfer phase.

6.4.2 Data transfer

The complete temporal affordance for the WWW browser has already been presented in Chapter 4. When referenced resources were very large, but well defined by their HTML document (and particularly if they were simply pages of text), the browser would start to render the Web page before all the data had been transferred and continue to render the Web page as the data arrived. As more data arrived, information was added to the end of the Web page and the browser showed this by the scroll bar button on the right of the Web page getting smaller (indicating a longer page). Similarly, if the position of embedded objects was well defined in the HTML document a space was often left for them and the rest of the Web page rendered as the data was retrieved. This behaviour is a direct result of the Web page being made up of a number of real data objects. In explaining the process that unfolded before them, participants still clung to their belief that the Web page was a single data entity and explained this particular temporal feedback in terms of a single file. This had the effect on novice users of not touching any part of the interface until the whole Web page appeared to be rendered. The novice users were apprehensive of scrolling down if they believed that the whole Web page was not present. It was, however, not always clear that the Web page had finished downloading and on occasion participants were caught out and suddenly found that data would arrive and change the look of the Web page before them. This behaviour often caused confusion.

On a number of occasions when this occurred, participants revealed that they were unaware that the data had not fully arrived. A screen full of information appeared and they assumed that the whole Web page was present. After perusing a short part of the Web page, they would start to scroll down and then discover that the text finished prematurely. Extract 6.4 shows two novice users' explanations for this type of WWW behaviour. Participant I uses the phrase *"I suppose the computer was just slow"* in explanation for the delay. This is of particular interest as it is an example of an understanding that was uncovered in the previous chapter. This user does not have a clear understanding of data transfer and instead relies on a concept of computer speed to rationalise delays they experience. The interface is not presenting a clear picture of data transferring but simply the concepts of Web pages appearing.

```
Subject I: I was surprised that the text was not there, so that is why I kept
clicking on the scroll bar.  I suppose the computer was just slow, it appeared later.

Subject E: I think something went wrong that is why I kept trying to scroll down.
```

Extract 6.4: Insufficient progress affordance

As well as causing problems for the less experienced participants, the fact that the browser rendered the Web page before the data transfer was completely finished also prompted the expert participants to adopt strategies to elicit data transfer performance. On a number of occasions, Participant D was observed *tapping* the down arrow on the keyboard during data transfer. The effect of this tapping was that as data arrived and the browser rendered the next part of the Web page, pressing the down arrow would scroll the page. If the data transfer process was slow, the page would scroll only occasionally and slowly. If on the other hand the data transfer process was fast, the page would continually scroll down. Participant D used this fact to determine if an interrupt action was appropriate by continually pressing the scroll bar during data transfer: as data arrived the page would scroll a little then stop while more data arrived then continue scrolling. The speed at which the data arrived and subsequently the Web page scrolled down provided the participant with a method of determining the latency of the connection. On two occasions this resourceful strategy resulted in the data transfer process being stopped and the same Web page requested again. This is an example of how possessing some understanding of the underlying technology has a direct influence on a user's actions.

6.4.2.1 Interrupting the data transfer process

The use of an interrupt action to stop the download process was extensive during the observational study. Only two participants never employed these actions and both were novice users who during their re-enactment interview explained that they were unaware of what the buttons did and had, up to then, never used them. Other studies have logged the frequency of these interrupt events, such as the Tauscher and Greenberg (1997) study that presented the statistic that 3.1% of all events were *reload* events. Similarly, the Catledge and Pitkow (1995) study stated that only 1% of all interaction events were what they termed *interrupts/asynchronous aborts.* Although both these studies used the Mosiac browser, which provides slightly different methods to interrupt the data transfer process, there is no reason to doubt that these results could be generalised to the use of the Netscape 3 browser. What was previously questioned in Chapter 4 was the conclusions that Catledge and Pitkow drew from these results when they argued that the relative infrequency of the interrupt actions meant *"the population as a whole was insensitive to retrieval latency".*

> **Scenario:** 22 seconds through the data transfer process nothing had been rendered by the browser (the screen was blank). D positions curser over stop button. 4 seconds later he clicks on the stop button and then immediately clicks on the reload button. D waited a further 16 seconds before clicking stop and clicking on back returning to a previous page.
>
> **Subject D:** There was something wrong with the Web page, the server was locked or something.
>
> **Interviewer:** What do you mean by the server being locked?
>
> **D:** Well, maybe it was the request I don't know, but I thought there was a problem so I started to think it was maybe better to request it again. If the server had too many people using it maybe another request would get a connection.
>
> **I:** So you thought that too many connections were made to the server and you weren't being aloud to disconnect?
>
> **D:** Well maybe, and if that was the case a reload may have got me a connection.
>
> **I:** Didn't you think there would be some feedback about this problem.
>
> **D:** Well no, I've never seen any, is there?
>
> **I:** How did you go about reloading?
>
> **D:** Well you can see I paused to reconsider and give it a last chance, then I stopped the download and reloaded the page.
>
> **I:** You only waited a few seconds before stopping again.
>
> **D:** Well yes, I didn't get a connection again so I gave up. It probably was nothing to do with connections, something else was wrong, maybe to do with the server being down, I don't know. Sometimes the cache can get locked and images get stuck so reload can free them up and get the page. Maybe it was just slow because America was awake?

Extract 6.5: Participant D, an expert user's interrupt action

It is not in doubt that the frequency of use of the interrupt actions, when compared to the activation of hyperlinks and the use of the *Back* button, is small. To claim that this implies that the user is insensitive to retrieval latency is, however, strongly disputed here.

Participants gave in-depth reasoning when asked about the use of an interrupt action and also showed a great awareness that it was an option available to them throughout the data transfer process. Indeed, one of the major results to come from this study was that participants showed an acute sensitivity to retrieval delays before, during and after activation of a hyperlink.

The main difference between the experts and the novices was revealed in the willingness of the experts to use the interrupt actions and, not surprisingly, their knowledge of what it did. With the exception of the two novice participants who did not understand the use of the interrupt buttons, all the participants gave adequate explanations of what these buttons did. The more expert the participant, the less apprehensive they were before performing an interrupt action. The less experienced participants pondered and deliberated before applying what they felt was a solution of last resort. Extract 6.5 and 6.6 highlights this difference, and present the rationale of an expert and novice for interrupting (or not interrupting in the case of the novice in this example) the data transfer process.

Scenario: (downloading Web page from the Times newspaper Web site) 25 seconds through the data transfer process A moves hand to mouse and starts to move the cursor. Stops for 10 seconds then places the cursor over the stop button and waits a further 23 seconds. The Web page then starts to render and is completed in 5 seconds.

Subject A: I had been waiting a while and so I thought something may be wrong and I was going to stop it but then I thought it was about 10 o'clock so lots of people in their offices would be reading the Times.

Interviewer: So therefore, it would be slow?

A: Well yes that is what I thought. I was afraid to stop it if it just going to appear in a few seconds. When lots of people use a web page its slow right.

I: And if nobody is using it its fast?

A: Yes that's what I think, when something is obscure I am probably the only person in the world looking at it so it will be fast.

I: OK, so how long do you think you would have waited for before stopping?

A: Well I was very unsure whether to stop the thingy, didn't know whether it was right thing to do. I was deliberating over the fact, but I really was unsure what was happening and why it was slow. Then it just appeared.

Extract 6.6: Participant A, a novice user's interrupt action.

In general, the expert participants felt they were applying a degree of control on the unpredictable nature of the data transfer process by interrupting the download. They were stopping an unpredictable situation and regaining control because the resolution of their previous action was in doubt. The less experienced participants, in contrast, felt that they were suspending the processes of the system and, due to a lack of feedback to describe what was happening, losing a degree of control and stepping into the unknown. The less experienced participants were always left uncertain about whether they had done the right thing. Would the requested Web page have appeared in the next few seconds and hence had they just done something that had wasted their time? This uncertainty often left the less experienced WWW user in a dilemma and this resulted in long periods of hovering over the interrupt buttons before a decision was made (Extract 6.6).

The analysis of the interrupt actions made it clear that participants continually assessed the data transfer process and tried to anticipate if a problem had occurred in order to justify stopping the download. The analysis of the participants' explanations did not, however, reveal that participants were making their decisions based on a diminishing utility of the resource as has been hypothesised by Johnson's decaying utility curve (Chapter 4). The explanations that participants gave for interrupting the data transfer were always based on whether the system had in some way broken down, resulting in the resource not being transferred at all. There was no notion of a resource becoming less valuable with respect to time and suddenly reaching a threshold where the participant no longer believed it was worth the wait. All the explanations were based on the potential of a temporary disturbance, or a permanent system failure.

6.4.3 Multiple windows

The use of multiple windows whilst browsing the WWW has been observed in other work (McManus, 1997; Parker, 1997). What this process involves is opening a hyperlink into another window, which contains another instance of the browser. This result can be achieved in Netscape 3 by clicking the right mouse button on any hyperlink and the referenced resource is downloaded and displayed in a new browser leaving the user able to continue with their current Web page and have the opportunity to activate some other hyperlinks.

Subject C: …this was a bit more meaningful explanation, and so I opened it in a new window.
Interviewer: Why did you open it in a new window?
C: The reason I opened it in a new window was because if it is very slow I can go back, if it's quite a slow site I can go back and look at the rest while some things happening.
I: right, so how did you know when the other one had finished downloading?
C: Well you see it flicks up on the task bar at the bottom of the screen, and then it flicks back, there see. But I was in no rush to go and look at it but I knew it was there when I wanted it. And now I have pressed Alt Tab to switch between the windows.

Extract 6.7: Participant C, multiple windows

This conscious activation of another browser window while browsing was only observed in the browsing sessions of three of the expert participants during the observational study. The reasoning for this was always given as a means to circumvent temporal delays in the data transfer process. Extract 6.7 shows one such temporal reasoning. This was a very complex behaviour that had an impact on the communications, but when questioned on how the multiple browsers would affect the performance of the communications the participants were unsure and seemed prepared to accept a reduction in communication performance in favour of always having information available to read.

As well as the user-led action to produce multiple windows, a hyperlink could also be downloaded into a new window due to the HTML of the referencing hyperlink. An HTML document author could force a browser to generate a new browser for a particular referenced resource. Whenever this behaviour was observed, participants were always open to confusion. Even the participants who opted to use multiple browsers by using the right mouse button, were often unaware that a new window had been spawned.

6.5 Summary

The aim of this investigation was to highlight the usability problems that users face and to discover the important features of the browser that users employ to reveal more information about the system. The interaction between the user and the WWW revolves around hyperlink activation and the data transfer process. WWW use is described by a cyclic model that divides the interaction into three phases: assessment of the value of the referenced resource; the data transfer process; and evaluation of the download process and the Web page's content. These three phases are delimited by hyperlink activation and by the completion (or interruption) of the data transfer process. Figure 6.7 shows a diagram of this complete cyclic model of WWW use.

An important finding of this investigation is that a significant amount of a user's browsing session is taken up with the data transfer process (on average 46%). During data transfer, users continually assessed whether it is appropriate to interrupt the process. Experienced users often employ an interrupt or reload action when they feel that the delay is due to a disturbance. Experienced users feel that they are regaining *control* by interrupting the download. In contrast, the less experienced users are very wary of interrupting the download. They are unable to interpret what is actually happening and they are not confident that their judgement, that system disturbance had occurred, is correct. This does not mean that less experienced users are not sensitive to retrieval delays, only that they are unsure whether the delay is normal behaviour or a disturbance. It also implies that the use of the interrupt actions is a positive function of expertise.

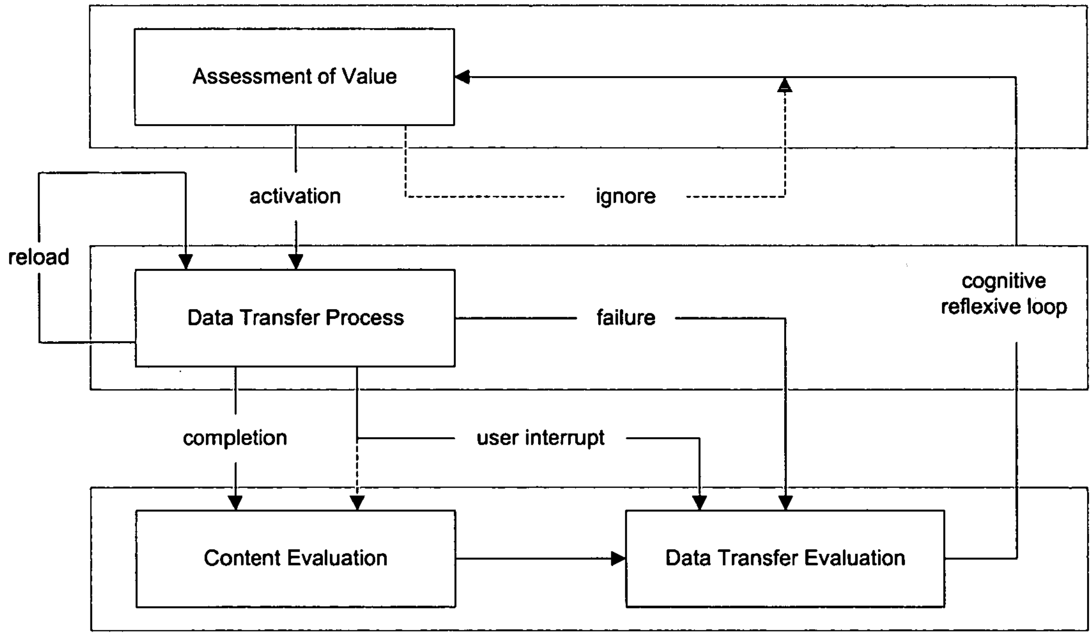


Figure 6.7: The cyclic model of WWW use.

Two models to describe the WWW have now been presented in this dissertation. The first is the three-domain model of hypertext, which can be used to categorise characteristics and properties of the system. The second is the cyclic model of interaction that revolves around the activation of hyperlinks. Taken together these two models describe the system and its use. To describe the WWW fully, the interface must provide users with knowledge of the information, space and time issues of the referenced Web page **before** hyperlink activation to help them to assess the value of the Web page in their particular circumstance. The interface must also provide users with knowledge concerning the information, space and time issues of the Web page **during** data transfer to help them cope with the unpredictability and inconsistency of this process. Finally, users must be able to evaluate the information, space and time issues of the Web page **after** data transfer in order to inform their assessment of the value of any further hyperlinks. Presently, the browser does not provide all of this information and this results in usability problems. By presenting information to the user on different levels of abstraction, support for improvisation can be enhanced and usability improved.

During the assessment of the value or relevance to the user of a referenced resource, a variety of methods are employed. Despite the transient nature of WWW information, if any prior knowledge of the referenced resource's content is known, users assume that the content is still the same. In general, however, users have no previous knowledge of a resource and so are left to assess the value and relevance of a referenced Web page using information provided in the referencing Web page, the activating object, and most interestingly, the URL of the HTML document. Throughout the ethnographic investigation, users employed a number of varied uses for the URL to justify their reasoning for hyperlink activation. The extensive use and the variety of explanations that stemmed from the URL, imply that it is a very important part of the browsing activity. Factors such as the referenced resource's location, type, content, and download time are all at some stage predicted directly from the URL. What exactly users can extract from the URL is not very clear and it was, therefore, highlighted as one of the salient interaction issues that must be investigated further: Chapter 7 is dedicated to this end.

Chapter 7

The Universal Resource Locator

The observational investigation presented in Chapter 6 revealed that participants often used the URL of the HTML document as a source of information concerning the content of the referenced Web page and the potential quality of the data transfer process. When participants were being interviewed and asked to explain the actions they performed and the understanding they possessed, the URL was often used as the evidence to support their explanations. The argument that this dissertation has presented centres on the suggestion that the abstractions of the Web page and the hyperlink conceal the mechanisms that determine the behaviour of the WWW and therefore hinder usability. The URL, having been designed as part of the underlying technology but prominently displayed in the browser, provides a window through the abstractions onto the mechanisms below. This is why the URL has become such an important element in the interaction between the user and the WWW: the URL provides a method to circumvent the hyperlink abstraction and has the potential to provide information about the referenced resource and the quality of the data transfer process that is otherwise obscured.

This chapter isolates the URL from the browser and presents a detailed examination of its make-up and properties. From this investigation, it becomes clear what system information the URL really does posses and what users should be able to confidently and consistently extract from it: what the URL *affords* the user about the referenced Web page and the data transfer process. This is followed by the presentation of an investigation undertaken to reveal firstly, if users are able to consistently extract and confidently use this information and secondly, if users ever extract information that is erroneous and potentially misleading.

7.1 Returning to the URL

In an article published on the WWW, Nielsen (1999) proposed that the URL was an important part of using the WWW. He argued that in principle, since the URL was a "*machine-level addressing scheme*" it should not have to be used or known by the user. Nevertheless, in his opinion users did use navigation mechanisms that involved "*exposure to raw URLs*". In recognition of the URL's contribution to usability, Nielsen argued that in order to increase the usability of the whole site, the designers of Web sites should use URLs that conformed to a number of properties. The URL should be short, easy to type and have an easy to remember domain name. It should be consistent in visualising the Web site's structure and be "*hackable*", allowing the user to move to higher levels of the information architecture by removing the end of the URL. Finally, the URL should be persistent and never succumb to linkrot. Although Nielsen's argument was backed only by informal observation and introspection, it would appear to agree, at least in principle, with the findings presented in the previous chapter; the URL is an important element in the interaction between the user and the WWW.

The URL is a formatted string that is used to locate a WWW resource. It identifies the protocol for communication (for the WWW this is the HTTP), a computer connected to the Internet and, the location within this computer of a WWW resource. The main constituents of the URL were introduced in Chapter 3 and they are briefly reiterated here.

- **Scheme** - specifies the protocol to be used to access the resource
- **Host** - specifies the computer in which the resource is supported
- **Path** - specifies the location of the resource relative to the host using a Unix based file naming scheme
- **Query** – server dependent operation

The rules that govern how these parts are combined to make a legitimate HTTP URL were specified and illustrated in Figure 3.9. Some of these parts are now investigated in a little more detail to show first, the technical reality of the URL and second, to show the constraints placed upon its construction that reveal certain features about the resource and, in particular, the computer the resource is stored on. Figure 7.1 shows the terms used in this discussion to describe the different parts of the URL and how they are combined in the construction of the URL.

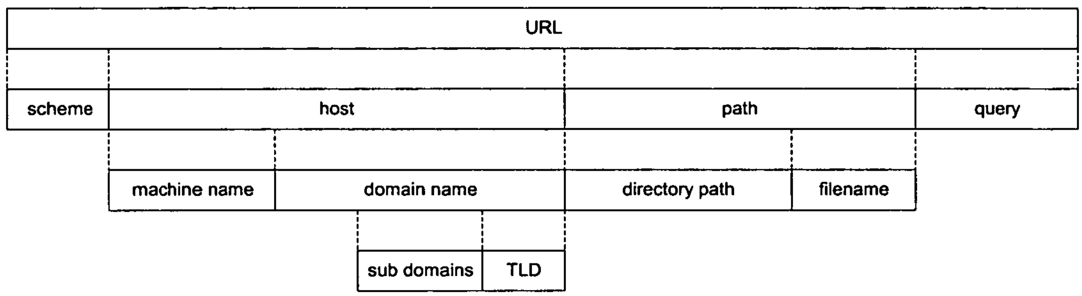


Figure 7.1: The constituents of the URL

7.1.1 Scheme

The scheme indicates what communication protocol is to be used to locate the referenced resource and how data transfer is to be attempted. Although the URL can be used by a number of different protocols, this work confines itself to the Hypertext Transfer Protocol. It is the HTTP URL that is of interest and this means that all the URLs that were studied in the work that led to this dissertation started with the string '`http:`'. As the HTTP dominates the network communications involved with the WWW, the '`http:`' scheme has become very common. In fact, it is so common that most modern browsers now assume it, and will add it to HTTP requests if it is absent from the user's input[1]. This feature of the modern browser not only demonstrates the dominance of locating and transferring resources using the HTTP, but also shows a realisation by the browser developers that the URL is used by WWW users.

7.1.2 Host

The host identifies a computer connected to the Internet with which communication takes place using the protocol specified in the scheme. The actual relationship between the host

[1] Netscape Communicator, Microsoft Explorer, Hot Java, Lynx

and a computer it addresses is rather complex and the next section explains why the host does not uniquely identify a computer, and a computer is not uniquely identified by a host.

7.1.2.1 The non-functional relationship between a host and an IP address

The DNS was introduced in Chapter 3, where it was explained that when it was given a host name it returned an IP address. One of the properties of the DNS is that it can be set up to return one from a set of different IP addresses when requested with the same host name. This facility is used by the DNS to balance the load of the network traffic and is often implemented by the DNS cycling through a list of IP addresses as each request for a particular host is received. The general philosophy is that each of these IP addresses corresponds to machines that contain exactly the same information, often called mirror sites. Every HTTP request to each of these machines will return an identical copy of the same WWW resource and the network load balancing will be hidden from the user. There is, however, nothing in the technology to stop each of these machines containing different information and therefore serving different HTTP responses. The result of this property is that, technically, the relationship between host name and IP address resolved by the DNS is **non-functional**.

There is also nothing to stop two host names resolving to the same IP address. In fact, this is often actively encouraged for two reasons. First, it can be used to implement a number of virtual servers on the same physical machine that will naturally result in a cost saving. Second, it allows a number of URLs to advertise the same Web page. This results in two URLs with different host names resolving to the same WWW resource. This is not as problematic as a host name resolving to a number of IP addresses but it does show that relationship between host name and IP address is many-to-many (Figure 7.2).

To further complicate the resolution of host names to IP addresses through the DNS, there is also the distributed nature of the DNS that may result in a delay when the DNS records have to be updated. As host names and IP addresses change, the DNS must be updated. The DNS is hierarchical and globally distributed, and there will be a delay in updating the whole system. This delay can be quite considerable (*i.e.,* days) and so after a change, the same host name may resolve to different IP addresses depending on which part of the DNS hierarchy is actually used. Requesting a WWW resource using the same URL, but from WWW clients connected to the Internet in different places, can result in different WWW resources being served even though the requests were sent at the same time.

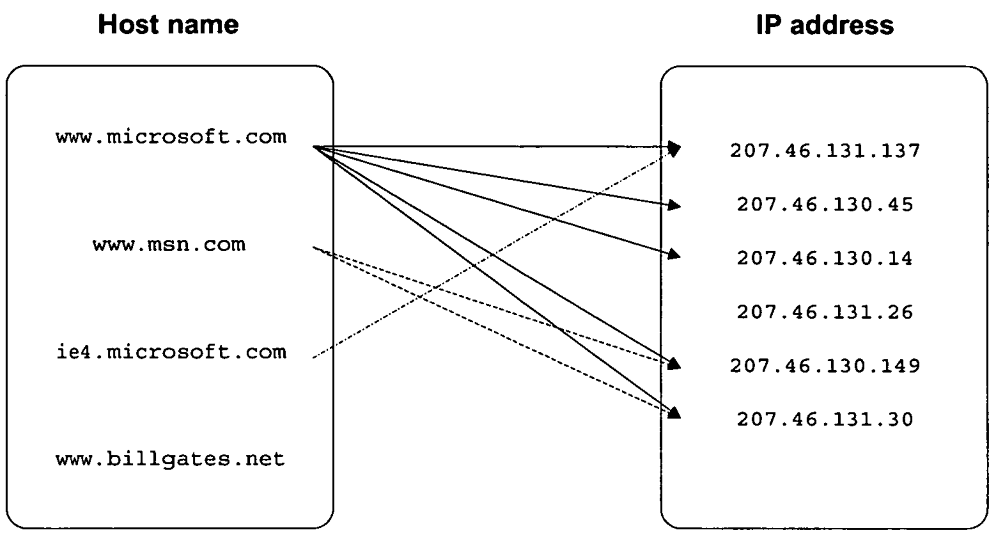


Figure 7.2: An example relationship between host names and IP addresses

To summarise, although the host is often considered as the human-readable address of a computer connected to the Internet, it does in fact only provide an address to one of a number of computers at any one time, and it may not be unique. The host's relationship with a computer and subsequently the URL's relationship with a WWW resource have the following properties: non-functional, non-unique and temporally variant.

7.1.2.2 Domain name

The string that makes up the host can be split up into two parts, the *machine name* and the *domain name,* with both parts being separated by a dot. Traditionally, the machine name for WWW servers has been the string 'www' to distinguish them from other Internet applications. As the URL has become dominated by WWW use this *de facto* convention has eroded and the machine name has started to be replaced either by an empty string or by a string that has the potential to be descriptive about that particular machine or the information it stores.

The domain name is the part of the host that identifies the institution or owner of the machine described by the machine name. The domain name space is theoretically implemented with a hierarchical structure with the boundary between the hierarchies also being a dot. The right most part of this hierarchy (the now famous dot com) is called the top-level domain name (TLD). The domain name is the part of the URL that must be registered with the Internet community, as the string must uniquely identify a domain.

The governance of the domain name has changed from a single body, called the Internet Society, to the Internet Network Information Centre (InterNIC), which set up regional registries to deal with the registration of domain names from different parts of the world. This distribution of registration led InterNIC to take the decision to split up domain names based on the country of registration. Any domain names registered from the United Kingdom's regional registry, called Nominet UK, must therefore, have the TLD '.uk'. In fact, every country that appears in the ISO 3166 list of internationally recognised two-letter country codes is eligible for top-level domain status and their two-letter country code is used as their TLD[2].

The first problem that arises from this system is that the rules, which each registry enforces by way of a contract with anyone wishing to register a domain name, are decided locally by that registry. Some registries might enforce that an applicant can only register a domain name with their countries designation if it has been established that the place of business is in this country. Others registries may not. It is not at all clear, therefore, what TLDs an applicant is allowed to own. The general rule of thumb is that there are five general TLDs (.com, .net, .org, .edu, .int) that can be directly registered from the North American registry with no geographic restrictions (but with the possibility of other restrictions). All the rest (.gov, .mil, two-letter country codes) **may** have some form of geographic restrictions though often they do not.

7.1.2.2.1 Sub domains

Under the TLD, the hierarchy has generally been flat. There is, however, a wide variation in the structure of the domain name as again, this is controlled at the local registry level. In

[2] As it happens the UK is the only exception to this rule as it does not appear on this list and is in fact only partially represented as Great Britain (GB). England, Scotland, Wales and Northern Ireland are all presently encompassed within the .uk domain name although there is currently a campaign from the Scottish to have its own top-level domain (.sco) but the present system does not allow it, http://www.scot.net/policy.html. The crown dependencies of the UK do, however, get their own domain names as they appear in the list of country codes. Isle of Mann - .im. Jersey - .je. Guernsey, Alderney and Sark - .gg.

some registries, the second-level domain names are based on generic categories (like Australia[3] and the UK) in others they may be based on political geography (like the USA[4] and Sweden[5]).

Table 7.1 shows the complete list of second-level domain names within the UK. There are various rules that must be obeyed to have a domain name accepted into each category and there are a number of offices that verify the requests for a domain name within each. In addition to these Internet community rules, there may also be some legal constraints. In the case of UK registered domain names, these legal restrictions have been dealt with under English law. For example, the '`.ltd.uk`' domain was intended to allow a unique mapping between a company name (registered under the Companies Act with Companies House) and a domain name. It became apparent, however, that if a UK public company used this domain name they could be committing a criminal offence under the 1985 Companies Act[6]. The remedy was the introduction of the '`.plc.uk`' domain. The choice of a domain name has the potential, therefore, to be restricted not only by the registry itself but also by the laws of the country of the registry.

---

[3] The Australian sub domains are listed at http://www.aunic.net/policies.html.

[4] The United States uses a hierarchy of the form `<entity-name>.<locality>.<state-code>.us` (RFC 1480) although the uptake of this standard is very small as the generic TLDs still dominate.

[5] The Swedish registry (NIC-SE) introduced a policy in January 1998 which permitted only corporated companies to use the .se TLD. All other enterprises had to register under a domain name that included a second-level domain province code. The aim was to prevent conflict between companies with similar names. The result however, was a 40% drop in registrations for NIC-SE and an abandonment of the .se country code. Companies turned to the generic .com, .net and .org TLDs and started to use other country codes where restrictions on registration were minimal (InterNet News, 1998).

[6] Section 33 of the Companies Act 1985 states:
1) A person who is not a public company is guilty of an offence if he carries on any trade, profession or business under a name which includes, as its last part, the words "public limited company" or their equivalent in Welsh ("cwmni cyfyngedig cyhoeddus").
2) A public company is guilty of an offence if, in circumstances in which the fact that it is a public company is likely to be material to any person, it uses a name which may reasonably be expected to give the impression that it is a private company.

Zones	Second-level	Top-level	Example
Commercial	CO	UK	www.electronic-media.co.uk
Non-Commercial organisation	ORG	UK	www.y2k.org.uk
Limited Company	LTD	UK	www.plastica.ltd.uk
Public Limited Company	PLC	UK	www.scottishpower.plc.uk
Academic	AC	UK	www.dcs.ed.ac.uk
Government	GOV	UK	www.inlandrevenue.gov.uk
Schools	SCH	UK	www.shsb.essex.sch.uk
Ministry of Defence	MOD	UK	www.mod.uk
National Health Service	NHS	UK	www.cmht.nwest.nhs.uk
UK Internet Networks	NET	UK	www.inter.net.uk

Table 7.1: Top and second level domain names within the UK.

7.1.2.2.2 Third-level sub domains

Third-level domain names may also be necessary particularly when there is the potential for the existence of a large number of hosts wanting the same domain name. For example, in the UK there are numerous cases of schools sharing the same name. To resolve which school gets what domain name a further level of hierarchy is needed below the '`.sch.uk`' domain. This third-level of domain is achieved in England and Wales by using an abbreviation of the local education authority area, then a fourth-level domain representing the school name (for example, shsb.essex.sch.uk is the domain name for the Southend High School for boys located in the Essex education authority). The same principles for third-level domain names are employed in the NHS and academic domains[7].

7.1.3 Path

The part of the URL that is presented to the server (identified by the host) is the path. The path is constructed using the Unix convention for file systems, which usually results in the

[7] Failure to have sufficient sub domains means that there will be a high chance of conflict in domain names. To get round this the rules on domain name assignment are being stretched. The URL `http://www.dateline.uk.com/` shows one solution to this problem. This reorganisation of the domain hierarchy is potentially in breach of the guidelines but domain registries are now offering a whole multitude of alternatives, for example, ProWeb http://www.couk.com/ offers a selection of alternatives for the UK including www.<your-name>.ukltd.com, .ltduk.com, .ukplc.com, .plcuk.com, .ukco.com and .couk.com.

path containing a directory path and the name of the resource separated by the "/" character. The name of the resource often has a file extension giving some indication to the client browser, and potentially the user, of the resource's type (though this may also be explicitly encoded elsewhere). The syntax for the path within the HTTP URL is the usual Unix construction with directory hierarchy separated with the "/" character.

If the path is empty, the server has an opportunity to supply a default resource. Similarly, if the resource name is empty, but a directory path is present, the server can also supply a default resource from that directory. This feature introduces an interesting behaviour of the browser. If only a directory path is present the path string should end in a "/" character. If this is the case, the browser will request the WWW resource and will subsequently receive the default resource from that location. If, however, the "/" character is not present the path is in fact incorrectly specified, and the server will believe that the client is requesting a particular resource that is not in the directory and so return a *"file not found"* error. Omitting the last "/" character is such a common mistake, however, that if this error is returned, browsers will automatically add the "/" character to the end of the URL and re-send the request giving the illusion that the initial URL was valid. Take the following URLs as an example.

- `http://www.bbc.co.uk/sport/cricket`
- `http://www.bbc.co.uk/sport/cricket/`

An HTTP request containing the first URL will ask the server for a resource named cricket from the directory sport. The second requests a default resource from the directory sport/cricket/. If only the second option is valid, most modern browsers will render the same resource (the default resource from sport/cricket/) without the user being aware of their mistake. This clearly has an impact on the network traffic and potentially its performance as the number of HTTP requests and responses is increased.

7.1.4 Query

The query is the part of the URL that can be used as an input for the resource specified in the path. The query can contain information like the user's name or the location of the previous Web page. This information can then be extracted from the query, processed by the resource, and used to personalise the Web page presented to the user.

7.2 The use of the URL

It has already been proposed in this dissertation that the typical WWW user assumes that there is a one-to-one relationship between the URL of an HTML document, and the Web page that this document defines. This was clearly illustrated in the drawings produced in Chapter 5 and further demonstrated by the explanations presented in Chapter 6. Put more simply, users associate a single URL with a hyperlink and so believe that a URL references a Web page (Figure 7.3). This belief is encouraged through the hypertext model presented to the user through the interface. A single URL can be entered in the browser's address bar to retrieve a Web page and a single URL appears at the bottom of the browser when the activating object of a hyperlink is moved over by the pointer. It appears to the user that the URL is the main mechanism in the request and retrieval of a Web page and users therefore believe that the URL, in some way, defines the Web page. It is the WWW user's understanding of this relationship that is of interest to this work and forms the basis for the experiments presented in the next sections.

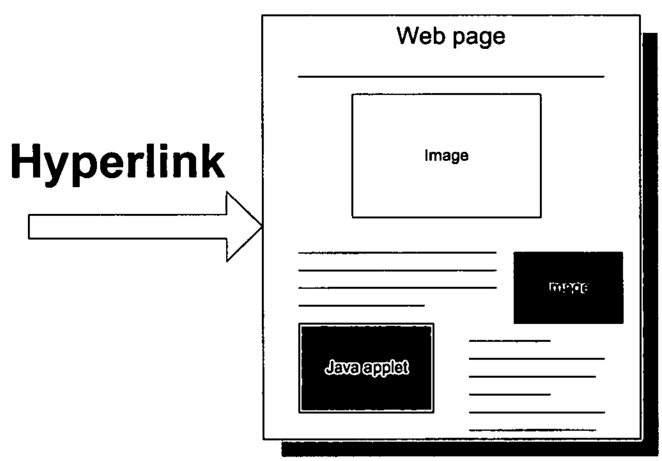


Figure 7.3: The hyperlink model, one hyperlink, one URL, one Web page

The technical reality of the URL is, however, that it does in fact only reference a single WWW resource. A Web page is a collection of these resources and so may involve a number of different URLs (Figure 7.4). Furthermore, the way the URL references a resource has a number of properties that would appears to contradict the WWW users' belief that one URL references one Web page. These properties are:

- **Non-functional** – The DNS may resolve a host to different IP addresses.
- **Non-unique** – Different URLs may reference the same resource.
- **Temporally variant** – The distributed hierarchical nature of the DNS means that different DNSs may contain different tables at the same time.

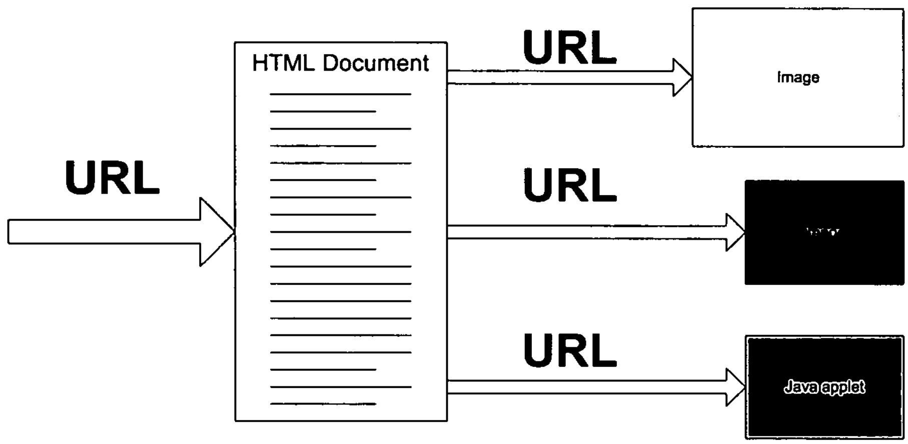


Figure 7.4: The technical reality of the Web page and its associated URLs

The final point about the URL that needs to be kept in mind is that it is made up of a number of distinct parts. As has been explained, the host contains a hierarchy that is delimited by dots and the path contains a file structure separated by the "/" character. The three major parts (scheme, host and path) are also separated by the "/" character. The result is that the signature of a URL is a string with sub strings separated by dots and slashes. It is therefore, reasonable to assume that understanding which sub string belongs to which part of the URL takes considerable expertise.

Although it has been shown that the URL has the potential to contain considerable amounts of information about the referenced resource's location and content, the question that now is investigated is: how (if at all) is this information extracted by the WWW user; what use do they make of it, and why? The rest of this chapter presents the findings of two laboratory experiments concerning the URL, what information users can extract from URLs, and how confident and accurate they are with their predictions. The first is entitled *the URL affordance experiment* and concentrates on determining **what** information is extracted from the URL. The second experiment, entitled *the URL delay experiment* focuses on the data transfer process by introducing an artificial delay and seeing how this affects participants' ability and confidence in extracting this information.

7.3 URL affordance experiment

The prime objective of this experiment was to determine what information WWW users could uncover from a URL concerning the Web page that it *indirectly* references. The basic hypothesis was related directly to the results of the focusing study presented in the previous chapter: the WWW user can use a URL to extract a rich set of information related to the whole Web page that it indirectly references. A corollary to this was that WWW users believe that the URL references, and potentially describes, the entire Web page entity and not just an HTML document.

This experiment also aimed to elicit the *type* of information that WWW users could extract about the referenced Web page, the data transfer process and, the system as a whole. To help accomplish this aim three different types of information are highlighted. These types fall into categories that have already been presented in this dissertation as information, time and space. The experiment aimed to uncover how, if at all, the WWW user extracted the following types of information from the URL.

- **Content** of the implicitly referenced Web page,
- **Temporal properties** of the ensuing data transfer process,
- Physical **location** of the data that has to be transferred.

7.3.1 Participants

As in the previous investigations, participants were unpaid volunteers, recruited from under-graduate and post-graduate students at the University of Edinburgh. The set of participants for this experiment contained seventeen members, six of whom claimed to be expert WWW users, two intermediate WWW users and, nine who claimed to be novice WWW users (Table 7.2). Every participant had used the WWW previously.

Participant	Gender	Age	Expertise
A	M	23	Expert
B	M	25	Expert
C	M	26	Expert
D	M	27	Expert
E	F	21	Novice
F	F	21	Novice
G	M	22	Intermediate
H	F	21	Novice
I	F	21	Novice
J	F	23	Novice
K	F	18	Novice
L	M	22	Intermediate
M	M	26	Expert
N	F	20	Novice
O	F	21	Novice
P	F	22	Novice
Q	M	20	Novice

Table 7.2: Participants' gender, age and expertise

7.3.2 Data collection

Participants were presented with ten different URLs one at a time, and asked to explain what they could predict about the referenced resource. To ensure that all predictions were based solely on the URL (internal validity) participants were shown each URL as text printed on a piece of paper. It was only after a participant had completed their prediction, that the URL was used with a browser to request a WWW resource for display.

a	**http://www.lmnet.com/~bobh/shakeit.avi** An audio visual movie of a baby dancing
b	**http://www.national-lottery.co.uk/game/results.html** The United Kingdom's national lottery results
c	**http://www.stfturist.se/engleska/index.htm** The Swedish tourist federation home page in English
d	**http://www.useit.com/alertbox/980614.html** Jackob Nielsen's alertbox for the 14th June 1998, an article entitled Fighting Linkrot.
e	**http://www.corrsweden.org/links.htm** The interesting links page from the Corrs (an Irish pop band) Swedish fans Web site in English.
f	**http://www.dcs.shef.ac.uk/~wmlh/ipcat97book.htm** Proceedings of the Information Processing in Cells and Tissues conference in 1997 authored by Professor Holcombe at the department of computer science, University of Sheffield, United Kingdom.
g	**http://home.netscape.com/newsref/pr/newsrelease558.html** A news release from the Netscape Corporation.
h	**http://www.telegraph.co.uk/et?ac=000200771016135&rtmo=a4KJddWJ&atmo=99999999&P4_FOLLOW_ON=/98/10/9/sgadc09.html&pg=/et/98/10/9/sgadc09.html** An article from the sports golf section of the electronic telegraph published on 9th October 1998 about the Alfred Dunhill cup.
i	**http://cello.cs.uiuc.edu/cgi-bin/slamm/ip2ll/** A form that allows one to uncover the longitude and latitude of a machine from its IP number. Published by the computer science department of University of Illinois, America.
j	**http://www.kl98.com.my/gamesex/default.html** The homepage containing a shockwave animation for the 1998 Commonwealth games exhibition held in Kualar Lumpur, Malaysia

Table 7.3: List of URLs for the URL affordance experiment.

After data transfer had been completed, a participant was given time to analyse the resource and then asked to judge their initial prediction and whether, retrospectively, they could now extract any more information from the URL. All utterances were recorded and the time it took to complete each experiment ranged from 16 minutes to 48 minutes. This resulted in the collection of over six hours of audio data. All of this audio data was subsequently transcribed for analysis.

The set of URLs used in the experiment was selected by randomly extracting 25 URLs from the WWW browser cache of the author's desktop computer. From this set, ten were chosen which it was felt provided a good range of URLs. The criteria for this judgment were that there was a number of different TLDs and a range of path lengths. The ten that were chosen are presented in Table 7.3 with a textual description of the Web pages they address. In an attempt to reduce the effect on the results of participants learning during the experiment and becoming more confident and accurate with their later predictions, the order in which the URLs were presented to the participants was pseudo-randomised. There were three different sheets of paper each presenting the URLs in a different order; these three lists were used in a strict rotation.

During a participant's predictions of the referenced resource from the URL, an experimenter talked with the participant. A concurrent verbal protocol method was used to encourage any taciturn participants and investigate the participant's broader understanding of concepts used specifically during the prediction. If participants struggled to provide any information concerning the referenced resource, the experimenter encouraged participants to make predictions relating to the referenced resource content, location and the potential quality of data transfer.

7.3.3 Results

When forced to use the URL to predict information about the referenced resource *before* hyperlink activation, participants were often able to supply surprisingly rich and detailed forecasts. Not only could participants extract information from the URL concerning the referenced Web page's general content, but also showed impressive ingenuity to reveal more subtle information, like for example, the date of Web page authorship, the type of server communication would take place with, the name of the author and the topology of the network. Extracts 7.1 and 7.2 show examples from expert participants' predictions that, as a whole, form a rich and accurate picture of the referenced resource.

http://www.dcs.shef.ac.uk/~wmlh/ipcat97book.htm
I: What can you tell me about this URL? S: Well I can tell you that it is a Sheffield University, and it is in someone's personal home directory because it has a tilda for Unix account, home directory, and there is some acronym 97 which presumably means that it is the proceedings or something to do with the year 1997, that is what I would assume I don't know that it will be yeah. Book indicates that it will be, might be an online version of a document that is found in another form somewhere else, like a paper document maybe and htm indicates that it is hypertext. I: What about download time, can you predict anything? S: No, I suspect, well my intuition would say that book htm could be fairly fast because it could be a big text file representation of a book, yeah, but then again it doesn't really mean anything, it is just me implying that.

Extract 7.1: a rich set of predictions from expert participant D. In particular, the location of the resource (Sheffield), the type of the resource (Hypertext), the year of publication (1997), the author and institution (wmlh from the University of Sheffield), the type of server (Unix) and when pushed, a guess about the data transfer process implied from a prediction about the type (fairly fast due to it being a text file).

http://www.lmnet.com/~bobh/shakeit.avi
S: Ermm, it looks like a company page, it could be, I M net makes me think that it might be an ISP but I'm not sure, it might just be a company IM networks or something. Again, it looks like it is in somebody's own space so it could be somebody who subscribes to that ISP or it could be one of the employees of the company. And shakeit.avi is an avi movie, shake it means that it could be somebody shaking their bottom or something maybe a dancer? I: Well OK, S: it will be slow I: why is it going to be slow? S: ermmmm, because dot com makes me think American company and also avi's are normally enormous

Extract 7.2: a rich set of predictions from expert participant B. The location of the resource (America), the type of resource (avi movie file), the author and institution (both perhaps from IM networks) and, the quality of the data transfer process (implied from the American location and the type of the resource).

As expected, the confidence and diversity of predictions grew in proportion to expertise. Expert participants consistently provided confident explanations that contained information concerning the informational content of the Web page, the location of the data and, the temporal properties of the data transfer process. In contrast, novice participants tended to be unable to make any predictions about the data transfer process and were left with very general predictions about content or location.

7.3.3.1 Type of predictions

The two extracts presented so far (7.1 and 7.2) show a number of different types of predictions concerning information about the referenced resource's content, the location of the server, and the time taken to transfer the data to the client. These predictions are rich in information due to the ability of the participants to extract information (a reflection of their expertise) and the construction of the URLs themselves. The richness and accuracy of participants' predictions varied widely throughout the experiment although, it was very unusual for a participant to be unable to predict anything[8]. Analysis of the transcripts revealed thirteen different types of predictions commonly used by participants. These thirteen different types of explanations that participants used in their predictions are presented in Table 7.4 with a reference to an example.

What the results of this experiment highlight are that two factors affected the predictions that participants gave. The first was the structure and content of the URLs themselves, and the second was the expertise of the participants. Different URLs provide different types of information that is dependent on their construction, the information contained in their structure and, in the meaning that participants associate with each sub string. In addition, different participants bring different amounts of knowledge and experience with them and this, not surprisingly, appears to affect their ability to extract information from the URL. The findings presented in Table 7.5 show the different types of information that participants could extract from each URL: how the URLs construction and content affects predictions. The findings presented in Table 7.6 show which participants could elicit which types of predictions: how expertise affects the predictions.

Prediction	Explanation	Example
Contents	A prediction concerning the contents of the Web page.	Extract 7.5
Type of resource	The type of the resource referenced by the URL.	Extract 7.1 and 7.2
Author	The author of the Web page	Extract 7.1 and 7.2
Institution	The Institution or organisation that published the Web page	Extract 7.1, 7.2 and 7.5
Date of Publication	The date the Web page was published	Extract 7.1
Language of Web page	The language that the Web page is published in.	
Country Location	The country where the Web page is located.	Extract 7.1 and 7.2
Specific Location	A more specific location than country, like a city or area.	Extract 7.1
Delay implied from resource	A prediction concerning the download time based on a prediction about the resource.	Extract 7.1, 7.2 and Table 7.7
Delay implied from location	A prediction concerning the download time based on a prediction about the location of the resource.	Table 7.7
Delay implied from network topology	A prediction concerning the download time based on an understanding of the network topology.	Table 7.7
Delay implied from usage	A prediction concerning the download time based on a prediction of how many people are using the Internet at that moment.	Table 7.7
Type of server	The type of operating system that the server of the referenced resource is running.	Extract 7.1 and 7.2

Table 7.4: A list of predictions used by participants

7.3.3.2 URL affordance

The first important feature of Table 7.5 is that there are four types of predictions that were frequently used by participants when making predictions about the majority of the URLs. These were (in order of URL coverage) the country location, the type of the resource, the institution that published the resource and the content of the Web page itself. The fact that the content of the Web page is so frequently predicted is due to it being explicitly asked for in the experiment. The content of the referenced Web page is probably the most important single type of prediction that can be made about a resource and although it appears from Table 7.5 that the URL affords this type of information, it is perhaps misleading. In general, participants were very unsure when predicting what the content of the referenced resource was going to be. The result was that predictions about a Web page's content were unspecific and generally poor. The other types of predictions were not always explicitly asked for

[8] This is a further example of Norman's demand structure that was introduced in Chapter 4: participants are compelled to answer the interviewer's question even though it was generated on the

during the experiment and their existence is due to participants revealing this information of their own accord.

	URLs									
	a	b	c	d	e	f	g	h	i	j
Contents	X	X	X		X	X	X	X		X
Type of resource	X	X	X	X	X	X	X	X		X
Author	X					X				
Institution	X	X	X	X		X	X	X	X	
Date of Publication				X		X		X		X
Language of Web page			X		X					
Country Location	X	X	X	X	X	X	X	X	X	X
Specific Location						X			X	
Delay implied from resource	X					X			X	
Delay implied from location			X			X	X			
Delay implied from network topology						X				
Delay implied from usage						X	X			
Type of server	X		X			X	X			

Table 7.5: Participants' types of predictions for each URL

	Participants																
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
Contents	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Type of resource	X	X	X	X			X	X				X	X				X
Author	X	X		X				X				X	X				X
Institution	X	X	X	X	X	X	X				X	X	X	X	X	X	X
Date of Publication	X	X	X	X	X			X			X		X	X	X		
Language of Web page		X	X	X					X				X	X	X	X	
Country Location	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Specific Location	X	X	X	X	X	X			X	X		X	X		X		X
Delay – resource		X	X	X			X					X	X				
Delay – location	X	X											X				
Delay -network topology		X															
Delay - usage		X											X				
Type of server	X			X													

Table 7.6: The use of different types of explanations in participants' predictions (expert participants highlighted in grey)

spot and not what they really believe.

Predictions concerning the location of the resource were made with considerable confidence. It has been shown that the country code TLD explicitly provides some form of spatial information, although this has the potential to be misleading. The country location of a resource is the only type of prediction that was always available from the URL and it was the information that all participants appeared able to predict (Table 7.6).

The type of the resource was another form of prediction extracted from every URL, with the exception of one. The interesting fact is, however, that the one URL that did not afford any resource type predictions was the URL that did not contain a filename. This would suggest that the filename part of the path afforded the participant the information concerning the type of the resource. This is, of course, not surprising as the filename extension has already been highlighted as an area where the type information of the resource is encoded.

URL	Extracts
http://www.stfturist.se/engleska/index.htm	"there is no reason why it could not be quick as there is a quick link to Sweden"
http://www.dcs.shef.ac.uk/~wmlh/ipcat97book.htm	"I would expect it to be very quick as we are all on the academic network, and academic sites tend to be quick and I imagine not a lot of people would look at that file."
http://www.lmnet.com/~bobh/shakeit.avi	"avis are normally enormous" "avi file, so it will take time"
http://cello.cs.uiuc.edu/cgi-bin/slamm/ip211/	"will be slow because it has to run a script at the other end"
http://home.netscape.com/newsref/pr/newsrelease558.html	"will be quite slow 'cause its very congested"
http://www.dcs.shef.ac.uk/~wmlh/ipcat97book.htm	"book htm could be fairly fast because it could be a big text file"

Table 7.7: Predictions concerning the data transfer process

Where the URLs appeared to fail in being informative was in providing information concerning the data transfer process. With the exception of URL f, the URL did not afford very much information for the participants to use in predicting the delay. Table 7.7 shows some of the different types of predictions made by participants concerning the data transfer process. All the explanations about the data transfer process are implied from other properties extracted from the URL, for example, the predicted location of the server or the size of the resource. Overall, predictions about the data transfer process were scarce and often very unsatisfactory.

7.3.3.3 The effect of expertise

Table 7.6 shows clearly that the more expert participants provided richer predictions about the referenced resource. Where expert participants really excelled was in making predictions about the data transfer process, an area where the less experienced participants struggled. The principal difference in technique between the expert and the novice participants was in the ability to break the URL down into its constituent parts (scheme, host, path and query) and then use the knowledge of these parts in their predictions. The expert participants were able to apply their understanding of the URL's structure to help them confidently reveal information concerning the referenced resource. The most important distinction that users made concerning the URL's structure was between the host and path. By isolating the host, the expert participants could apply their understanding of the host's domain structure like, for example, the country code or generic TLDs to help them in their predictions. Similarly, their understanding of the file structures helped them to isolate the filename and type extension from the path.

Novice participants, in contrast, showed a very rudimentary understanding of the URL's structure. The novice participants, on occasions, used a country code in their prediction but in general, their explanations appeared to have a foundation based around a sub string of the URL. This sub string was usually a recognisable *cueword* that could appear in any part of the URL, host, path or filename. Once isolated, a novice participant's predictions centred on their understanding of this cueword.

Extract 7.3 and 7.4 show an expert participant and a novice participant breaking down the same URL. The expert participant clearly breaks down the URL and uses the host information to reveal (from the TLD) that the resource is probably going to come from Sweden, and what the company name is. The expert user then proceeds to use the path to determine some contextual information and the filename type extension to determine the type of server. The novice participant, however, does not appear to understand the structure and, for example, is unaware of the country codes to help her predict a location. She is left forming an argument for her predictions around a single cueword. In this case, the cueword is the (apparently) rather unenlightening *engelska.*

http://www.stfturist.se/engelska/index.htm
I: What can you say about this one? S: Well it is something in Sweden I think, ermm dot SE, engleska probably means English, and index.htm probably means they are using a DOS box to write the stuff on [three letter extension rather than four]. Whatever stfturist is it will be something to do with them. Company name? I: Right OK, what can you tell me about how long it is going to take to download? Can you predict anything? S: Not really, there is no reason why it could not be quick as there is a quick link to Sweden, but it all depends.

Extract 7.3: Expert participant A applying their understanding of the URL's structure to aid their prediction

http://www.stfturist.se/engelska/index.htm
I: ..., like what country it might be from? S: might be German I: OK, you think it might be German. S: because of the engleska, sounds German. I: and what about content? S: Well, I would have to guess that it is, ermmm, a German book, errmm, in English maybe. Maybe a German book about English?

Extract 7.4: Novice participant E forming a prediction around the sub string engelska

Although Extract 7.4 shows a novice participant failing to make a successful prediction, the strategy of picking a cueword and then basing a prediction on it, was often very fruitful. When attempting to predict the contents of the resource referenced by the URL `b`, for example, participants, regardless of their expertise, would immediately identify the clearly recognisable sub string `national-lottery` from the domain name. Using this cueword as the basis of their predictions, rather than any structural knowledge of the URL, minimised the advantages of expertise. As shown in Extracts 7.5 and 7.6, the novice and expert participants' explanations are very similar.

```
                http://www.national-lottery.co.uk/game/results.html

S: It is highly likely to be the UK national lottery, errm, results of the game but
it doesn't tell you which one so you would have to assume it's the most recent.
```

Extract 7.5: Participant A (expert)

```
                http://www.national-lottery.co.uk/game/results.html

S: Ummm, the national lottery results in the UK, I don't know, probably the latest
results
I:      OK
```

Extract 7.6: Participant N (novice)

Unfortunately, if a novice user extracted a cueword that was in some sense misleading, their predictions could go very wrong. A cueword may be misleading due to the participant's interpretation of the word not being related to the referenced resource in any way. The existence of an understandable cueword gives the novice participant a basis for some reasoned argument and this can result in a much-distorted prediction. For example, URL i caused confusion for the novice participants on a number of occasions, and subsequently led to unsatisfactory predictions (for example Extract 7.7).

```
                    http://cello.cs.uiuc.edu/cgi-bin/slamm/ip211/

S: err
I:   anything spring to mind
S:                           music page
I:                                      Right
S:                                           Umm Well slamm
I:                                                          slamm what is slamm? don't
know
S:   No, umm does it actually give you pictures, I don't know
I:                                                            why do you say pictures?
S:  I just did, how about buying cellos page, or maybe how to play a cello.  Well it
is definitely something to do with cellos.
```

Extract 7.7: Participant J (novice) employing the cueword cello.

In contrast, expert participants tended not to make a similar mistake with this particular URL, as they were able to use the URL structure to eliminate the cueword '`cello`' from their predictions, understanding that it was the name of a machine. It did not, therefore, have a high probability of being related to the content of the referenced resource. The expert participants used the other parts of the host and the path to base their prediction on (Extract 7.8).

http://cello.cs.uiuc.edu/cgi-bin/slamm/ip211/
S: Right, well cs means computer science usually, U I U C is something like university of California, edu means ermm probably American university errrmmm, so computer science U I U C. Cgi-bin probably means it is a cgi-bin script, slamm probably means the person can't spell and I P two hundred and eleven that's totally meaningless. The fact that it doesn't have with a dot at the end suggests that they got a phtml index file there, but that doesn't give you any useful information.

Extract 7.8: Participant A (expert)

7.3.3.4 Functional belief of URL and Web page relationship

Finally, one of the aims of this experiment was to investigate participants' beliefs concerning the relationship between the URL and the Web page it indirectly references. At no time throughout this experiment did any participant express a belief that the URL did not relate directly to the Web page that was presented to them on the WWW browser. Every participant, regardless of their expertise, was content to make predictions about the Web page.

7.3.4 Summary of URL affordance experiment

This experiment has shown that participant can make rich and accurate predictions about a Web page from the URL of its HTML document. What is also clear is that the amount and types of information that a participant can extract varies greatly between URLs. The inconsistent nature of the information contained in a URL and the fact that parts of it are null (like the path or filename) mean that users can never be assured of any reliable information about the referenced Web page.

During this experiment, participants were best at extracting information about the spatial location of the data from the URL. This spatial information tends to be located in the host part of the URL and in particular, the TLD country code and sometimes, the other domains of the host like, for example, `shef.ac.uk`. Spatial information is therefore mainly contained in the host part of the URL and it should be noted that this is the only constituent of the URL that is always present. The path and filename are not consistently available for inspection as they are often blank. It appears, however, that if the path and filename are present contextual predictions, and especially predictions concerning the type of the resource, become first, easier to elicit and second, far more accurate. Overall, the URL did not present the user with very much information concerning the data transfer process and all participants were left attempting to explain the process using implication from predictions concerning the resource's size or the server's relative location.

Expertise was directly reflected in the ability to elicit information from the URL and this again suggests that using the URL in this way must be a common strategy used while browsing the WWW: users get better at understanding URLs because they are constantly having to elicit information from them. The main difference between the expert and novice participants was in their ability to understand first, how the constituent parts of the URL where combined in the URL string and second, what each part related to: the host was the location of a computer, the path was the file hierarchy of a data store *etcetera*. All participants would try to pick a *cueword* from the URL string and use this as evidence to resolve what the content of the referenced Web page was. This strategy, although sometimes fruitful, often led to erroneous predictions for the novice participants, as the cueword they chose was frequently inappropriate. The extraction of unsuitable cuewords was minimised by the more expert participants as their understanding of the URL's structure enabled them to eliminate some potentially misleading words like, for example, the machine name.

Despite the inconsistency of the URL in providing information about the referenced Web page, participants still believed that the URLs often held clues to the referenced Web page's content and location. Even though some of the expert users had a good understanding of the technology, no participants showed any signs of not believing that the contents of the URL related directly to referenced Web page.

7.4 URL delay experiment

The second URL experiment built on the first by introducing a rating mechanism so that participants could rate how confident they were about their predictions before hyperlink activation, and how accurate they felt these predictions were once the resource had been transferred. The introduction of these ratings allowed for a more empirical analysis to be undertaken. The second URL experiment also provided an opportunity to investigate a further ten URLs which broadened the findings of the whole study and therefore increased the external validity of the results in general.

The second URL experiment also introduced an artificial delay into the data transfer process. This delay was introduced in order to investigate what effect it would have on the participants' rating of the data transfer process. The hypothesis is that the unpredictable and abstract nature of the data transfer process in normal WWW operation means that participants will not recognise this artificial delay and their explanations will adapt to its presence.

7.4.1 Participants

The same recruitment policy was used as in the previous experiment and 18 participants with various degrees of WWW expertise were used in this experiment (Table 7.8).

Participants	Gender	Age	Expertise	Artificial Data Transfer Delay (seconds)
A	M	21	Intermediate	3
B	F	22	Novice	8
C	M	19	Expert	10
D	M	27	Expert	10
E	M	20	Intermediate	10
F	M	24	Novice	10
G	M	19	Expert	0

H	M	23	Expert	0
I	M	26	Expert	5
J	M	40	Novice	10
K	M	28	Expert	10
L	M	29	Expert	3
M	F	25	Novice	0
N	M	31	Intermediate	4
O	M	26	Intermediate	2
P	F	23	Novice	8
Q	F	21	Expert	5
R	F	24	Novice	2

Table 7.8: Participants for the URL delay experiment

7.4.2 Data collection

The second URL experiment employed the same method as the first experiment with a different set of URLs (Table 7.9), but added two important modifications. First, participants were asked to rate how confident they were about their predictions, and then after data transfer, to rate how accurate they believed their predictions were. A one to six scale was used for both ratings with one being the least confident/accurate and six being the most. Participants were asked to rate their confidence in, and the accuracy of, their predictions with respect to three criteria: information, space and time. For each URL, participants would explain what they thought the referenced resource content was, where they believed the resource to be located and, how long it was going to take to download. Participants would then rate how confident they were of these predictions before activating the hyperlink and awaiting data transfer. Once the Web page was rendered the participant would then be asked to rate how accurate they believed each of their predictions had been.

After all the experiments had been completed, each participant's content predictions were shown to two independent observers. Having seen the Web pages themselves, these observers rated (on the same one to six scale) how accurate they judged the participants predictions were. The experiment therefore collected three different sets of ratings. The participants' predictions concerning the content of the referenced Web page, the location of the data and the download time; the participants' accuracy ratings of their predictions; and the independent observers' accuracy ratings of the participants' content predictions.

K	**http://uk.samba.org/** The Sun site resource centre located at imperial college London.
L	**http://www.ozsports.com.au/cricket/** A Web page of Australian cricket links.
M	**http://www.whitehouse.gov/WH/glimpse/tour/html/index.html** A picture of the USA's presidents Whitehouse which can be used to link to different Web pages about the parts of the house.
N	**http://www.inter.net.uk/** The home page for a UK based internet service provider.
O	**http://www.pota.to/** A picture of some potatoes.
P	**http://www.scottishpower.plc.uk/** The home page to Scottish Power plc.
Q	**http://www.comlab.ox.ac.uk/oucl/conferences/prg-sun/index.html** The computer laboratory at Oxford University.
R	**http://www.comlab.ox.ac.uk/oucl/people/james.worrell.html** The homepage of a student, James Worrell, studying at the computer laboratory at Oxford University.
S	**http://www.comlab.ox.ac.uk/idgp/usingz/** A page explaining some of the examples provided by the book Using Z.
T	**http://www.microsoft.com/jobs/features/feature_02.htm** An article about a particular job at Microsoft.

Table 7.9: List of URLs for the URL delay experiment.

7.4.2.1 Introducing a delay

The second modification was that the WWW browser used to download and display all the Web pages was connected to a proxy server (Figure 7.5) that introduced an artificial delay to the data transfer process. There are a number of points during the data transfer process where a delay might be introduced by a proxy server.

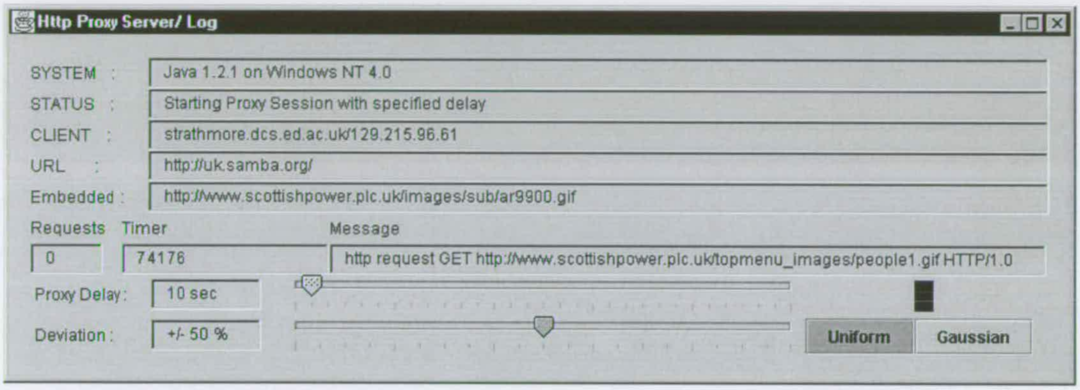


Figure 7.5: The Interface to the proxy with artificial temporal delay

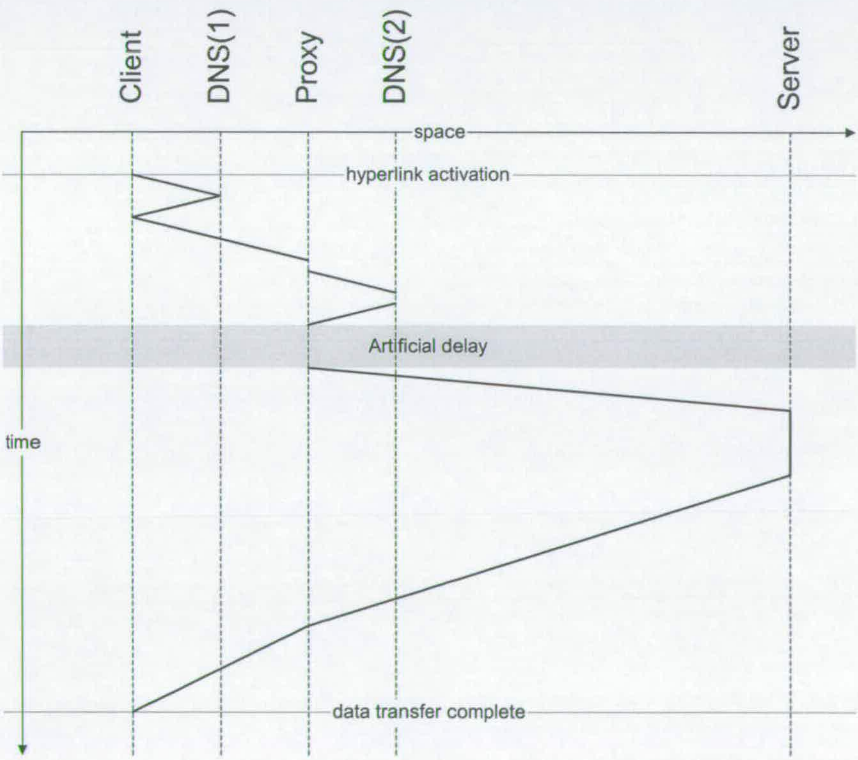


Figure 7.6: The introduction of an artificial delay

The delay had to be introduced without the participant being aware that it was artificial and so care had to be taken to produce system behaviour that was as *normal* as possible. The delay was therefore introduced before the proxy sent every HTTP request (Figure 7.6). Each participant was randomly assigned a basic delay set between 0 and 10 seconds. This delay was then used as the basis[9] to delay the data transfer for every object that passed through the proxy. The result was a data transfer process delay that was artificially high with some participants having to wait as much as 15 seconds extra per HTTP request. Although the transfer of each WWW object is achieved with a degree of concurrency, some participants waited as much as 2 minutes in addition to the *normal* system delay for a Web page to download.

7.4.3 Results

All the empirical data collected during this experiment are presented in the six tables in Appendix I. Averages of participants' confidence and accuracy predictions with respect to content, location and download time are shown in Figure 7.7. The overall mean for the participants' ratings was 4.5. The mean of participants' ratings of how confident they were about their predictions was 4.4, and the mean for participants' ratings of how accurate their predictions had been was slightly higher at 4.6. Taken as a whole, these results suggest that participants were able to make confident predictions about a referenced resource from a URL and participants considered these predictions to be accurate. Considering the normal for the ratings is 3.5, all of the mean values are significantly higher.

Looking at the breakdown of the predictions presented in Figure 7.7 immediately suggests that participants were most confident at inferring information concerning the location of the referenced resource (mean 4.9), followed by the actual content of the resource (mean 4.2) and least confident about predicting the download time (mean 3.9). This confirms the results of the URL affordance experiment that also showed the URL to regularly supply some information with respect to location and some aspects of the resource's content and to supply little or no information concerning the data transfer process. Considering that the data transfer process was altered with an artificial delay, the ratings concerning the data transfer process are of particular interest and will be returned to in detail in section 7.4.3.2.

---

[9] A normal distribution used to apply a random delay around the seed.

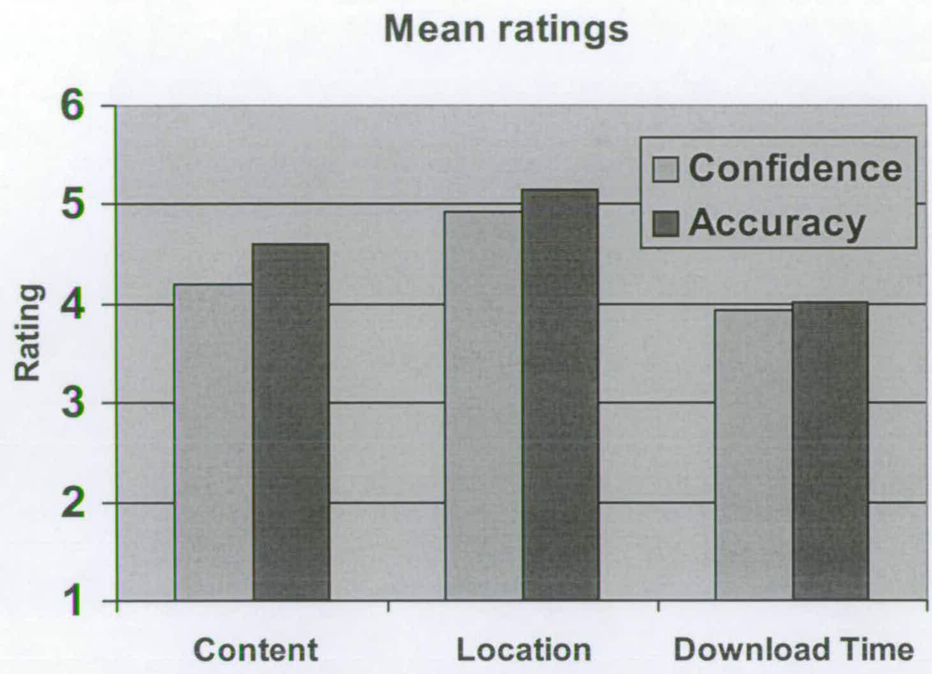


Figure 7.7: Mean rating for all participants

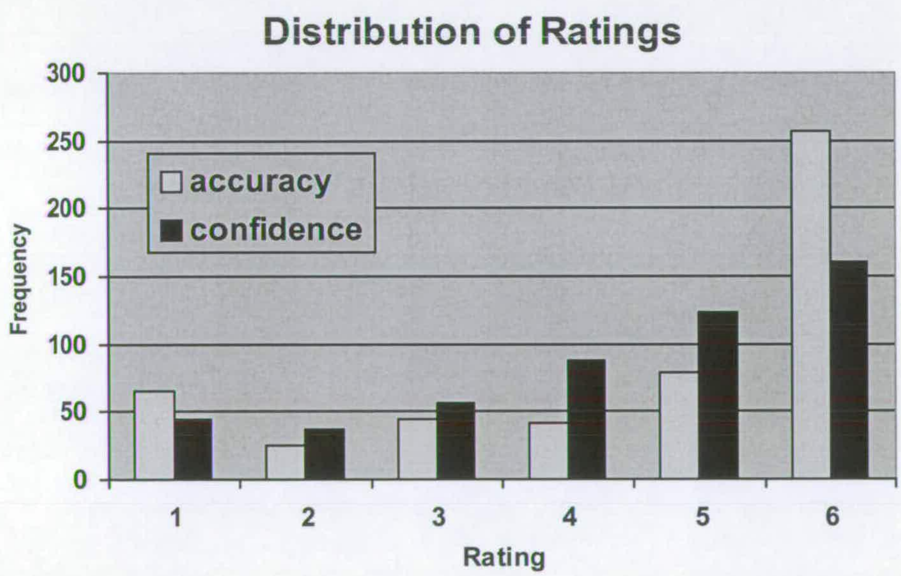


Figure 7.8: Top heavy distribution of ratings

Before any statistical analysis can take place, the distribution of the data needs to be investigated. Figure 7.7 shows that the mean ratings are all well over the normal. A distribution of the ratings (Figure 7.8) presents a distribution that is top heavy with the mode for both the confidence and accuracy rating being six. In particular, the accuracy ratings are very top-heavy implying that participants were answering a polarised question concerning how accurate their predictions were to which the answer was often ascribed the maximum rating of six. This may have been due to the accuracy predictions being a form of *self-assessment* for the participants. After the data transfer process when the resource was rendered on the screen, participants convinced themselves that their predictions had been correct, highlighting the positives and possibly neglecting some of the negatives. This behaviour produced the very high number of maximum ratings as participants were convinced that they had made correct predictions about the referenced resource.

7.4.3.1 Accuracy of participants' content predictions

The first analysis presented here intends to uncover the relationship between how confident participants were about their predictions with how accurate their predictions turned out to be. Comparing the confidence rating of participants' content predictions with how accurate the independent observers rated their predictions to be should reveal whether participants were falsely optimistic or pessimistic about their ability to predict from the URL, or whether the URL provided enough information to enable the participants to accurately make their predictions. Figure 7.9 shows the basic pattern that the results should conform to, with a best fit falling below the normal suggesting a pessimistic participant and a best fit falling above an optimistic participant. The problem with the data collected by this experiment is that it is discrete and producing a best fit in a scatter plot is not possible. The results are therefore presented as a 3D histogram that shows the frequency of the value pair (in this case the confidence rating of a participant and the average accuracy rating of the observers). Figure 7.10 shows this 3D histogram comparing the participants' confidence with their content predictions with the independent observers' rating of their accuracy.

Despite the fact that a large proportion of the value pairs are six, six, the histogram shown in Figure 7.10 still shows signs of tending towards the optimistic side of the normal. There are far fewer values that show a participant having a low confidence but eventually proving to be accurate, than participants being confident but actually being inaccurate. Although far from conclusive, this does suggest that participants were slightly over confident with their predictions.

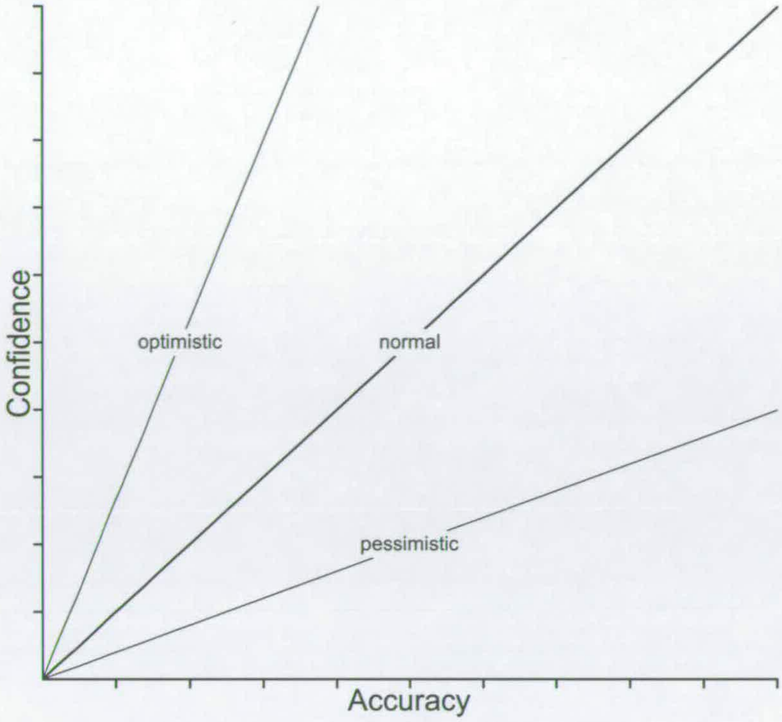


Figure 7.9: Regions of the confidence accuracy graph.

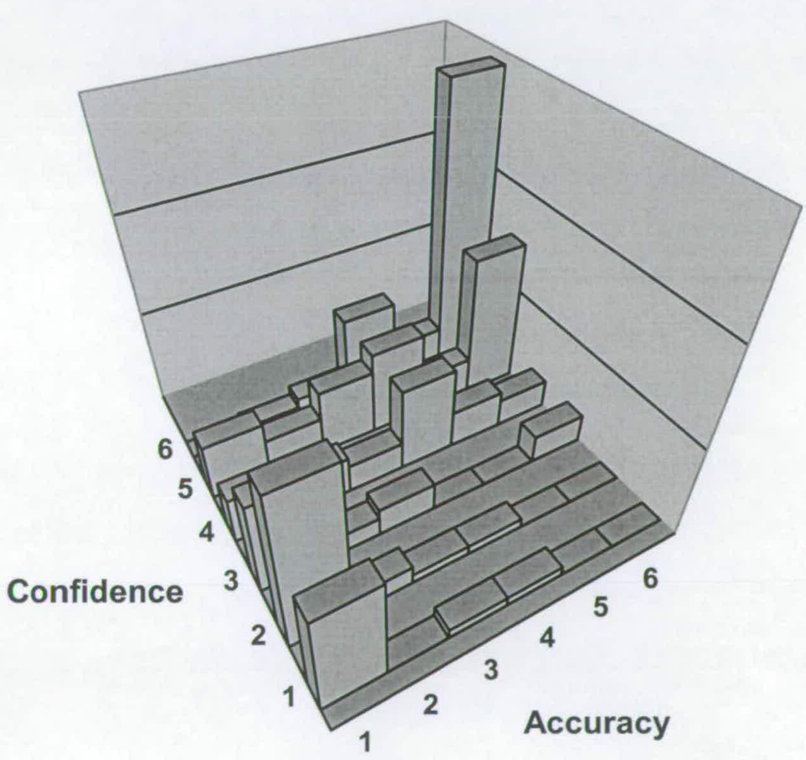


Figure 7.10: Participants' confidence rating with respect to Web page content, compared with the independent observers' accuracy ratings.

7.4.3.2 The data download predictions

The URL affordance experiment concluded that the URL presented poor information with respect to the data transfer process. Participants implied details concerning the data transfer process from knowledge obtained about the resource's location and content. The same was found in this experiment with participants implying the data transfer properties from other factors (Extract 7.9 and 7.10).

```
                          http://uk.samba.org/

I: How long do you think it is going to take to download?
S:                                                        It depends how busy it is
right now.
I:         How could you tell?
S:                               Well you can't really tell, but errr, after about
twelve o'clock our time America wakes up, the United States wakes up, and then it
gets a little busy, a little slow, so the mornings are the best time.
I:                                                               And how about now?
S: Now it should be OK, it should be yeah, I am pretty confident it will be OK.
```

Extract 7.9: Participant J (Novice 10 second delay) confidently predicting properties of the data transfer process.

```
                   http://www.ozsports.com.au/cricket/

I:Now how long do you think this one is going to take?
S:                                                  probably quite a lot longer as
this one will be a bit more popular and, does it make a difference as it has a longer
address, I would think so.
I:                         Why do you think that?
S:                                            because it has more to go through,
if it is one word it has to look up one word but this one will take a bit longer.
```

Extract 7.10: Participant B (novice, 8 second delay) predicting the properties of the data transfer process.

To confirm this result, there should exist a correlation between location and content predictions and the data transfer predictions. When participants produced highly confident location predictions they would also produce confident download time predictions, and *vice versa*. Figure 7.11 shows the 3D histogram comparing the download predictions and the location predictions for all the participants and URLs. It has already been shown that the distribution of the results was top heavy and this is clear from the Figure 7.8 graph. There is, however, no doubt that there exists some proportionality between the location prediction and the download time prediction as high confident ratings with respect to location often produced high download time prediction. In other words, there are very few predictions where a participant was very confident within one domain and not at all confident in the other.

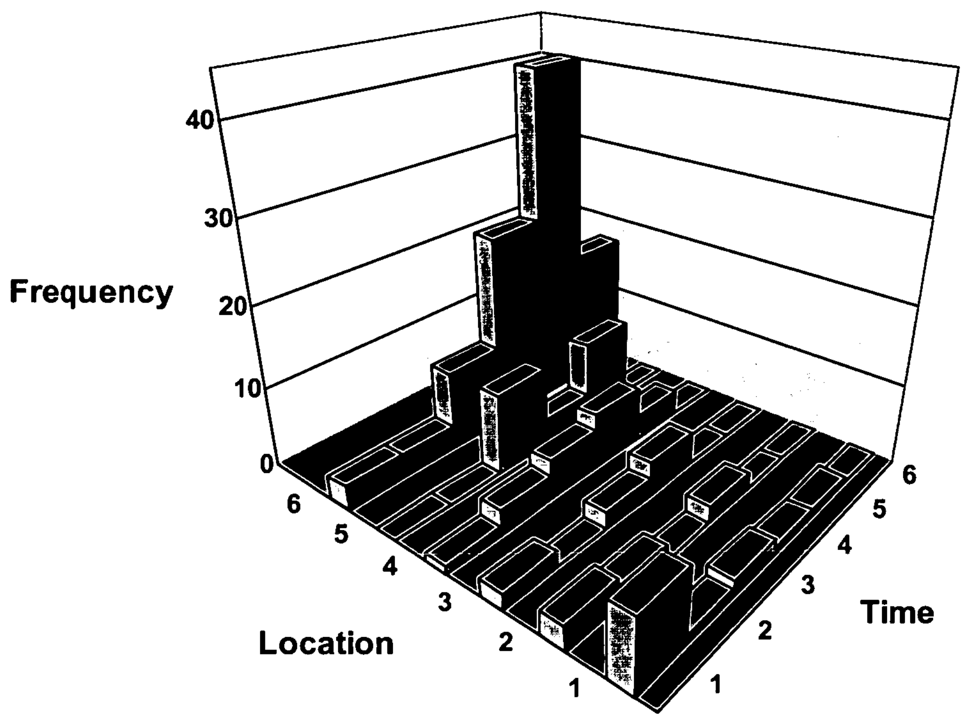


Figure 7.11: 3D histogram for the confidence of predictions with respect to location and the download time.

In contrast, the 3D histogram comparing the confidence of the content and download time predictions shows a far less exaggerated correlation (Figure 7.12). Although confident content predictions often produced confident download time predictions, the difference is far more spread showing a significant number of predictions with low confidence with respect to content, but high confidence with respect to download time.

The graphs presented in Figures 7.11 and 7.12 nicely illustrate the differences in the relationship between the confidence in content and location predictions and the confidence in download time predictions. It appears that the more confident a participant was about the location of the data, the more confident they were in their time predictions. There are, however, statistical methods that can also uncover correlations between the variables although they must be used with some care due to the small sample size and the non-normal distribution of the data collected.

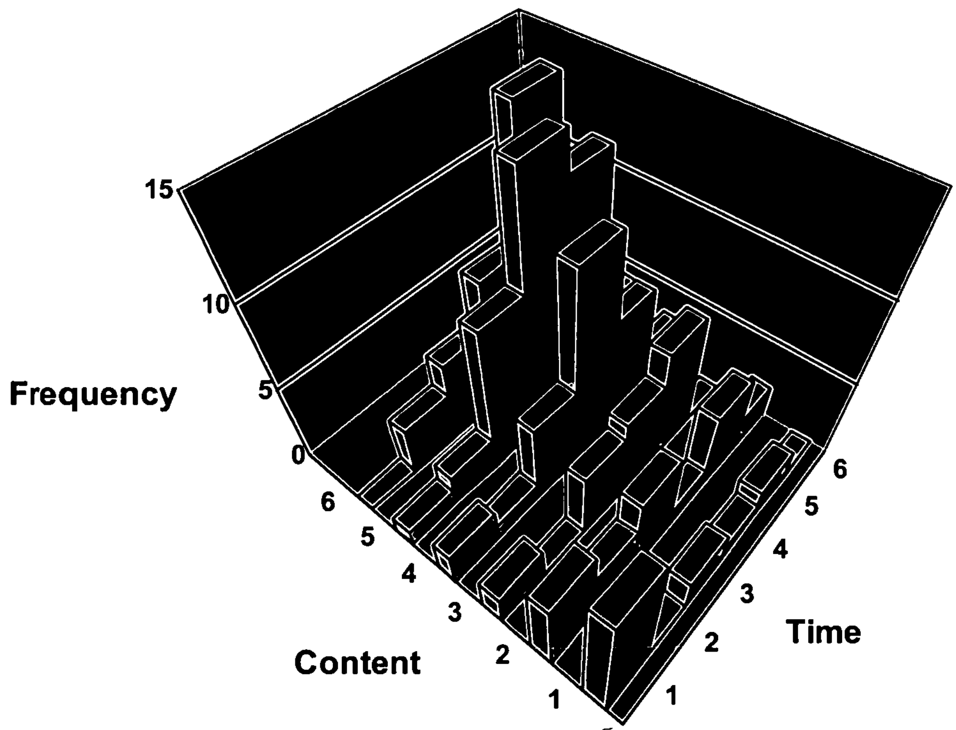


Figure 7.12: 3D histogram for the confidence of predictions with respect to content and the download time.

A correlation coefficient measures the relationship between two (or more) variables. The correlation coefficient can range between –1.0 and 1.0 with –1.0 indicating a negative correlation, 1.0 indicating a positive correlation and 0.0 indicating a lack of correlation. By calculating the linear correlation coefficients between the six data sets collected by the URL delay experiment, any proportionality between the data can be uncovered. For example, if the URL presented the participants with a good and accurate affordance concerning location (as has already been suggested) not only would the mean values be high, but there would also exist a positive correlation between the confidence of a location prediction and the accuracy of the prediction. In other words, when a participant was confident about a prediction they would also be correct; the URL would not be misleading.

Table 7.10 shows all the correlation coefficients between the six sets of data collected from the participants by the experiment. It is clear from the table that there is a strong correlation (0.79) between the confidence and the accuracy of location predictions. Similarly, there is also a strong correlation between the confidence and the accuracy of content predictions (0.68). In contrast, the correlation between the confidence and accuracy of the data transfer process predictions is less significant (0.44). These three results show that the confidence of a participant's content and location predictions shows some proportionality to how accurate they subsequently felt they were. When making predicting about the data transfer process, however, the participants show much less certainty that a confident prediction would result in an accurate prediction and *vice versa*. Again, it confirms the result that the URL provides far more information concerning the location and content of the referenced Web page than the data transfer process and, what is more, participants believe they elicit this information accurately.

		Confidence			Accuracy		
		Content	Location	Time	Content	Location	Time
Confidence	Content	1	0.478891	0.426639	0.680158	0.506466	0.31948
	Location		1	0.610395	0.509352	0.790011	0.464474
	Time			1	0.324618	0.432114	0.436639
Accuracy	Content				1	0.530996	0.300395
	Location					1	0.468027
	Time						1

Table 7.10: Correlation Coefficients

7.4.3.3 The insensitivity to the artificial delay

One of the aims of this experiment was to investigate the effect that the artificial delay had on the participants' predictions. It has already been shown from the data in Figure 7.7 that participants' confidence and subsequent accuracy in their ratings of the data transfer process were on average lower than the predictions concerning the referenced resource's content and location. What is apparent, however, is that the ratings still appear to be very high (confidence mean of 3.9 accuracy mean of 4.0). The accuracy of the temporal prediction is particularly high especially when one considers that an artificial delay was introduced.

No participant performed any actions or made any utterances to show that they realised that an artificial delay had been introduced into the data transfer process. Concluding from this that participants believed that the system was performing normally, suggests that participants were not eliciting anything directly useful from the URL about the data transfer process, otherwise they would have identified the existence of the artificial delay. Secondly, the size of the artificial delay appeared to have no effect on participants' predictions.

Drawing on the investigation into the URL presented at the beginning of this chapter and the explanation of the technology behind the Internet presented in Chapter 3, the fact that participants were unable to extract any information concerning the data transfer process should not be surprising. What, however, is of interest is that participants believed they were not only making confident predictions about the data transfer process (mean 3.9) but, more importantly, they were being fairly accurate with these predictions (mean 4.0). It would appear that participants had an inflated opinion of their ability to understand the data transfer process from the URL, or that their explanations factored in the unpredictable nature of the process, thereby adapting their predictions to explain the extra delay that was added.

7.4.4 Summary of URL delay experiment

The URL delay experiment clearly shows that the participants were unable to elicit any accurate information from the URL relating to the data transfer process. If participants were able to elicit useful information, they would have been able to uncover the artificial delay that was introduced by the proxy server throughout the experiment. Not even one participant suggested the presence of any extrinsic delay even though the data transfer process of a Web page was increased, on occasion, by as much as two minutes. The URL does not afford any consistently accurate information about the data transfer process before, during, or after hyperlink activation.

The in-depth investigation into the URL undertaken at the beginning of this chapter showed that the properties of the communication channel are not encoded within the URL: the fact that participants cannot elicit them is therefore not surprising. What, however, is surprising is how highly participants rated the accuracy of their predictions concerning the properties of the data transfer process. Participants from all levels of expertise consistently gave themselves high ratings of how accurate they had been in predicting the data transfer process. What became apparent was that participants' confidence with their temporal predictions positively correlated with how confident they were with their spatial predictions. The URL does provide some consistent spatial information in the host - the TLD and possibly other domains – and what is more, the host is always present in every well-formed URL. Users can always make an educated guess about the location of the referenced HTML document from the host, and they then proceed to depend on this to make their predictions concerning the data transfer process. The fact that location is consistently provided by the URL is perhaps one reason why the visual representations presented in Chapter 5 are so geographically dependent.

The experimental participants also gave themselves high accuracy ratings for their content predictions when compared to the independent observers. One possible reason for this over confidence is the belief that the URL is directly related to the referenced Web page, when in fact it is only a pointer to a HTML document. Understanding that the Web page is made up of a number of components and the URL is only a pointer to one of them is crucial if users are to appreciate the complexities of the system and hence to look for explanations for the behaviour they experience.

7.5 Summary

The findings of the investigation presented in Chapter 5 produced a set of users' visual representations of the WWW. Subsequently, in Chapter 6, these visual representations were shown to provide the basis for a set of heuristics, which users employed in conjunction with other more situated knowledge and assumptions, to explain the system behaviour they experienced in a certain circumstance. What the observational investigation uncovered was that the main source for the explanations that users provided was the URL of the referenced Web page. The URL was not only used to extract information about the present Web page, but its continual use also helped shape the visual representations and heuristics that users possessed. Having uncovered the use of the URL as one of the salient issues of the interaction as a whole, the experiments presented in this chapter aimed to show exactly what information the URL affords the user and how successful users are at extracting it.

The URL is definitely employed by users to help assess the referenced Web page's content and during the data transfer process as a means of explanation for a perceived system disturbance. What the experiments presented in this chapter highlight is that the URL is not a reliable source for this type of information. Although some participants were shown extracting rich and accurate information about a referenced Web page from a URL, with a different URL the same participant struggles to extract anything useful. In general, the URL consistently provided some spatial information concerning the location of the HTML document data. From this limited spatial knowledge and some situated assumptions – like, for example, the amount of people using the WWW at a particular time of day - users try to make predictions concerning the data transfer process. In addition, users appear to be overly confident about their ability to make these predictions from the URL.

The hypertext abstraction conceals the dynamic workings of the system and the contents of referenced Web pages. The fact that users resort to a strategy that employs the URL – an inconsistent source of system information - when trying to discover the contents of the referenced resource and explain the disturbances of the system, is a very bad reflection on the browser interface as a whole. The information that users try to elicit from the URL – the referenced Web page's content, location and the temporal properties of the data transfer - must be presented to the user in a more consistent and predictable form. This is the aim of the browser interface extension whose design and example implementation is presented in Chapter 8.

It was claimed in Chapter 5 that through the accounts of how chess players understood the game of chess, a new chess interface could be constructed that would highlight the relationships between pieces and show the important structures of the position. Having undertaken a number of investigations and experiments to determine the user's understanding of the WWW, this dissertation is now in a position to do the same for the WWW: redesign the interface and show the WWW user the important relationships and structures of the WWW that have been uncovered. Chapter 8 introduces the *hyperlink lens* – an extension to the browser – that redefines the hyperlink to provide document metadata and time affordances, and hence provide the user with a better interface: an interface that presents information from within the WWW abstraction

Chapter 8

The Hyperlink Lens

The work presented in this dissertation has identified that users struggle to break down the major interface abstraction of the Web page before hyperlink activation and during the data transfer. What this chapter will describe is a proof of concept implementation of a WWW interface, or rather an addition to the existing browser, that allows the system mechanisms that are presently hidden behind the abstract model of the WWW to be presented to the user. First, the Web page is presented as a collection of resources rather than a single entity and second, the hyperlink is redefined to allow the spatial, temporal and informational properties of the referenced resources and the defined data transfer process to be presented to the user for assessment. For reasons explained later in the chapter, this implementation has been called the *hyperlink lens*.

8.1 The hyperlink lens

The *hyperlink lens* is an application that simulates how the WWW browser could be enhanced to reveal the underlying mechanisms of the system. In particular, the *hyperlink lens* breaks down the abstraction of the Web page to provide the user with information that they are unable to elicit with the present WWW browsers. The design principle adopted presents system information and preview information to the user without making a commitment as to how it may actually be used. The *hyperlink lens* reveals the underlying mechanism of the system in order for the user to understand the system behaviour when, or if, circumstance decrees.

The abstract model of the WWW presented in Chapter 2 was a network of information objects connected by hyperlinks. The course of work undertaken for this dissertation has shown the inadequacies of this model in explaining the actual behaviour of the system and, the way that WWW users are unable to circumvent it in order to explain the real behaviours they witness. To reveal the underlying mechanisms of the system the *hyperlink lens* was designed to present a redefined hyperlink; a hyperlink whose attributes are not only the destination URL but, are also a much richer set explaining the time, information and spatial properties that have been shown to interest the WWW user. The *hyperlink lens* is a window through the abstract model of the WWW into the real workings of the system below.

8.1.1 Magic Lenses™

Technically, the implementation of the *hyperlink lens* is based on *Magic Lens* filter technology that was identified as a new user interface tool in 1994 (Stone, Fishkin, Bier, Buxton and Baudel, 1994, and Stone, Fishkin and Bier 1994). A magic lens filter is a movable region of screen real estate that is interactively positioned over the interface of an application to affect the appearance of the structures viewed through it. The technology is based on the metaphor of physical lenses like, for example, a magnifying glass that can be positioned over a newspaper magnifying only a small portion of its text. The filters that magic lenses apply can vary from the general "*multiple-application*" filters, which apply across applications and yet still have semantics specific to every one, to the uniquely specified "*application-specific*" filters. Although magic lens technology was designed as a visualisation tool that would alter the appearance of objects in graphical environments, Stone *et al.* suggested that the magic lens technique could be employed to display all sorts of properties of the region under the lens. In particular, they proposed that the magic lens technology could be used in hypertext systems.

> *"In a hypertext system, lenses could be used to highlight links and enable their operation. This provides a solution to the visual clutter caused by making the links always visible. Lenses could also be used to activate the links. In this*

case, the lens could display either the attributes or contents of the destination node." (Stone, Fishkin, Bier, Buxton and Baudel 1994, p 309)

Although the *hyperlink lens* is not implemented using exactly the same techniques as the magic lens filters presented by Stone *et al.,* it does have a number of common properties. Principally, the *hyperlink lens* reveals the properties of the hyperlink *under* the lens. The fact that (as shall be shown) the actual lens is not placed directly over the activating object of the hyperlink does perhaps break the metaphor of the magic lens filter. Nevertheless, the extract above does suggest a system with similar properties to the *hyperlink lens* and, therefore, it is at least a derivative of magic lens technology if not a direct application of it.

8.1.2 Design and Evaluation

The implementation presented in this chapter is a working system that simulates how a Web browser could be enhanced to provide the user with document metadata and time affordance. The objective of this implementation was to show first, how more information could be presented to the user and second, how this information helps the user induce more appropriate explanations of the system behaviour that they experience. This chapter mainly addresses the first of these objectives by focusing on the design and implementation details of the *hyperlink lens* but section 8.3 includes a discussion on how the second objective might be achieved.

The development of this implementation involved a number of design and evaluation cycles and progressed from paper-based sketches to a java applet based system. The full system presented in this chapter is still only a simulation of how the *hyperlink lens* could look and behave. This chapter not only presents the details of the implementation but also shows how the simulation was developed including some of the design processes that were used to finalise what sort of information should be presented.

The design of the *hyperlink lens* evolved from paper-based sketches of possible features. The sketches were two colour printouts of an interface that presented various types of information about a particular Web page at a particular moment in time (*i.e.*, during data transfer). These paper-based sketches were presented to three focus groups each comprising of three people. Table 8.1 presents the expertise of each subject.

First, the focus group was asked to discuss and explain to the experimenter what they believed each feature was presenting to them about the hypertext. Each focus group was then given a list of the features present on the sketches and encouraged to discuss the merits of each in turn. The discussion was directed to reveal how useful each subject felt each feature was After the general discussion, each subject was asked to rate the perceived usefulness of each feature using a one to eight scale where eight was very useful and, one was useless. After this evaluation of the interface, each member of the focus group was asked to suggest other information that could be added to improve the usefulness of the interface.

Focus Group	**Subject**	**Gender**	**Expertise**
A	1	M	Expert
A	2	M	Expert
A	3	M	Expert
B	4	F	Novice
B	5	M	Novice
B	6	M	Expert
C	7	M	Intermediate
C	8	M	Intermediate
C	9	M	Expert

Table 8.1: Focus Groups

The reaction to the prototype lenses was mixed. The focus groups revealed that the prototypes presented far too much information; "*the interface is too crowed*", "*very colourful and confusing*" and "*bewildering*". Some of the information present, however, was deemed very useful and the subjects stated that it would help them make an informed decision about whether to activate a hyperlink. The consensus from the focus groups was that a form of *hyperlink lens* would positively influence a user's behaviour whilst browsing the WWW.

In particular, the focus groups highlighted the abstract of the Web page and the thumbnail picture as being the most potentially useful pieces of information. These two features also averaged the highest ratings of usefulness and the focus groups predicted that the main benefit of this information would be to stop unnecessary hyperlink activation. Although some of the features presented to the focus groups were not rated very highly on usefulness, the focus groups did suggest that information concerning the data transfer process would be "*interesting*" and explain more clearly what was happening and, may even help them to detect a system disturbance. The information collect from these focus groups was used to isolate the important features which should be present on the *hyperlink lens*. The features present on the *hyperlink lens* described in the rest of this chapter are a direct result of this formative evaluation. A further account of this design process and this usability study is presented in Stanyer and Procter (1998; 1999).

8.1.3 The Application

The *hyperlink lens* is a separate window used in conjunction with a conventional Web browser. When a user places the mouse cursor over the activating object of a hyperlink, the *hyperlink lens* displays information about the referenced Web page (Figure 8.1). Once the user activates a hyperlink, the *hyperlink lens* provides a temporal affordance during the data transfer process by way of animation.

The basic data transfer concept presented in Stanyer and Procter (1998; 1999) was termed the *channel*. To provide the user with appropriate information about the data transfer process a communication resource was specified that connects two computers together and allows information to flow between them in either direction. By restricting the notion of a channel to the overall connection between the browser and the server, the channel becomes an abstraction synonymous with the HTTP communications. All information presented to the user by the *hyperlink lens* concerning the data transfer process is relevant to this channel concept.

The information that the *hyperlink lens* provides is split into three sections and each is presented on a separate pane in the window.

- **Structure** - the structural make up of the referenced Web page, the number and size of each information object.

- **Overview** - preview information about the referenced Web page's content that allows the WWW user to better assess the value of the referenced Web page before committing to hyperlink activation
- **Rating** - facility to provide information compiled by other users or user groups who have previously assessed the Web page and submitted a rating

The discussion of this implementation will centre on the structural information, as this is where the constituent parts of the Web page are shown. It is through the structural section of the hyperlink lens that the model of the WWW is changed and where the hyperlink is redefined.

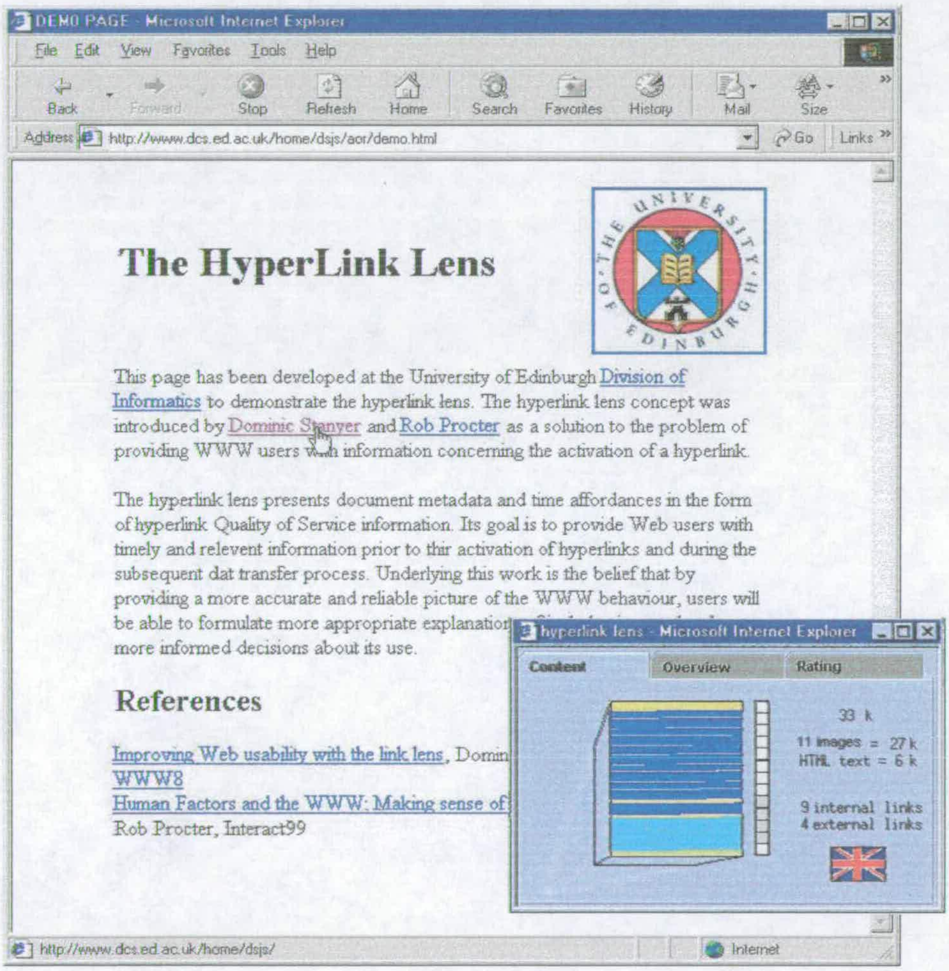


Figure 8.1: The *hyperlink lens*

The structural view of the *hyperlink lens* presents the WWW user with a detailed break down of the Web page referenced by the hyperlink. The two main features of this information are:

- the components of the Web page
- the hyperlinks that originate from the Web page

This information includes the size of the HTML document that defines the Web page and the type and size of the embedded objects it references. The result is a dissection of the Web page into its constituent parts. The first advantage of presenting this information to the user is that it provides a realistic interpretation of the Web page. The WWW user is in a position to understand the components of the Web page and the implications this may have on the subsequent data transfer process. The Web page may be graphically intensive and require the download of a number of large files. Similarly, knowledge concerning the number and type of hyperlinks that emanate from a referenced Web page may provide the user with the facility to better assess the potential value of activating the hyperlink.

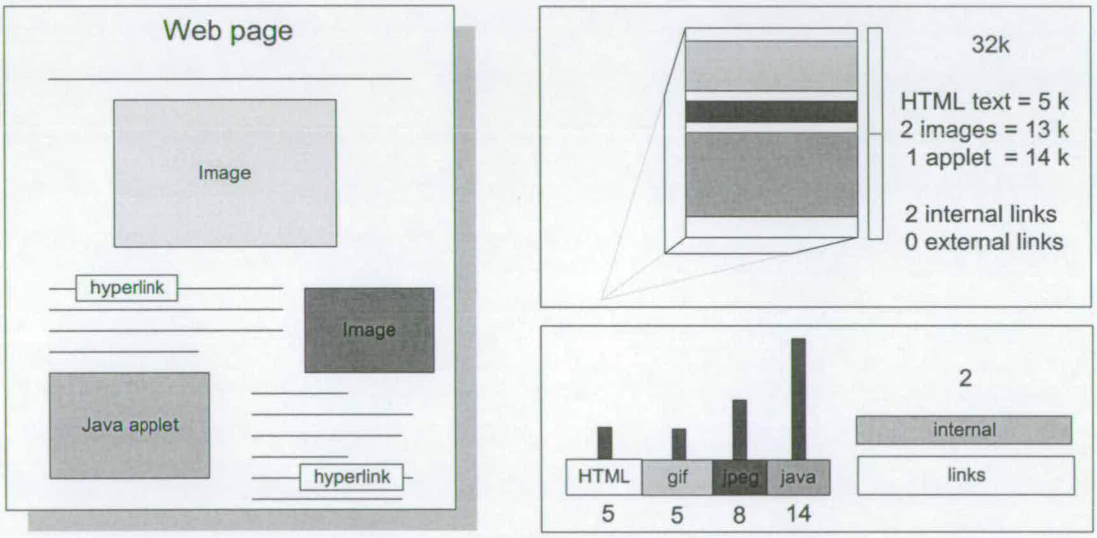


Figure 8.2: Two visualisations of the Web page constituents

Figure 8.2 shows two visualisations of the Web page on the left. In Stanyer and Procter (1999), the bottom visualisation was used to describe the constituents of the Web page. The present implementation, however, uses the top visualisation, as it is more scalable when Web pages contain a large number of embedded objects. The visualisation is based on a prism that presents the relative size of each embedded object on its face. The solid distance (depth) of the prism shows a total size of the Web page using an exponential scale[1]. The example in Figure 8.2 shows a Web page of size 32k and therefore has a depth along the near corner to the vanishing point of 12 pixels. As well as this graphical visualisation, the components are also displayed in text.

The *hyperlink lens* also provides some information concerning the hyperlinks that originate from the referenced Web page. Some Web pages contain a large number of hyperlinks and some contain none, ending any thread of interest that a user is following. Knowledge of the number and type of hyperlinks that a Web page contains is, therefore, of potential interest to the user when assessing whether to download a Web page. The *hyperlink lens* provides information about the number of hyperlinks that a Web page contains and whether each hyperlink references another Web page that is within the same domain of the referencing Web page (internal) or from another domain (external).

8.1.3.2 Temporal Affordance

Once the hyperlink is activated, the structural view is automatically shown so that some temporal affordance properties can be displayed to the user. The concept of temporal affordance was introduced in Chapter 4 and Conn's eight requirements were specified. The *hyperlink lens* addresses a number of these requirements by colouring in the visualisation of the Web page as data arrives

Figure 8.3 shows four stages of this process with the visualisation in various stages of download as data arrives at the client (or more precisely a proxy). Moving clockwise from the top left frame:

1) Frame one shows the visualisation before any data has arrived.
2) Frame two shows the progress after the HTML document, that defines the Web page, has been downloaded but none of the subsequent embedded objects have arrived.
3) Frame three shows the embedded objects in a state of transition.
4) Frame four shows the successful completion of the task.

---

[1] 2 pixel is 1 Kbytes, 4 is 2 Kbytes, 6 is 4 Kbytes et cetera, up to 50 pixels which represents greater than or equal to 16 Gbytes.

4) Frame four shows the successful completion of the task.

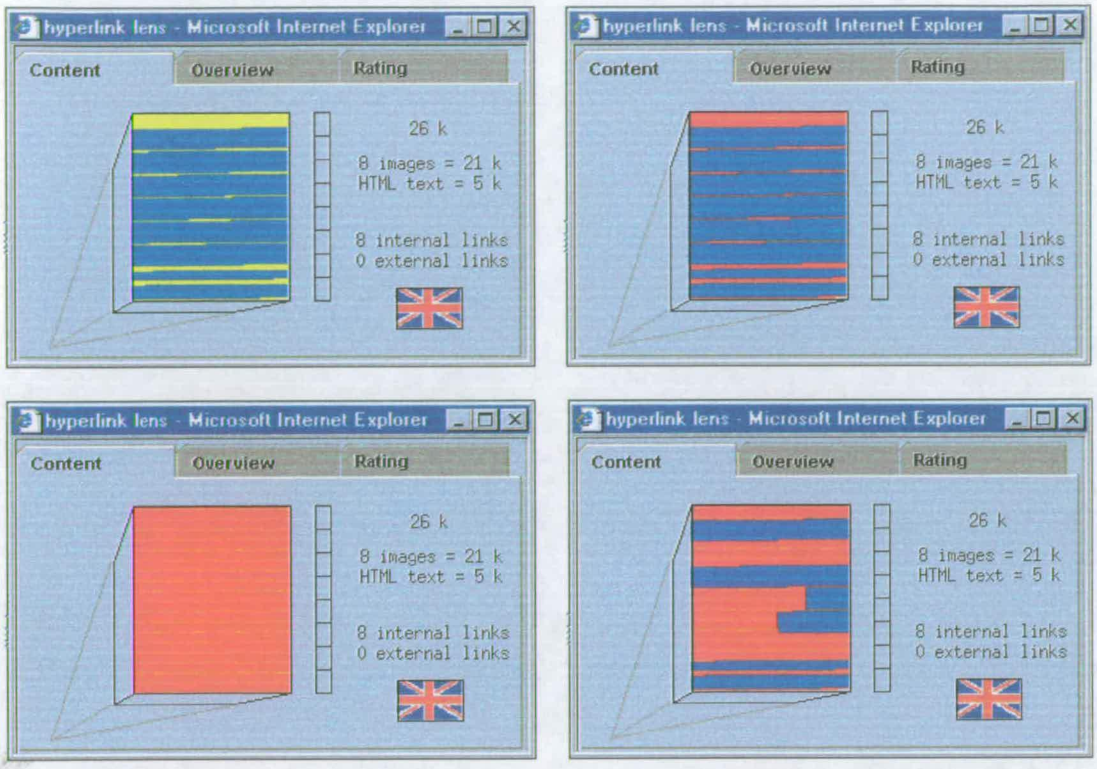


Figure 8.3: Temporal progress affordance (starting top left and progressing clockwise).

The areas of temporal affordance that this simple animation and task description tackle are:

- **Acceptance** – by showing the referenced Web page, the *hyperlink lens* provides the user with knowledge that there exists a number of objects for download after activation of the hyperlink. The *hyperlink lens* shows the user that the referencing has been correctly specified and that the destination node has not succumbed to linkrot[2].
- **Scope** – the visualisation of the Web page components, displays to the user the scale of the task to be completed. The scope of the task is, however, not completely satisfied as the *hyperlink lens* does not display a prediction for the delay in seconds.
- **Initiation** – although this requirement is already fulfilled by the Web browser the *hyperlink lens* provides a further source of initiation information.

[2] The architecture section of this chapter shows that linkrot may still occur but constant updating of the structural information will minimise the risk.

parts, progress is far more easily displayed. The visualisation also clearly copes with the non-deterministic nature of the download process.

- **Heartbeat** – the animation of the data transfer is a form of heartbeat although it does not state anything about the continued state of the task in terms of the communication.
- **Exception** – the *hyperlink lens* does not fulfil this requirement, as it does not explicitly show communication failure. It will however, present the user with an indication as to where an error may have occurred. If an embedded object has not been received, it is clear from the visualisation that this object is the cause of the failure.
- **Completion** – this is clearly fulfilled by the animation.
- **Remainder** - the animation shows the remainder of the data transfer tasks to be accomplished although, again, there are no explicit times provided.

In summary, a simple visualisation of the Web page showing its component parts, coupled with an animation during the data transfer process provides the user with information that has the potential to fulfil a large majority of the temporal affordance requirements.

8.1.3.3 Overview

The overview of the *hyperlink lens* provides preview information concerning the contents of the referenced Web page. The problem that a preview attempts to address was identified in hypertext literature as long ago as 1987 (Conklin, 1987) where it was observed that a hypertext user often left the referencing node only to immediately return after determining the uselessness of the referenced node. There have been a number of proposed solutions to this problem some of which have started to appear in the current WWW browsers. Bieber, Vitali, Ashman, Balasubramanian and Onias-Kukkonen (1997) categorised the hypertext functions that address this type of problem as "*global and local overviews*". One example of the WWW currently supporting a basic form of the link preview is the textual hint supplied by Microsoft Internet Explorer by way of the alt attribute in an anchor element (Figure 8.4).

When a user moves the cursor over the activating object of a hyperlink, the WWW browser presents the user with a short piece of text (the contents of the alt attribute). This textual hint supplies a user with some context in order to decide if a hyperlink is worth activating but is dependant on the author providing it and is limited to text. This type of simple textual overview was derived directly from hypertext systems where it has been called a "*gloss*" (Zellweger, Chang and Mackinlay, 1998) or a "*scent*" (Furnas, 1997).

```
<a href="http://www.dcs.ed.ac.uk/" alt="Department of Computer
        Science at the University of Edinburgh">home</a>
```

Figure 8.4: A textual hint

One piece of work that directly addressed this problem of hyperlink overviews for the WWW provided users with a "*visual preview*" of the referenced Web page of a hyperlink (Kopetzky and Mulhlhauser, 1999). Using an implementation based on a proxy server that introduced dynamic HTML layers to any downloaded HTML document, a *thumbnail* image of every referenced Web page could be added to the current Web page. The result was a *proof of concept* system that would present a preview image of the referenced Web page when a hyperlink's activating object was hovered over by the cursor.

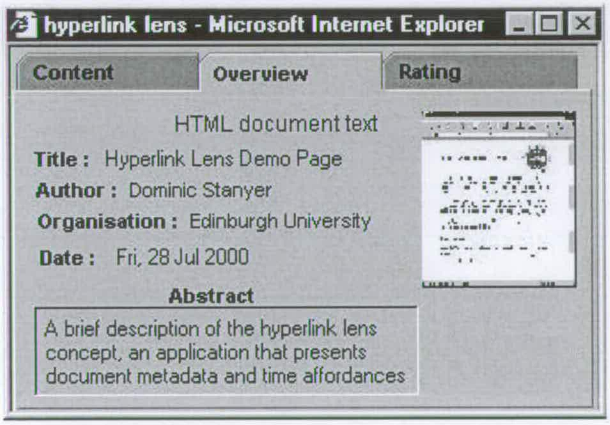


Figure 8.5: The overview section

The overview of the *hyperlink lens* combines these textual and visual forms of preview to provide a rich set of information concerning the referenced Web page. Textual information includes the possibility of an abstract of the Web page, the title, author and organisation and the date when the Web page was published (Figure 8.5). The visual preview is simply a thumbnail similar to the Kopetzky and Mulhlhauser application (it is in fact a pixel resizing jpeg thumbnail rather than the bi-linear resampling suggested by Kopetzky *et al.*). In summary, the overview provides a textual and visual preview of the referenced Web page.

8.1.3.4 Rating

The rating section of the *hyperlink lens* is part of a broader programme of work that aims to add social affordances to the Web through the use of e.g., collaborative filtering to produce user ratings of Web pages. It is hoped that this rating information will add additional Web page quality and relevance information that is personal to the user (Procter and McKinlay, 1997). The rating section is not discussed further in this dissertation, as it is itself a complex area of ongoing research.

8.1.4 Summary

The hyperlink lens presents document metadata and time affordances in the form of hyperlink Quality of Service information. Its goal is to provide Web users with timely and relevant information prior to their activation of hyperlinks and during the subsequent data transfer process. Underlying this work is the belief that by providing a more accurate and reliable picture of the WWW behaviour and content, users will be able to formulate more appropriate explanations concerning its behaviour and make more informed decisions about its use. The rest of this chapter presents the actualities of the implementation starting with the basic architecture and proceeding with some of the real design decisions.

8.2 Architecture

The *hyperlink lens* presents information about a resource when the mouse pointer is moved over a hyperlink's activating object. The *hyperlink lens* must have the information about the referenced resource **before** the activation and subsequent download of the actual resource. In fact, every hyperlink contained on a Web page must have an associated information object that describes the Web page indirectly referenced by the hyperlink through the URL of the referenced HTML document. This description is what is used to provide the information visualisations presented previously in this chapter. This descriptive information object is an attribute of the hyperlink itself; every hyperlink should posses an information object that describes the Web page it references. The architecture of the hyperlink lens is based on this redefining of the hyperlink.

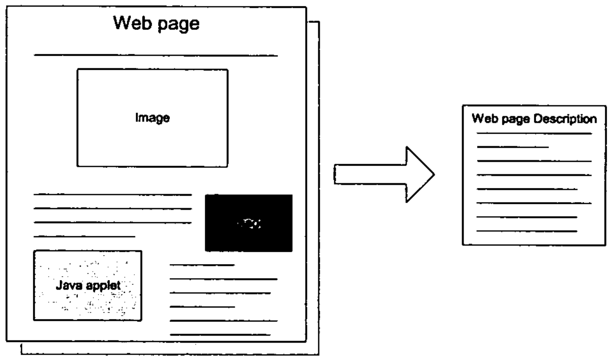


Figure 8.6: The Web page description

Every Web page (or in general every WWW resource) can be described by an information object (Figure 8.6). Traditionally, this information object has been referred to as metadata and is associated with the object it describes. The *hyperlink lens* differs from the traditional metadata approach in two ways:

- what is described by the metadata and,
- where the metadata is stored.

8.2.1 The Dublin Core metadata initiative

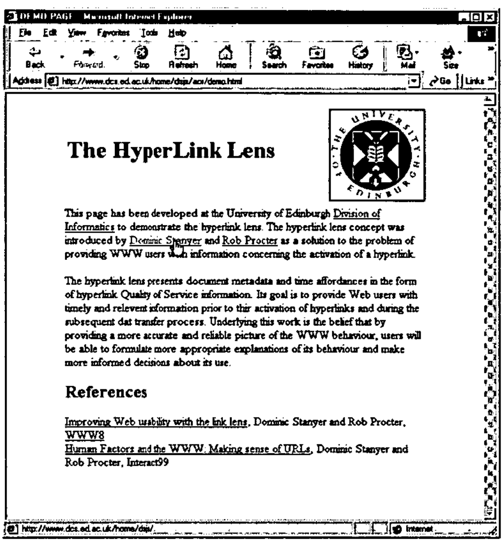
The Dublin Core (Dublin Core Metadata Initiative, 2000) is a simple content description model designed specifically for electronic resources. It provides a metadata element set intended to allow authors to add descriptions of their Web resources to facilitate discovery. Proponents of Dublin Core claim that it has characteristics that positively distinguish it from other metadata approaches.

- **Simplicity** – elements have commonly understood semantics
- **Semantic interoperability** – common set of descriptors
- **Extensibility** – a simple element based approach

Figure 8.7 shows a Dublin Core metadata description for an example Web page. The information that this description contains is contextual; the name of the Web page; a list of keywords; the author; the publisher and, an abstract or description. There is also some information about the actual object that, in this case, is the HTML document rather than the

Web page: type and format. Overall the Dublin Core metadata describes the contents of the Web page, and some limited technical information about the HTML document.

This is the first difference between this type of metadata description and the descriptive information object used in the *hyperlink lens* system: the contents of the Web page is described along with its structure. The descriptive information object used in the *hyperlink lens* system describes the HTML document and **all** the embedded objects. This allows the *hyperlink lens* to display both the structure of the Web page, and the more traditional overview.




```
<?xml version="1.0" ?>
<Dublin Core>
   <Title>Hyperlink Lens DEMO PAGE</Title>
   <Creator>Dominic Stanyer</Creator>
   <Subject>The hyperlink Lens; Division of
Informatics; Improving Web usability with the
link lens</Subject>
   <Description>A brief description of the
hyperlink lens concept, an application that
presents document metadata and time
affordances.</Description>
   <Publisher>Edinburgh
University</Publisher>
   <Date>2000-07-28</Date>
   <Type>Text<Type>
   <Format>text/html || 4081 bytes</Format>

<Identifier>http://www.dcs.ed.ac.uk/home/dsjs
/aor/demo.html</Identifier>
</Dublin Core>
```

Figure 8.7: Dublin Core metadata description encoded as an XML head element generated by Dcdot (`http://www.ukoln.ac.uk/cgi-bin/dcdot.pl`).

The second difference between the traditional metadata approach and the *hyperlink lens* is where the metadata is stored. In the example presented in Figure 8.7, the description is encoded as XML metadata and is added to the head element of the HTML document that defines the Web page. Alternatively, the same Dublin Core description could be encoded as HTML meta elements. Either way, the metadata is always stored **in** the HTML document that defines the Web page. To retrieve the metadata that describes a Web page a client must download at least part of the HTML document. In the *hyperlink lens* architecture the descriptive information object (the metadata) is associated with the hyperlink that references the Web page (Figure 8.8). The hyperlink is not only a pointer to a Web page but it is also a description of the Web page and its constituents. If a client has the referencing Web page, it also has the description of every Web page referenced.

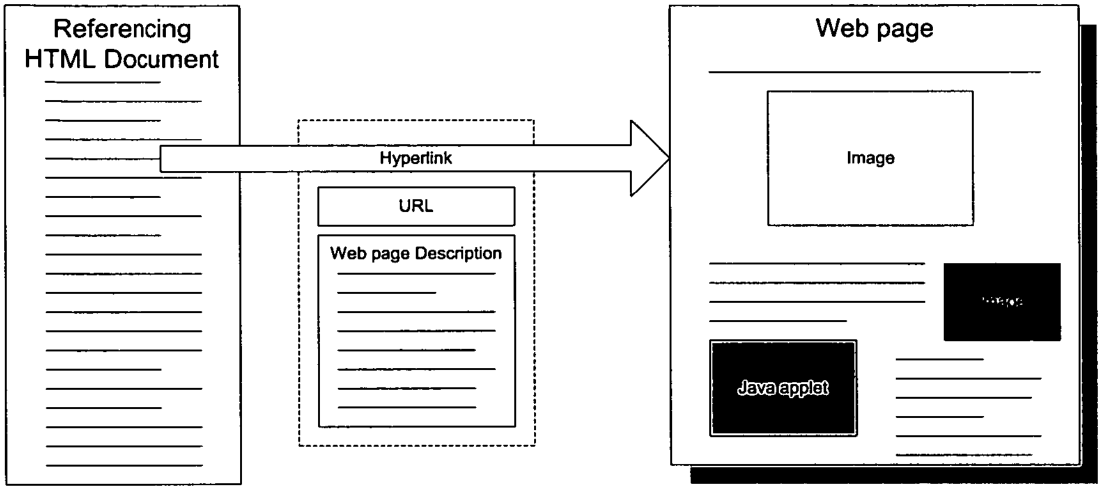


Figure 8.8: The Hyperlink's Attributes

Chapter 3 contained a detailed description of the HTML and in particular the anchor element, which defines a hyperlink. To recap, the anchor element in its most basic form contains a start tag and end tag which together surround the activating object of the hyperlink. The start tag must contain an href attribute whose contents is the URL of the referenced object. It may also contain a number of other attributes like, for example, the alt tag whose use has already been explained. The information object that describes the referenced resource is another attribute of the anchor element and therefore an attribute of the hyperlink itself. The architecture of the *hyperlink lens* insists on this new attribute being added to the anchor element of the HTML (Figure 8.9). The descriptive information object has been entitled the **annotated object reference** and the actual attribute, the **aor**.

```
<A
        href="http://www.dcs.ed.ac.uk/"
        aor="annotated object reference"
        alt="Computer Science Homepage">home page
</A>
```

Figure 8.9: The introduction of the descriptive information object

8.2.2 The Annotated Object Reference

The information contained in the annotated object reference is what is used by the *hyperlink lens* to display the structure and overview information. The annotated object reference is actually encoded using the XML mark-up style shown in Figure 8.7. The reason for employing this approach is twofold. First, the annotated object reference is stored in an HTML document and so should be human readable and human editable. Second, by using a similar mark-up scheme as other metadata methods, like the Dublin Core initiative, the chance of extensibility is enhanced. The annotated object reference is therefore a hierarchy of tagged information. It is split into two categorise to reflect the two sets of information: structural information and contents preview information.

8.2.2.1 Structural Information

The structural part of the aor encodes whether the resource is a Web page and if so, how the Web page is in reality made up of a number of objects. The aor must provide information to the *hyperlink lens* concerning the type and size of these fundamental elements. Furthermore, the aor also provides information concerning the hyperlinks from this Web page. This information can be generated automatically without the need for any author intervention. The implementation of the *hyperlink lens* breaks down the structural information into three sections; the resource's type, the Web page's layout and the hyperlinks it contains.

8.2.2.1.1 Type

The type of the resource simply states whether the resource is a Web page or another basic type. The present implementation uses the Unix *file* command to determine the type of a resource but a MIME type (as discussed in Chapter 3) could be considered more appropriate and in keeping with standard Internet typing. The structural part of the aor therefore would contain the following text if it were a Web page.

```
<type>HTML document text</type>
```

8.2.2.1.2 Layout

The layout encodes the embedded objects that make up the Web page. A Web page is an HTML document and a collection of embedded objects that is rendered by the user agent to produce the multi-media information object on the screen. The layout constituent of the aor tries to capture this by encoding the Web page as a list of text objects and other media objects. Each text keyword is followed by a pair of numbers representing the amount of text (in bytes) for display in the Web page (visible text) and the text that is part of the mark-up and will not be directly shown to the user (invisible text). Each type of object is followed by the size of the object in bytes and then a letter to indicate if it is located on the same server to the HTML document. The following text shows the layout for one particular Web page.

```
<layout>text(477,6698);image(1282,I);text(16,688);
image(1942,I);text(123,123)</layout>
```

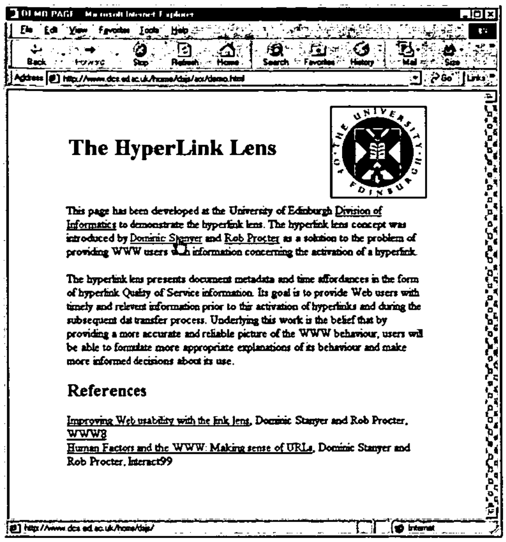
This Web page has two internal images of size 1282 bytes and 1942 bytes surrounded by text with 616 bytes of visible text and 7509 bytes of invisible text (i.e. HTML mark-up). A WWW client must therefore transfer 11349 bytes of data in order to render this Web page.

8.2.2.1.3 Hyperlinks

The final part of the structural information contains a list of hyperlinks on the Web page. This is simply a list of letters to represent whether each hyperlink on the Web page points to a resource with the same domain name (internal) or another domain name (external).

```
<hyperlinks>eeeii</hyperlinks>
```

The complete structural information for an example Web page is shown in Figure 8.10. It is from this information alone that the prism visualisation is constructed and the size of every object determined to the animation during data transfer.




```
<structure>
  <type>HTML document text</type>
  <layout>text(634,63);
    image(7654,i);
    text(2246,1138)</layout>
  <hyperlinks>eeiiiei</hyperlinks>
</structure>
```

Figure 8.10: The structural information of the annotated object reference

8.2.2.2 Contents preview information

The information encoded for the overview is basically a subset of the Dublin Core metadata shown in Figure 8.7. The elements presently are limited to the title, creator, publisher, date and description. It is clearly a simple process to add additional elements from the Dublin Core set of elements, or indeed from another metadata set.

```
<aor>
      <structure>
            <type>HTML document text</type>
            <layout>text(634,63);image(7654,i);
                    text(2246,1138)</layout>
            <hyperlinks>eiiiei</hyperlinks>
      </structure>
      <overview>
            <title>Hyperlink Lens Demo Page</title>
            <creator>Dominic Stanyer</creator>
            <description>A brief description of the hyperlink lens
                        concept, an application that presents
                        document metadata and time
                        affordances.</description>
            <publisher>Edinburgh University</publisher>
            <date>2000-07-28</date>
      </overview>
</aor>
```

Figure 8.11: The annotated object reference

In contrast to the structural information that can be collected automatically, the existence of overview metadata is dependent on the author of the Web page. A few of the metadata elements can be automatically generated and work is presently underway within the Dublin Core community to improve this. Nevertheless, an annotated object reference can only contain this overview information if it is specified in the HTML document of a Web page by its author. The annotated object reference for the example Web page is shown in Figure 8.11.

The final addition to the annotated object reference is the thumbnail image of the Web page. This is simply added to the annotated object reference as a further element with tag thumbnail. The problem with this element is, of course, that the jpeg encoded data appears in the annotated object reference and can be rather large and is not human readable. Although the present system allows the addition of the thumbnail, this approach is not ideal and is an area for further work.

8.2.3 Annotating the HTML document

The anchor element of a hyperlink must contain an aor attribute if the *hyperlink lens* is to display the structure and overview information about the referenced Web page. For the *hyperlink lens* to interact with a Web page in the way described at the beginning of this chapter, the HTML document that defines the Web page must define an aor attribute in every anchor element it contains. The HTML document must be annotated with an aor attribute in every anchor element.

To annotate an HTML document each anchor element must be extracted and an annotated object reference constructed for the resource it references. Constructing an annotated object reference involves collecting the structural and overview information. To collect the overview information, the HTML document of the referenced resource must be downloaded and the metadata extracted from the HEAD section of the HTML document if it exists. Collecting the structural information also involves downloading the HTML document of the referenced resource and then parsing it to extract all the embedded objects and hyperlinks. Each of the embedded objects must then be downloaded to determine their size and type. Constructing the structural information is therefore automatic, and not dependent on any author intervention. The process may take a considerable amount of time, as every hyperlink on the referencing Web page must be downloaded.

Once an annotated object reference has been constructed, it must then be added to the appropriate anchor element in the form of an aor attribute. When this process has been completed for every anchor element in the original HTML document, this document can be considered annotated. The annotating process can be undertaken at a number of locations in the network and at a variety of times in the data transfer process.

Server side annotation - before the HTML document is sent to a client it can be annotated. The server stores an annotated version of the HTML document and sends this file when an HTTP request is received. Using this method makes the annotation process a form of publication. When an author stores an HTML document on a Web server, the file is automatically annotated. It is the WWW server's responsibility to add the annotations and subsequently renew and update the annotations using an appropriate scheme.

Client side annotation – the annotation is undertaken after the HTML document is transferred. This can be achieved by a proxy using an approach similar to that taken by Kopetzky and Muhlhauser (1999) in their visual preview system. The disadvantage of this approach is that time taken constructing and inserting each aor is time directly added to the data transfer delay.

The program (written in C++) presently available for constructing annotated object references of Web pages and annotating HTML documents, takes a considerable amount of time to perform an annotation. This is due to the time involved with transferring the data for all the referenced HTML documents and their embedded objects. It is, therefore, presently infeasible to allow this additional temporal overhead during the data transfer process of a Web page. The *hyperlink lens* system relies on the server side annotation of HTML documents. The server side approach also clearly illustrates the major concept that the annotated object reference is an attribute of the hyperlink.

8.2.4 Implementing the Magic Lens

When the WWW client requests a Web page it is sent an annotated HTML document and a collection of embedded objects. Before displaying the Web page, the annotated HTML document must be modified to allow the *hyperlink lens* to display the information contained in each hyperlink's definition. A number of JavaScript functions are therefore added to the annotated HTML document by a proxy to allow the communication between the WWW browser and the lens. These functions define the *remote window* that contains the hyperlink lens and a method to communicate a hyperlink's aor to the remote window when the mouse pointer is moved over the activating object. The HTML document has therefore undergone a number of transformations before it is presented to the WWW browser and the *hyperlink lens* (Figure 8.12).

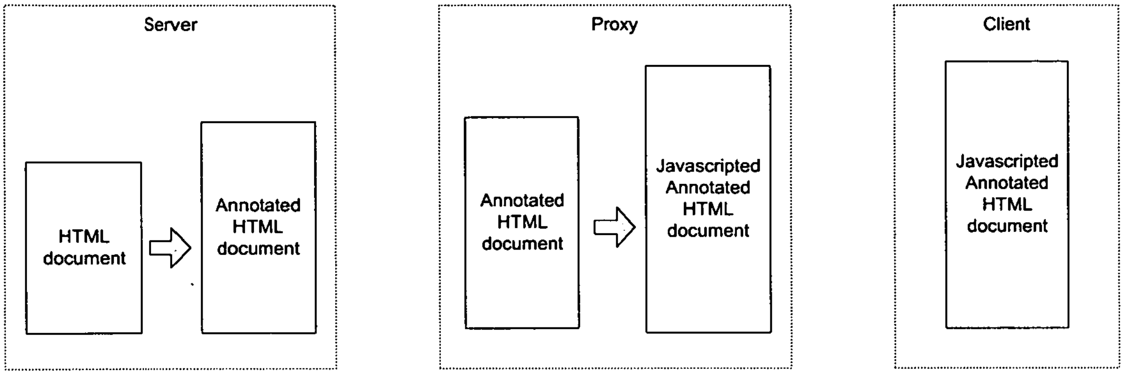


Figure 8.12: The transformation of the HTML document

Displaying the hyperlink information visualisation is undertaken by a Java applet that is executed in the remote window. This applet receives an input of an aor every time the user moves over an activating object. This input is implemented as an applet-parameter that allows the applet to parse and display the information it contains. Once hyperlink activation has started, the applet is then required to communicate with the proxy to show the temporal affordance of the communication. As the applet is aware of the number and size of every object that must be retrieved before the referenced Web page can be rendered, the proxy only has to tell the applet how much of a particular object it has received.

The three main elements of the magic lens implementation are shown in Figure 8.13.

- **Proxy** – adds the JavaScript functions to an annotated HTML page **before** sending this on to the browser and, communicates to the applet the amount of data it has receives **after** hyperlink activation.
- **Java applet** – parses and displays the hyperlink information and communicates with the proxy to animate the visualisation during the data transfer progress.

- **JavaScript** – defines the remote window that contains the Java applet and facilitates communication between the applet and the WWW browser so that the hyperlink lens changes as the mouse pointer moves around the Web page.

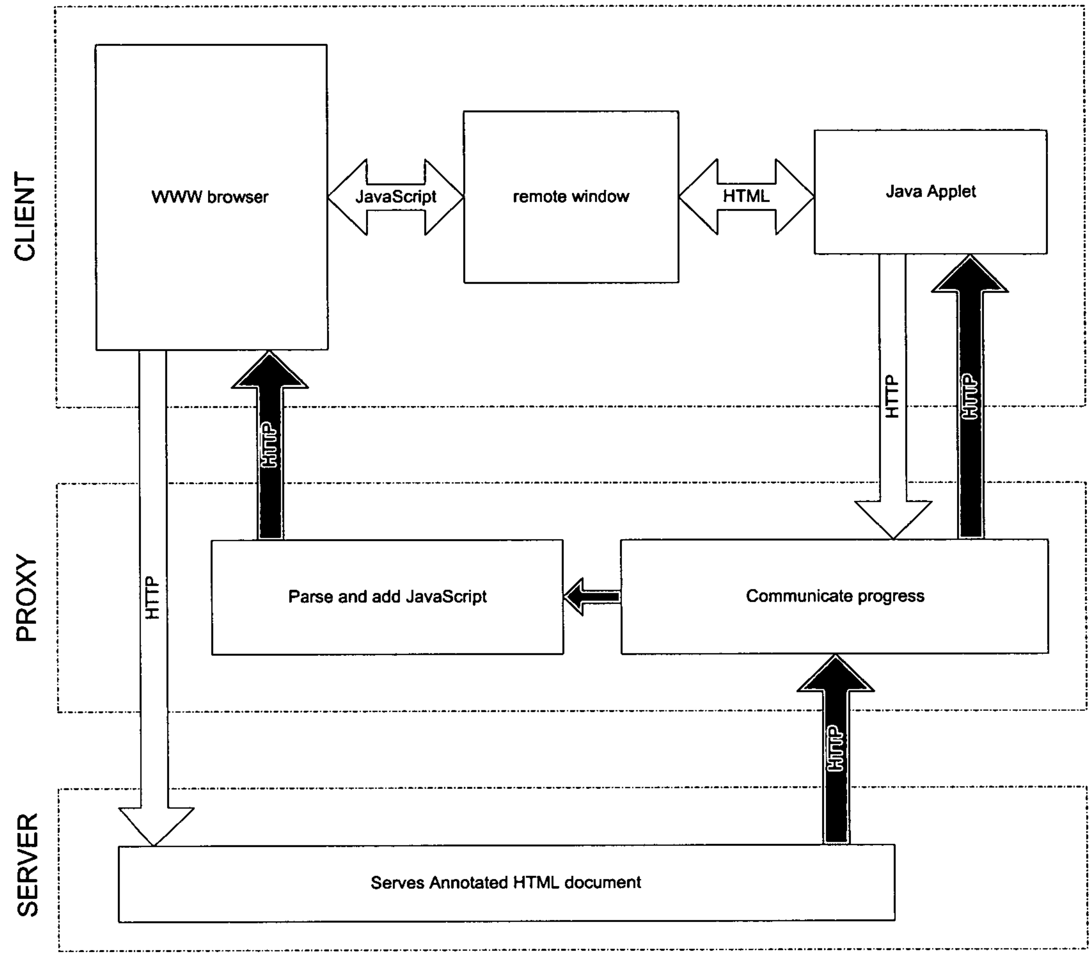


Figure 8.13: Hyperlink lens architecture overview

8.2.4.1 Flow of events

A typical sequence of events is described in the following steps.

1) A user activates a hyperlink that references an annotated HTML document.
2) The user's browser requests the resource from the proxy server.
3) The proxy server requests and subsequently receives the annotated HTML document.

4) The proxy parses the annotated HTML document and inserts JavaScript code to enable communication between the browser and the applet executed in the remote window.
5) The user's browser receives the annotated HTML document now enhanced with JavaScript code.
6) The user's browser requests and subsequently receives all the embedded resources referenced by the HTML document.
7) The Web page is displayed to the user.
8) Whenever the user moves the mouse pointer over the activating object of a hyperlink the hyperlink lens displays information about the referenced Web page.
9) When the user activates a new hyperlink the applet starts to communicate with the proxy and displays the data transfer process as steps 2-6 are repeated.

8.2.4.2 Applet Security

This overview of the hyperlink lens architecture presents a simple system that appears to require the user to set up the WWW browser to use a special proxy and, execute the Java applet locally. The actual implementation, however, prevents both of these user-actions from being necessary. To use the *hyperlink lens* a user only has to request and download a special initialisation Web page stored on the hyperlink lens's Web server. The user is then requested to enter the first URL of the Web page they wish to browse. From this point on, all communication takes place through the hyperlink lens's Web server. This Web server can now be considered the proxy in the architecture.

The initialisation Web page runs a cgi-script that acts as the proxy performing both the JavaScript addition and the communication with the applet. Indeed, what appears to be a rather convoluted approach to the implementation is in fact required to circumvent some of the restrictions enforced by the Java applet. Java applets that are downloaded over the network are untrusted (Flanagan, 1996). The Web browser therefore carefully restricts what an applet is allowed to do. One of these restrictions is that an applet is **not** allowed to create a network connection to any computer other than the one from which the applet was itself created. This restriction is the primary reason for implementing the proxy as a cgi-script on the hyperlink lens's Web server (*i.e.*, where the applet is stored). As this Web server is now both the proxy and the computer from where the applet was loaded, a network connection can be established and the progress information communicated.

This approach of implementing the proxy as a number of cgi-scripts executed on a Web server has an implication on the communication between the client and the WWW server. Normally, a proxy will receive an HTTP request from the client containing a normal URL and forward this on to the WWW server (as explained in Chapter 3). The location of the proxy is a parameter of the browser. By using a Web server as a proxy, all HTTP requests must first be communicated to this Web server (proxy) and then forwarded to the actual WWW server that is specified by the URL. All URLs must therefore be changed so encode the location of the proxy. This is achieved by using the query part of the URL. A URL specified in the href attribute of a hyperlink like, for example, `http://www.go.com/` becomes, `http://www.hyperlink_lens_Web_sever.com/cgi-bin/proxy?http://www.go.com/`. The proxy can now forward the original URL on to the correct WWW server, and then parse and change the data it receives before passing it on back to the client. This complication to the implementation does not change the architecture of the hyperlink lens system; it only enforces a few further changes to HTML documents as they pass through the proxy.

8.3 Summary and Limitations

The hyperlink lens system presents a working system that simulates how a Web browser could be changed to present document metadata and time affordance encoded in the hyperlink. If a Web browser were to implement a hyperlink lens, the actual change to its implementation would be small and simple to accomplish. The real change to the WWW is getting structural information about the referenced resource into the definition of a hyperlink. This is a radical departure from the model of the WWW that is presently used.

Annotating HTML documents increases the amount of data that needs to be transferred from the Web server to the client. The more hyperlinks a Web page contains the more aors it will contain and hence, the size of the transferred data will increase. Similarly, a Web page that contains many images will require a longer aor than a Web page with just a few. The impact that this increase in HTML document size will have on data transfer process is an area for further work. It must, however, be noted that by providing the hyperlink lens there may be less redundant requests for Web pages that turn out to be useless and so an increase in HTML document size must be traded with a decrease in pointless Web page transfers.

Another potential advantage of enforcing a publication process is that the HTML document can be checked. Not only does this mean that the existence of every referenced resource can be confirmed, but also the HTML document itself can be syntactically checked and optimised (*i.e.,* by extracting whitespace). What is more, the annotation process does not have to be enforced. If an HTML document has not been annotated the *hyperlink lens* will not be able to present any metadata, an indication in itself of quality.

As it stands, the major limitation with the *hyperlink lens* architecture is that it does not cope with dynamically created HTML documents. Dynamic HTML, as the name suggests, is a method to construct HTML documents that are dependent on the contents of the HTTP request. Different clients can receive different HTML documents allowing users that have accessed the server before to be served personalised Web pages. If the content of the Web page is not known before the actual request, how can it be encoded in an aor and associated with a hyperlink. In reality, the actual difference between dynamically generated Web pages is small and it may be possible to make general assumptions about Web pages' content and presentation.

This chapter presented the design and implementation of the *hyperlink lens* architecture. Clearly, this particular implementation is only the first stage of the design lifecycle. The design needs to be analysed and evaluated before the new cycle can begin. Section 8.1.2 explained how the focus groups were used to help in the initial design of the present implementation, and a similar process needs be undertaken before the design can be augmented. The aim of the *hyperlink lens* is to help users understand the information space and the unpredictable dynamic behaviour of the system. The *hyperlink lens* therefore needs to be analysed using a method similar to that used for the Web browser experiments presented in the previous chapters of this dissertation. This analysis would form the basis for an evaluation of the present *hyperlink lens* design and provide some useful feedback for the next iteration of the design. Are users extracting more appropriate explanations of the behaviour they experience than with a normal WWW browser? Do these more appropriate explanations improve the usability of the WWW?

The evaluation process should involve situated observation of the user whilst using the *hyperlink lens* enhanced Web browser. It is clear that the *hyperlink lens* will be a new interface feature and so users will need to become accustomed to its use. The learning process will be an interesting part of the evaluation and potentially very useful with respect to the design improvements, but it would need to be factored out of any comparison to the experiments of just the Web browser. How a subject's explanations and understanding alters through the use of the *hyperlink lens* is what needs to be elicited. If a comparison were to be made with the experiments presented in this dissertation, subjects would need to be recruited and trained on the enhanced WWW browser. The experimental methods employed would need to determine whether a user's understanding formed solely through the use of a normal Web browser could be improved through the use of the *hyperlink lens*.

A possible experiment could involve only using subjects who had never previously used the WWW. By exposing a sample of subjects to the WWW using the enhanced Web browser and comparing their descriptions of the information space and explanations of the system behaviour to another sample of subjects who simply used the traditional Web browser. Clearly, this form of evaluation would take a considerable amount of time but it would reveal the impact that the *hyperlink lens* would make to the usability of the WWW.

In addition to the observational experiments, the *hyperlink lens* could also be analysed for its ability to provide information before hyperlink activation. Using a similar method to the URL experiments presented in a previous chapter, subjects could be presented with the *hyperlink lens* output and asked to predict what the referenced Web page contained. This could then be compared to the predictions based solely on the URL. It is hoped that these experiments will show that by revealing the underlying mechanisms of the Internet communication and the information space that surrounds a particular Web page, the user is better able to form explanations and understand the behaviour they experience.

It is clear from this discussion that the *hyperlink lens* is only the first stage of the design process. It needs to be developed and refined if it is to present the appropriate information to the user about the information space and the dynamic properties of the WWW. The continuation of this design lifecycle is beyond the scope of this dissertation but the experiments presented in previous chapters form a guide as to how this design process should continue.

The *hyperlink lens* demonstrates how the WWW abstractions of the Web page and the hyperlink can be enhanced to present more system information to the user. The observed behaviour of WWW users, presented in this dissertation, highlighted the need for the interface to provide a clearer picture of the actions it performs. The *hyperlink lens* does not destroy the Web page and hyperlink abstractions, which have been shown to be central in users' mental models and understanding of the system, it augments them by revealing the production of system actions and the consequence of a user's action.

Chapter 9

Conclusion

The work presented in this dissertation has shown, through a number of field studies and laboratory experiments, the problems users have in extracting sufficiently complete explanations through the interface. The investigations uncovered the strategies users employ to cope with system disturbances, the improvised nature of their actions and their resourcefulness in the quest for more enlightening system information. The hypertext abstraction the WWW imposes on the underlying Internet technology restricts the system's ability to explain why it behaves as it does in a given situation. Users are left trying to circumvent the abstraction to reveal properties of the system that will directly, or through implication, answer the questions of their interrogation.

This final chapter will consolidate all of the material presented thus far. It has been shown that there exists an asymmetry in the human interaction with the WWW due to the hypertext abstraction constraining the needs of the user. This results in users employing coping strategies to evade the abstraction barrier. The solution is an interface that resolves the asymmetry by augmenting the WWW abstraction and providing more appropriate information - the *hyperlink lens*. The chapter will finish with a presentation of the areas where further work should be encouraged.

9.1 The hypertext abstraction

This dissertation began by presenting various models of hypertext. The WWW is based on the simplest of these models and the hypertext abstraction that it imposes on the system consists of two major components – information nodes and the connections between them. For the WWW the information node is the Web page – a collection of multimedia objects referenced and positioned by a defining HTML document – and the connections are hyperlinks – pointers embedded in a HTML document that reference the location of another resource. When a hyperlink is activated, the resource it references is transferred to the user's display using Internet technology. If this resource is a Web page, the HTML document and all the embedded objects it references are transferred from the servers, where they are stored, to the client. Once data transfer is accomplished, the Web page can be constructed and displayed to the user. The Web page is a conceptual entity that only exists as an object when it is completely rendered in a user's browser. The Web page is a manifestation of its constituent data objects.

The hypertext abstraction attempts to shield the user from the complexities of the data transfer processes. The abstraction splits the system up by concealing the Internet technology behind the Web page and the hyperlink. In a similar fashion to the London Underground Diagram presented in Chapter 1, the hypertext abstraction is relevant and useful in some situations but too restrictive in others. The abstraction hides the workings of the system. Throughout the investigations, users' explanations employed the Web page and hyperlink concepts. Indeed, in the users' visualisations of the WWW the vast majority used the Web page as the fundamental element. When the WWW is in *normal* working order the Web page is an abstraction that users can conceptualise and successfully use: the underlying system mechanics are transparent to the user. When system breakdown or disturbance occurs, however, – usually in the form of delays in data transfer - users start the processes of inspection and practical problem solving and the Web page abstraction becomes constrictive, and as this is the only abstraction that the interface supplies, users struggle to explain the reasoning behind the disturbance. The result is that users are unsure what the best strategy is to resolve the problems: does the user interrupt and reload or simply wait?

The Web page abstraction is not sustained during WWW use. Rather than experiencing a deterministic response to the activation of a hyperlink, the user often copes with an unpredictable delay related to Internet performance fluctuations and failures. The reality of the data transfer processes initiated in response to a Web page request reveal themselves in the form of delays. Due to these disturbances, users' understandings of the system change as they gain expertise. To begin with, users accept the hypertext abstraction and explain hypertext activation as a movement through the hypertext. The novice user navigates through the hypertext. As users gain experience, the reality of WWW use becomes apparent and this basic understanding starts to change. The unpredictable temporal behaviour of the data transfer process becomes a central concern and users' understandings are adapted so that each Web page has an associated delay. Hyperlink activation is no longer a navigational action that initiates seamless travel to a new Web page. Users start to realise that the Web page is coming to them. The hypertext abstraction has started to fail.

Users acknowledge that the delays they experience after hyperlink activation are inconsistent. The time it takes to download a Web page changes from one circumstance to the next. Explaining to the user the reasons for this form of variance is unsatisfactory. What users can do is give plausible accounts why *different* Web pages take a different amount of time to download. This reasoning centres on uncovering the geographical location of the data. Users associate a Web page with a geographic location. From this piece of spatial information, users explain temporal behaviour. Making an accurate prediction about the actual behaviour of the Internet technology is outside the understanding that most users' possess, but using a simple set of heuristics that describe the relationship between space and time and eliciting a geographical location for a Web page, users do generate explanations for delays. The extensive use of this strategy, associating a Web page with a geographical location, reveals that some part of the interface must provide this spatial information. Although it is not part of the hypertext abstraction, users still manage to extract this information. This demonstrates the resourcefulness of users as they circumvent the hypertext abstraction to elicit further properties of the system, and the restrictive nature of the hypertext abstraction.

These general accounts of WWW understanding do provide an insight into how understanding develops with expertise. The accounts do not explain, however, the resourcefulness and improvised nature of users' actions when they actually experience a disturbance. Although users hang on to the concepts of Web pages and hyperlinks, the hypertext abstraction continually collapses when users are faced with a disturbance. As these types of disturbance happen, users try to find out first, what is happening and second, why it is happening. The hypertext abstraction is at odds with the behaviour the user is experiencing and even though some effort has been made by browser developers to provide information about the data transfer process, it is inappropriate and insufficient. The result is that users look elsewhere for explanation.

9.2 The Universal Resource Locator

The URL is a reference to a single WWW resource: an HTML document, an image, a movie *etcetera*. It was designed to be internal to the system, buried in the anchor elements of an HTML document and used to facilitate hyperlink activation. Its appearance in the browser, in the address bar and the status bar, and the fact that the user can directly manipulate it led to the URL becoming a central source of system information. The URL is not part of the hypertext abstraction - it is an internal but visible system mechanism – and so when users attempt to interrogate the system for more appropriate information when a disturbance occurs, the URL is easily at hand and has the potential to reveal some of the workings below the abstraction. Although not designed to help the user in this way, the URL is a window through the hypertext abstraction onto the Internet technology below. It is a key element in a user's inspection and subsequent understanding of the system and a resource for improvising when the abstraction fails

The URL possesses some interesting properties potentially at conflict with its observed use by users. It is non-functional – a URL can reference a number of different resources, it is non unique – different URLs can reference the same resource, and it is temporally and spatially variant – the same URL requested from different locations can reference different resources. Although these properties are often transparent to the user, due to the URL's actual use by authors and the inability of the user to know what other users see, relying on the URL to describe what it references is fraught with inconsistencies. Moreover, users employ the URL in their attempts to describe a Web page when in fact the URL only references an HTML document.

What is surprising, therefore, are the rich and often accurate predictions that users are able to extract about a Web page. Indeed, users were very confident about their ability to predict the content, spatial location, and the download behaviour of a Web page solely from the URL of the HTML document. However, this confidence is often misplaced when users were required to make very specific predictions as the URL actually affords very little to help the user predict the behaviour of the data transfer process. The only consistent information the URL contains is in the domain name, and it is in the domain name where spatial information is provided, in the form of a TLD.

The URL is an inconsistent and untrustworthy resource with which to predict the content, location, and potential data transfer properties of a Web page. The fact that users employ this resource is a reflection on the lack of appropriate information deliberately offered by the interface. The information that users struggle to accurately reveal from the URL must be presented in a more appropriate fashion by the interface itself.

9.3 The hyperlink lens

Having shown the asymmetry in the interaction between the user and the WWW, and presented some of the improvisations and resourcefulness of users in their attempts to break down the abstraction barrier, this dissertation presents an enhancement to the browser that addresses this problem. The implementation – *the hyperlink lens* - centres on redefining the hyperlink to provide more appropriate system information by showing document metadata and the production of the data transfer actions. By deconstructing the Web page abstraction and showing the actual communication process of each object provides the user with a solid basis for explanation when disturbances occur.

The interface provides different levels of abstraction but does not commit itself as to how these abstractions may actually be used. The document metadata and time affordances presented through the hyperlink lens show the system's data transfer actions, and an explanation of the system's behaviour. Exactly how this information is used is dependent on the particular situation in which users find themselves. This dissertation has argued that disturbances centre on the unpredictable nature of the data transfer process and the very fabric of the information space with the possibility of linkrot. Various situated examples of these disturbances were observed and presented in this dissertation. The reason that the hyperlink lens is predicted to help in these situations is not simply because it has been designed to do so, but rather that it provides a consistent picture of the production of actions.

The *hyperlink lens* is an example of how the interface can present users with information to aid their enquiries when their circumstances demand. It was designed to present the information that users may need to know when faced with disturbances but struggle to find through the browser. The *hyperlink lens* augments the Web page and hyperlink abstractions.

9.4 Further Work

To complete the thesis a few suggestions for further research are presented. A lot of material has been covered in this text; spanning topics such as, situated action, affordance, ethnomethodology, hypertext, the Internet, chess expertise, experimental validity, causation, ethnography, magic lenses and Dublin Core. There is, therefore, much that could be investigated further. To contain the discussion, the following three topics have been selected as they offer the most potential for expansion.

9.4.1 Continued investigation of WWW users.

Throughout this dissertation, the work of other researchers was compared and contrasted with the findings generated by the field studies and laboratory experiments of this work. The idea was to provide a basis for verifying the external validity of the work undertaken in this dissertation: showing that the findings could be generalised across settings and populations as the observed user's behaviour and the Web pages they requested, were comparable to other studies. Often the findings of the investigations confirmed previous studies, but on occasion there was conflict – for example, whether the WWW is a recurrent system. The WWW is a vast information space whose content is growing and shifting. The people that use the WWW are numerous, diverse and continually changing in expertise and general computing knowledge. The information is continually changing, the technology that transfers the data is continually changing, and the users are continually changing.

The first area for further work must therefore be to continue the investigations into the characteristics of the Web page and the behaviour of the WWW user. This will not only help to form a better understanding of the WWW but it will also highlight the changes that have occurred.

9.4.2 The hyperlink lens architecture

The *hyperlink lens* demonstrates how the principled philosophy of interaction and the findings from a rigorous set of investigations on WWW use is applied to the design of an interface. Its implementation was complicated by the need to augment an existing browser. Changing the actual browser to implement some of the design principles demonstrated by the *hyperlink lens* would not be very difficult. What must first be undertaken, however, is an evaluation of the actual impact of the *hyperlink lens* on the usability of the WWW. This dissertation has presented a philosophy of interaction with some supporting experimental evidence, but the *hyperlink lens* is only a demonstration of how this approach can be applied, not necessarily how it should be applied. Users must be observed actually using the *hyperlink lens* and this investigation must be compared and contrasted with the investigations of WWW users presented in this dissertation. It is hoped that these investigations will confirm that usability is positively affected by *hyperlink lens* use.

As well as the usability studies of *hyperlink lens* use, there also needs to be some further investigation into the impact that the architecture will have on the data transfer process. In particular, the introduction of the annotated object reference into the HTML document will increase its size and therefore, affect the time taken to transfer the file. This must be contrasted, however, with the potential time gained through the elimination of unnecessary hyperlink activations due to the user being better able to decide on the relevance of a referenced Web page. Finally, the impact of compiling HTML documents to introduce the annotated object references is another area that needs investigation.

9.4.3 The mobile Web

Communications technology is changing from desktop computers fixed to wire networks, to mobile devices using wireless technology. As mobile devices develop, people have started to use what is often referred to as the mobile Web. Rather than just using the Internet technology of TCP/IP and HTTP, the mobile Web also employs the Wireless Application Protocol (WAP) to transfer data from Internet gateways to mobile devices. In a similar fashion to HTTP being designed to transfer HTML documents and other resources, the WAP is designed to transfer resources described in the Wireless Mark-up Language (WML)[1].

[1] The basic difference between HTML documents and WML decks are that the WML resources must be compiled before transfer.

The size limitations of mobile devices mean that screen real-estate is often restricted to only a few lines of text and simple images. This imposes severe constraints on the content of the *mobile Web pages.* Despite the difference between the WWW and the mobile Web, they are both distributed information spaces with information nodes and hyperlinks. Applying the philosophy of interaction proposed in this dissertation to this new area of mobile communication would be an area of further research worth pursuing. The delays incurred during data transfer should be explained to the user using a similar method to that proposed in this dissertation. The screen real-estate restrictions and the very nature of the mobile devices would be an interesting application of the interaction techniques and philosophy uncovered and used in this dissertation.

9.4.4 The social WWW

One feature of the WWW that has been referred to continually throughout this dissertation is the social nature of the information space. The author of a Web page provides a form of recommendation through the existence of a hyperlink, but a more accurate, or personalised, form of rating the quality or relevance of a Web page is desirable. Explicit ratings by other users can be collected and presented, and through filtering or weighting of this user population, different forms of social recommendation can be calculated. Similarly, implied ratings can be collected based on evidence such as the popularity of the Web page. The provision and use of Web page recommendations is an area of research that is related to the findings of this dissertation and an area of ongoing research.

9.5 Conclusion

There is disparity between the desires of the user to probe the WWW and to reveal explanations for the system behaviour currently being experienced, and the actual information offered by the interface. The user wants to explore and enquire in order to make sense of their current situation but the browser only offers a single static abstraction – Web pages connected by hyperlinks - in explanation of every circumstance. This dissertation has shown the existence and effect of this asymmetry, the coping strategies users employ and the resourcefulness of their enquires. The asymmetry in the interaction constrains users' understanding and hence the usability of the WWW. The problem is resolved by balancing the interaction, achieved through an interface that allows the user to explore its workings by presenting an account of the system, which explains in detail the production of its actions and the subsequent behaviour.

Bibliography

Albert, R., Jeong, H. and Barabasi, A. (1999), Diameter of the World-Wide Web, Nature 401, pp 130-131, 9th September 1999

Atkinson, M. (1984), Our masters' voices, Methuen, 1984

Augustine, M. A. and Covert, M.D. (1991), Simulation and information order as influences in the development of mental models, SIGCHI Bulletin 23, 1, pp 33-35, January 1991

Ayers, E. Z. and Stasko, J.T. (1995), Using graphic history in browsing the world wide web, Tech. Rep. GIT-GVU-95-12, Georgia Institute of Technology, Atlanta, May 1995

Barabasi, A., Albert, R. and Jeong, H. (2000), Scale-free characteristics of random networks: the topology of the world wide web, Physica A, 281, 2115 , 2000
`http://www.nd.edu/~alb/paper/PHYSICA281-2115.pdf`

Berners-Lee, T., Cailliau, R., Luotonen, A., Nielsen, H.F. and Secret, A. (1994), The world-wide web. Communications of the ACM 37, 8 (Aug. 1994), 76-82

Berners-Lee, T. (1999), Weaving the Web: the past, present and future of the World Wide Web by its inventor, Orion business books, 1999

Bieber, M., Vitali, F., Ashman, H., Balasubramanian, V. and Onias-Kukkonen, H. (1997), Fourth generation hypermedia: some missing links for the World-Wide Web, International Journal of Human-Computer Studies 47 (1997) 31-65.

Binet, A. (1894), Psychologie des grands calculateurs et joieurs d'echecs, Hachette Paris, 1894

Borgman, C, L. (1984), The user's mental model of an information retrieval system: effects on performance, Stanford University, Ph.D., 1984

Bray, T. (1996), Measuring the Web, the World Wide Web journal 1(3), 1996.
http://www5conf.inria.fr/fich_html/papers/P9/Overview.html

Brebner, G. (1997), Computers in communication, McGraw-Hill, 1997

Bush, V. (1945), As we may think, the Atlantic monthly, July 1945
http://www.ps.uni-sb.de/~duchier/pub/vbush/ubush-all.shtml

Catledge, L.D. and Pitkow, J.E. (1995), Characterizing Browsing Strategies in the World-Wide Web. Computer Systems and ISDN Systems: Proceedings of the 3rd International World Wide Web Conference. April 10-14, Darmstalt, Germany, Vol. 27, pp 1065 –1073

Cerf, V.G. and Kahn, R.E. (1974), A protocol for Packet Network intercommunication, IEEE Transactions on Communications, May 1974, vol Com-22, No 5, p.637-648

Chase, W. G. and Simon, H. A. (1973), Perception in Chess, Cognitive Psychology 4, pp 55-81, 1973

Cheswick, B. and Burch, H. (1999), Internet Mapping Project, 1999
http://www.cs.bell-labs.com/~ches/map/index.html

Cockburn, A. and Jones, S. (1996), Which way now? Analysing and easing inadequacies in WWW navigation. International Journal of Human-Computer Studies 45.(1). pp 105-129

Conklin, J. (1987), Hypertext: an introduction and survey, IEEE Computer, pp 17-41, September 1987

Conn, A. (1995), Time Affordances: the time factor in diagnostic usability heuristics, in Proc. CHI '95, ACM Press, pp 186-193, 1995

Cook, T.D. and Campbell, D.T. (1979), Quasi_Experimentation: design and analysis issues for field settings, Houghton Mifflin Company, Boston, 1979

Cotton, B. and Oliver, R. (1993), Understanding Hypermedia; from multimedia to virtual realty, Phaldon Press, 1993

Cunha, C.R., Bestavros, A., and Crovella, M. (1995), Characteristics of WWW client-based traces, Tech. Rep. BU-CS-95-010, Boston, MA, Computer Science Dept., Boston University, 1995.
http://www.cs.bu.edu/techreports/95-010-www-client-traces.ps.Z

Dix, A.J. (1994), Seven years on, the myth continues, (RR9405), 1994,
http://www.soc.staffs.ac.uk/~cmtajd/papers/myth95/7-years-on.html

Dodge, M. (1999), This is what cyberspace looks like, The Guardian, Online page 3, Thursday October 28, 1999

Dourish, P. and Chalmers, M. (1994), Running out of space: Models of information navigation, HCI'94, Glasgow, Aug. 1994
http://www.research.apple.com/people/dourish/papers.html

Dourish, P. and Button, G. (1998), On technomethodology: foundational relationships between ethnomethodology and system design, Human computer interaction, Vol 13, 395-432, 1998

Emerson, R. M., Fretz, R. I. and Shaw, L. L. (1995), Writing Ethnographic fieldnotes, The University of Chicago Press, 1995

Flanagan, D. (1996), Java in a Nutshell, O'Reilly & Associates, February 1996

Foss, D.J. and Deridder, M. (1987), Technology transfer: on learning a new computer-based system, in interfacing thought, ed. J.M Carroll, The MIT Press, Chapter 7, pp 159-183, 1987

Furnas, G.W. (1997), Effective view navigation, in Proceedings of the ACM CHI '97, Human Factors in Computer Systems, 1997, pp 367-374

Garfinkel, H. (1967), Studies in Ethnomethodology, Prentice Hall, 1967

Garland, K. (1994), Mr Beck's underground map, Capital Transport Publishing, Middlesex, 1994

Geertz, C. (1973), Thick description: towards an interpretive theory of culture, in (ed) Geertz, the interpretation of cultures, 3-30, Basic Books, 1973

Gibson J.J. (1966), The senses considered as perceptual systems, Boston: Houghton Mifflin, 1966

Giles, J. (1999), Why the world is your neighbour thanks to 19 clicks of separation, The Guardian, Online page 2, Thursday October 28, 1999

Goguen, J.A. (1992), The Dry and the Wet, in Information Systems Concepts: Improving the Understanding. Proceedings of IFIP Working Group 8.1 Conference, 1992

Goguen, J.A. (1994), Requirements engineering as a reconciliation of social and technical issues, in Requirements engineering: social and technical issues ed. M. Jirotka and J. Goguen, chapter 7, pp165-199 1994

Gray, W. and Salzman, M. (1998), Damaged Merchandise? A review of experiments that compare usability evaluation methods, Human-Computer Interaction Vol. 13, No 3, 1998

Greenberg, S. (1993), The Computer User as Toolsmith: The uSe, Reuse and Organisation of Computer-based Tools, Cambridge, MA : Cambridge University Press, 1993

Greenberg. S. and Cockburn, A. (1999), Getting Back to Back: Alternate Behaviour for a Web Browser's Back Button, Proceedings of the 5th Annual Human Factors and the Web Conference, Held at NIST, Gaithersburg, Maryland, USA, June 3rd 1999

Gronboek, K. Bouvin, N.O. and Sloth, L (1997), Designing Dexter-based hypermedia services for the World Wide Web, In Proceedings of Hypertext 97, Southampton UK, 1997

De Groot, A.D. (1965) Thought and Choice in Chess, Psychological Studies 4, Mouton and Co, 1965

Gwertzman, J. and Seltzer, M. (1996), World Wide Web cache consistency, in: Proceedings of the 1996 Usenix Technical Conference, Boston, MA, Harvard College, 1996 http://www.eecs.harvard.edu/~vino/web/usenix.196/

Halasz, F. and Schwartz, M. (1994), The Dexter Hypertext Reference Model, Communications of the ACM, 1994, 37(2): pp 30-39

Hartston, W.R. and Wason, P.C. (1983), The Psychology of Chess, Batsford, 1983

Heath, C.C., and Luff, P.K., (1992), Collaboration and control: crisis management and multimedia technology in London underground line control rooms, CSCW Journal 1:69-94, 1992

Heath, C.C., Jirotka, M., Luff, P.K. and Hindmarsh, J. (1993), Unpacking collaboration: interactional organisation in a city trading room, Proceedings of the Third European Conference on Computer Supported Cooperative Work, pp. 155-170, 1993

Hofstadter, D.R., (1979), Godel, Escher, Bach: An eternal golden braid, Penguin, 1979

Huberman, B.A. and Adamic, L.A. (1999). Growth dynamics of the World-Wide Web, *Nature*, vol. 40, pp 450-457, 1999

Jirotka, M. and Gougen, J. (1994), Requirements Engineering; social and technical issues, Academic Press, 1994

Johnson, P, and Kavanagh, J. (1996), Electronic gridlock: reducing download latency for video on the web, in A. Blanford, H.Thimberley, (Eds.), Proceedings of the 11th British Computer Society Annual Conference on Human Computer Interaction, HCI'96, August 1996

Johnson, P. (1997), What's the Web Worth? The impact of retrieval delays on the value of distributed information, In time and the Web, Stafford, June 1997

Kopetzky, T. and Mulhlhauser, M. (1999), Visual preview for link traversal on the World Wide Web, Proceedings of the eighth International WWW Conference, Toronto, 1999

Lansdale, M. and Ormerod, T. (1995), Understanding Interfaces: A handbook of Human-Computer Dialogues, Computers and People Series Academic Press, chapter 7. pp 164-187, 1995

Lawrence, S. and Giles, C. (1999), Accessibility of information on the web, Nature 400, p107-109, 1999

Leiner, B. (1994), Internet technology, Communications of the ACM 37, 8 (Aug. 1994), 32-339

McManus, B. (1997), Compensatory actions for time delays, Time and the Web, Staffordshire University, 19th June 1997
http://www.soc.staffs.ac.uk/seminars/web97/papers/barbara.html

Van Mannen, J. (1995), An end to innocence: the ethnography of ethnography, in representation in ethnography, eds Van Maanen, SAGE publications, pp 79-111, 1995

Miyake, N. (1986), Constructive Interaction, Cognitive Science 10, 1986

Monk, A.F. and Watts, L.A. (1999), Telemedical consultation in primary care: a case study in CSCW design, In M.A. Sasse & C. Johnson (eds.) Human-computer interaction - INTERACT'99, Edinburgh, UK. (pp. 367 - 374): IOS Press, 1999

Monk, A.F. and Watts, L.A. (2000). 'Peripheral participation in video-mediated communication, International Journal of Human Computer Studies, 52: 775-960, 2000

Moore, A. and Murray, B. H. (2000), Sizing the Internet, A Cyveillance Study, July 2000, http://www.cyveillance.com/

Mukherjea and Foley (1995), Visualizing the World-Wide Web with the Navigational View Builder, Computer Networks and ISDN System, Special Issue on the Third International Conference on the World-Wide Web '95, April 1995, Darmstadt, Germany.
http://www.cc.gatech.edu/gvu/people/Phd/sougata/www/mukh.html

Munzer, T and Burchard, P (1995), Visualizing the Structure of the World Wide Web in 3D Hyperbolic Space, Proceedings of the First Annual Symposium on the VRML Modelling Language, San Diego, December 14-15, 1995

Nelson, T.H. (1965), A file structure for the complex, the changing, and the indeterminate, Proceedings of the ACM National Conference, 84-100, 1965

Nelson, T.H. (1981), Literary Machines, Project Xanadu, 8400 Fredericksburgh, 1981

Nelson, T.H. (1995), The Heart of Connection: Hypermedia Unified by Transclusion, Communications of the ACM, 1995, 38(8): p.31-33, 1995

Newman, W. and Lamming, M. (1995), Interactive system Design, Addison-Wesley, 1995

Nielsen, J. (1990a), Hypertext and Hypermedia, Academic Press, 1990

Nielsen, J. (1990b), The art of Navigating through hypertext, Communications of the ACM, 1990, 33(3): p.296-310

Nielsen, J. (1994), Usability Engineering, Morgan Kaufmann Publishers, 1994
http://www.useit.com/papers/responsetime.html

Nielsen, J. (1995), Multimedia and Hypertext : the internet and beyond, Academic Press Ltd, 1995

Nielsen, J. (1998a), Using Link Titles to Help Users Predict Where They are Going, Alertbox for January 11, 1998, http://www.useit.com/alertbox/980111.html

Nielsen, J. (1998b), Fighting Linkrot, Alertbox for June 14, 1998,
http://www.useit.com/alertbox/980614.html

Nielsen, J. (1999), URL as UI, Alertbox, March 12, 1999,
http://www.useit.com/alertbox/990321.html

Norman, D. (1983), Some observations on mental models, In Mental Models, Ed. Gentner, D and Stevens, A, pp. 7-14, Lawrence Erlbaum Associates, 1983

Norman, D. (1988), The Psychology of Everyday Things. BasicBooks, 1988

Norman, D. (1995), Designing the Future, Scientific America, pp.159-160 September 1995

Norman D. (1999), Affordance, conventions, and design, Interactions 6,3, Pages 38-43, May 1999

O'Donnell, P, and Draper, S.W. (1995), How machine delays change user strategies, HCI'95 short paper, 1995

Parker, H. (1997), Temporal usability and disturbance management in interaction, Time and the Web, Staffordshire University, 19th June 1997, http://www.soc.staffs.ac.uk/seminars/web97/papers/helen.html

Petska-Juliussen, K. and Juliussen, E. (2000), Computer Industry Almanac, 8th Edition, ISBN 0-942107-07-1, 2000

Planas, M.A., and Treurniet, W.C. (1988), The effects of feedback during delays in simulated teletext reception, Behaviour and Information Technology, vol 7, no 2 ,183-191, 1988

Procter, R. and McKinlay, A. (1997), Social Affordances and Implicit Ratings for Social Filtering on the Web. In Proceedings of the fifth DELOS Workshop on Filtering and Collaborative Filtering, Budapest, Hungary, November 1997

Procter, R. et al. (1999), SELECT: Social and Collaborative Filtering of Web Documents and News. In Kobsa, A. and Stephanidis, C. (Eds.) Proceedings of the Fifth ERCIM Workshop on User Interfaces for All, Dagstuhl, November 28th-December 1st, 1999, p. 23-37

Ramsey, J., Barabasi, A. and Preece, J. (1998), A psychological investigation of long retrieval times on the world wide web, Interacting with Computers, 10:77-86, 1998

Randall, D., Hughes, J. and Shapiro, D. (1994), Steps towards a partnership: ethnography and system design, in: "Requirements Engineering: Social and Technical Issues," M. Jirotka and J. Goguen (eds.), Academic Press, London.

Ritchie, I. (1999), Wiring the World, Edinburgh University Informatics Colloquium 10th November 1999, http://www.coppertop.co.uk/present.html

Sears, A., Jacko, J. and Borella, A. (1997), Internet delay effects: How users perceive quality, organization, and ease of use of information, In Proceedings of CHI'97, pages 353-354, 1997

Sharrock, W. and Button, G. (1997), Engineering investigations: practical sociological reasoning in the work of engineers, in: G.C. Bowker, S.L. Star (eds.) Social Science, Technical Systems and Cooperative Work: Beyond the Great Divide. Hillsdale, N.J.: Lawrence Erlbaum: 79-104, 1997

Somerville, I. (1992), Software Engineering, Addison-Wesley, fourth edition, 1992

Stanyer, D. (1996), Requirements Engineering: Browsing in a Museum Environment, MSc dissertation, Oxford University, 1996

Stanyer, D. and Procter, R (1998), Link Lens: an enhanced link user interface for Web browser. In Stephanidis, C. and Waern, A. (eds.), proceedings of the 4th ERCIM Workshop on user interfaces for all, Stockholm, October, 1998, pp 34-47

Stanyer, D. and Procter, R (1999), Improving Web usability with the link lens, In Mendelzon, A et al. (Eds), Journal of Computer Networks and ISDN Systems, Vol. 31, Proceedings of the Eighth International WWW Conference, May 1999

Stone, M., Fishkin, K., Bier, E., Buxton, W. and Baudel, T. (1994), A Taxonomy of See-Through Tools. In Proceedings of CHI'94, ACM Press, 1997, pp 358-364

Stone, M., Fishkin, K. and Bier, E. (1994), The Movable Filter as a User Interface Tool. In Proceedings of CHI'94, ACM Press, 1997, pp 306-312

Suchman, L. (1987), Plans and situated actions: the problem of human machine communication, Cambridge, ISBN 0-521-33739-9, 1987

Sullivan, T. (1999), All things web – how much is too much? May 1999, http://www.pantos.org/atw/35654.html

Tauber, M. (1990), On Mental Models and the User Interface, Human-Computer Interaction, Ed Preece, J. and Keller, L., Prentice Hall, Chapter 15, pp.308-324, 1990

Tauscher, L. and Greenberg, S. (1997), How people revisit Web pages: Empirical findings and implications for the design of history systems. Int. J of Human-Computer Studies, Special issue on World Wide Web Usability 47.(1). pp 97-138

Thomas, S.A. (1996), IPng and the TCP/IP Protocols: Implementing the Next Generation Internet, Wiley, 1996

Treurniet, W.C., Hearty, P.J., and Planas, M.A. (1985), Viewers' responses to delays in simulated teletext reception. Behaviour and Information Technology, 4, 177, 1985

Wills, G.J. (1999), NicheWorks - Interactive Visualization of Very Large Graphs, Journal of Computational and Graphical Statistics, 8(2), pp. 190-212, June 1999, http://www.bell-labs.com/user/gwills/NICHEguide/nichepaper.html

Wolcott, H. F. (1995), Making a study "more ethnographic", in Representation in Ethnography, eds Van Maanen, SAGE publications, p.1-35, 1995

Woodruff, A., Aoki, P., Brewer, E., Gauthier, P. and Rowe, L. (1996), An Investigation of documents from the World Wide Web, The World Wide Web Journal, 1(3), 1996.

Wright, P and Monk, A. (1989), Evaluation for Design, In People and Computers V, eds Sutcliffe, V. and Macaulay, L. Cambridge, pp. 345-358, 1989

Wright, P and Monk, A. (1991), The use of think-aloud evaluation methods in design. SIGCHI Bulletin 23, 1 pp 55-57, Jan 1991

Wright, R. (1997), The man who invented the web, Time Magazine, May 1997, p.52-55

Yin, R. (1993), Applications of case study research, Beverly Hills, CA: Sage Publishing, 1993

Zellweger, P.T., Chang, B.W. and Mackinlay, J.D. (1998), Fluid links for informed and incremental link transitions, in Proceedings of the ACM Hypertext '98, 1998, pp 50-57

W3C (1997), HTML 4.0 Specification, W3C REC-html40-971218

RFC 1296 (1992), Internet Growth (1981-1991), January 1992

RFC 1480 (1993), The US Domain, June 1993

RFC 2045 (1996), Multipurpose Internet Mail Extensions: Format of Internet Message Bodies, November 1996

RFC 2046 (1996), Multipurpose Internet Mail Extensions: Media Types, November 1996

RFC 2047 (1996), Multipurpose Internet Mail Extensions: Message Header Extensions for Non-ASCII Text, November 1996

RFC 2048 (1996), Multipurpose Internet Mail Extensions: Registration Procedures, November 1996

RFC 2049 (1996), Multipurpose Internet Mail Extensions: Conformance Criteria and Examples, November 1996

RFC 2068 (1997), Hypertext Transfer Protocol—HTTP/1.1, January 1997

IE4 White Paper (1997), Microsoft Internet Explorer 4 white paper, June 1997 `http://www.microsoft.com/TechNet/IE/technote/IE4WP.asp`

Dublin Core Meta Data Initiative (2000), `http://purl.oclc.org/dc/`

GVU (2000), GVU's WWW user surveys, GVU Centre, College of Computing

Georgia Institute of Technology, Atlanta, http://www.cc.gatech.edu/user_surveys/

Reprinted from

COMPUTER NETWORKS

The International Journal of Computer and Telecommunications Networking

Computer Networks 31 (1999) 1533–1544

Improving Web usability with the link lens

Dominic Stanyer *,[1], Rob Procter [1]

Institute for Communicating and Collaborating Systems, Division of Informatics, University of Edinburgh, Edinburgh, Scotland, UK

ELSEVIER

Computer Networks 31 (1999) 1533–1544

COMPUTER NETWORKS

Improving Web usability with the link lens

Dominic Stanyer [*,1], Rob Procter [1]

Institute for Communicating and Collaborating Systems, Division of Informatics, University of Edinburgh, Edinburgh, Scotland, UK

Abstract

A number of factors may influence Web users' choice of which links to follow. These include assumptions about document quality and anticipated retrieval times. The present generation of World Wide Web browsers, however, provide only minimal support to assist users in making informed decisions. Web browser 'link user interfaces' typically only display a document's Universal Resource Locator (URL), whilst a simple binary colour change in the URL's anchor is used to indicate its activation history. The question then is, how do users deal with the problem of having to make such decisions when the information at hand is insufficient? We have been conducting an investigation of how users make link selections.The results show users often are forced to fall back on heuristics and improvising strategies drawn from past experience. Based upon these results, we present a prototype of the 'link lens', an enhanced link user interface designed to make such decisions easier and more productive for all users and help less experienced ones gain a better understanding of Web behaviour.

Keywords: Usability; User interfaces; Quality of service; Download delays

1. Introduction

The present generation of World Wide Web browsers provide users with too little information to help them decide whether to choose a link. Whilst there may be a number of factors that may influence this decision, the only information that the user can rely on being provided by a browser user interface is the document's URL, and an indication of its activation history. Sometimes, the user may be able to determine from contextual information whether to activate it or not (see Fig. 1). For example, by choosing an appropriate anchor, the author of the document in which the link is embedded can help the user to determine quality and/or relevance [19]; similarly, a description of the document's size and type may help the user to assess how long it will take to download, and whether he or she will be able to view it locally. Since this kind of information is entirely discretionary, however, authors often do not go to the trouble of including it.

Where such explicit information is lacking, users may still be able to infer useful information about the document and its download overheads. For example, the URL may contain implicit information such as the document's type and the physical location of its server. Such resourcefulness suggests that users would benefit if this kind of information were made more explicit, perhaps through the agency of an enhanced representation of the link. The selection and activation of a link is not the end of the user's dilemma, however. Download times are often long and unpredictable and may leave users uncertain

* Corresponding author.

[1] E-mail: {dsjs,rnp}@dcs.ed.ac.uk

1389-1286/99/$ – see front matter
PII: S1389-1286(99)00034-1

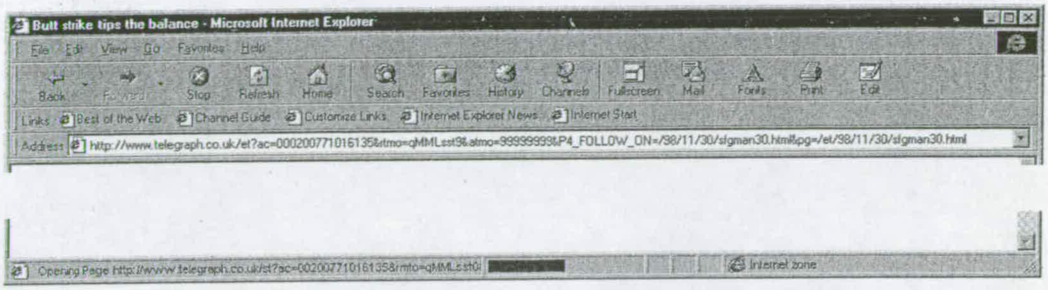


Fig. 1. A typical browser's 'link user interface'.

about when — or if — it will be completed. Yet, browsers (see Fig. 1) generally provide only limited feedback of download progress. If anything can be read into the scale and placement of this feedback, it is that it is unimportant. Yet numerous studies have shown that users of interactive systems are very sensitive to factors such as delays.

We have been conducting an investigation of how users make link selections and of their behaviour subsequent to link activation. To date, this has included re-enactment protocols of videoed Web sessions and interviews with Web users where they were asked to interpret a series of URLs. The results show how users fall back on heuristics and improvising strategies drawn from past experience. One important issue that emerges is that users' heuristics are often flawed simply because the picture of Web behaviour that they get from the link user interface is too abstract. Based upon these results, we discuss specific ways in which links might be enhanced to improve users' understanding of Web behaviour and provide more reliable information about its contents. The concept of enhanced or 'rich' links has been the subject of some discussion in the hypermedia literature (see e.g., [1]). Our objective here is to investigate what these might entail from a user perspective.

We begin by discussing some critical determinants of World Wide Web usability. We then present a model of link selection decision-making behaviour and relate this to our empirical observations of users. We then review relevant usability principles and finally, we propose ways in which users' decisions could be better supported through the 'link lens', a prototype of an enhanced link based upon Magic Lens [2] concepts of movable visual filters [2,27].

[2] Magic Lens is a trade mark of the Xerox Corporation.

2. Usability and the World Wide Web

The Web raises many usability issues [9]. In this paper, however, we focus on two specific ones which are fundamental to understanding how to enhance the link. The first of these is content: how users determine that the document referenced by a link meets their requirements. The second is the temporal behaviour of the link: how long the selected document will take to download. These two factors together: quality of the document and quality of service are key determinants of users' selection and use of links. These factors interact, however, so their combined influence cannot be determined a priori. For example, a user may be prepared to tolerate slow download where the document quality is known to be high. Conversely, if time is pressing, then a lower quality — but more speedily downloaded — document may be more acceptable.

Neither document quality, nor its corollary, relevance, can be properly assessed except in the context of use. The only certain way that a user has of determining these properties is to download the document. However, following this strategy blindly may require the user to invest a lot of time and effort. This is simply impractical when the user has many documents from which to choose, such as when picking from a list of documents generated by a search engine. One solution is to make use of whatever contextual information is available. For example, the choice of anchor may provide a meaningful summary of the content of the document which it references. The wider context in which the anchor appears may also provide further information.

The problems of relying on contextual information partly explain the current wide interest in developing a standardised document markup language such as the World Wide Web Consortium's Extensible Markup Language (XML) [29] so that content providers can add meta-data — information about information — to Web documents. One goal of adding meta-data is to facilitate the automatic processing of documents by Web applications such as search engines. We argue, however, that meta-data may also be an important resource for the design of an enhanced link. The first practical objective of our work is therefore to investigate how meta-data might be presented through an enhanced link.

As any Web user soon learns, the simple underlying abstraction of a seamless, distributed information space is rarely sustained in practice. Instead of an instant and deterministic response to link activation, the user typically must cope with unpredictable download times arising from variable latencies and network bandwidths. System response times in general are recognised as a major factor in determining the usability of an interactive system [4–6,14,21] and investigating how Web delays affect user behaviour is an area of growing interest [10,17,15,25].

Studies have looked at how users' strategies change to cope with extreme delays or system failures [7,17]. The unpredictable nature of Web download times means that users are unable to apply consistent coping strategies which leads to frustration and error. Attempts have been made to formalise this behaviour with an aim to helping developers design for download latencies [13,15]. Other studies suggest that users' perceptions of document quality on the one hand, and their tolerance of download times on the other, are not independent. In particular, they indicate that download times may influence users' perception of document quality and relevance. For example, users' ratings of *how interesting* a document is have been shown to decrease for longer download times [23,25]. Conversely, users' tolerance of download delays is influenced by their expectations of document quality [15]

Despite the well-known detrimental effects of unpredictable delays on users, browser user interfaces all but ignore the problem. Whilst abstractions are useful tools for protecting users from unnecessary details of system behaviour, rigid adherence to a single abstraction may deny them the information they need to act effectively [8,16]. The link user interface, we argue, is a good example of this problem. The account it offers of the behaviour of the underlying system is too abstract and so fails to provide users with the resources for informed decisions and effective improvisation when potential problems are detected. Our second practical objective, therefore, is to investigate how to provide all users with a more informative account of download behaviour [20] which is helpful both for decision-making and for learning.

3. Terminology

A distributed information space such as the Web consists of documents connected by links. A document may represent different types of content, it may be referenced by links, and also contain links that reference other documents within the information space. A document can therefore be referenced by a number of links each situated in a different document.

A link is embedded within a document and attached to part of latter's content by its *anchor* (see Fig. 2) [18]. With respect to a particular link, we will define the object in which the link is embedded as

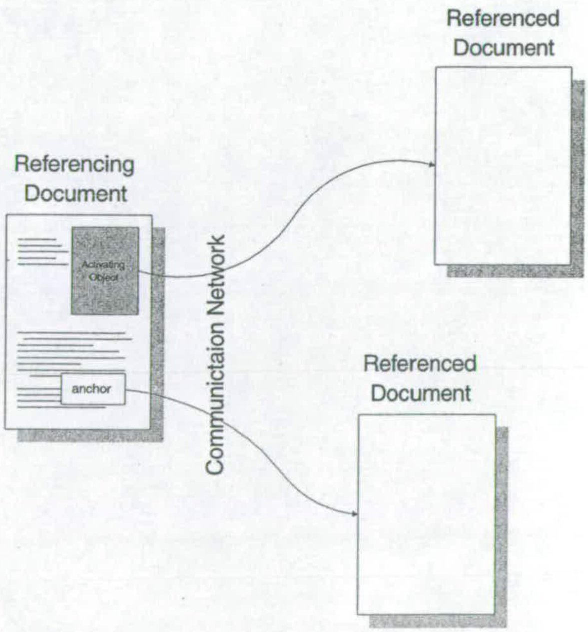


Fig. 2. The link.

the *referencing document*. Similarly, the object that the link points to will be defined as the *referenced document*.

As well as being identified with its anchor, referencing, and referenced documents, a link also has behavioural properties which become evident when it is activated and the referenced document is downloaded from the server to the user's client. This behaviour reflects the communication process which is implemented on the underlying network between the two locations.

4. Understanding Web user behaviour

A series of observational studies of Web user behaviour were undertaken in order to learn more about how users made link activation decisions. Fourteen subjects were recruited with varying degrees of Web experience and expertise in computer use, ranging from novices to experts. Data about subject browsing behaviour was automatically logged by using a proxy server connected to the Web client. This parsed the HTML pages that the subject downloaded and recorded the URL, number of images, number of links, number of applets and the download time from the proxy request to the proxy receiving the data. Each browsing session was also recorded using video and audio.

At the end of the session, all the collected data — video, audio and proxy logs — was reviewed and a detailed transcript of the session was produced. A re-enactment protocol was then employed to elicit from subjects the reasoning behind their actions [28]. All the utterances of the subject and the investigator were recorded on audio cassette and later transcribed.

Analysis of this data leads us to propose a three phase model of user link activation decision-making behaviour:

Assessment of value. The user must decide prior to activating a link whether the referenced document contains information that is useful with respect to the user's overall objective. Reaching a positive assessment will lead to the link being activated and the download of the referenced document will commence; a negative assessment will result in the link being ignored. In addition to the assessment of the referenced document's content, the user also makes an assessment of the likely download quality of service (QoS). The potential of a long wait (e.g., because of server or network congestion, low network bandwidth) for the referenced document may persuade a user to ignore the link, or at least ignore the link for the present until a more favourable time. Our subjects often tried to infer by the URL information about the probable length of the wait. Here, heuristics and rules of thumb prompted by previous experience often come into play. For example:

> "This page is on a University server and that means it should be quick because they have fastish servers." (Subject C)

The assessment of value process that the user undertakes prior to activating a link can therefore, be split into two parts:

(1) an assessment of content, and
(2) an assessment of download QoS, particularly the anticipated delay in downloading the document.

Download. Once the link has been activated, the document download begins. During this phase the user is exposed to the underlying behaviour of the communication network. As we noted above, users of interactive systems are often very sensitive to delays, especially where these are of an unpredictable extent. The problem takes on even greater significance for the Web, however, because browser user interfaces allow users to interrupt and abandon the download at any time. In our studies, observations of users choosing this option were very common.

Johnson has employed micro-economic cost-utility models and the notion of marginal utility to analyse Web user responses to download delays [15]. He demonstrated that users' tolerance of delays could be manipulated by increasing their expectations of document quality. Johnson's findings are evidence for the close relationship between the assessment of value phase and the download phase; by increasing the information available to the user before they activate the link, the user is not only better informed about its nominal value, but is also better able to make an informed decision about whether the document is worth waiting for.

Some of our subjects revealed interesting beliefs about the extent to which their actions could influence the download time. For example, several

```
protocol ``://'' hostname ``/'' [pathname] [filename ``.'' file_type]
```

Fig. 3. The constituent parts of a URL.

subjects articulated the view that slow downloads could be speeded up by stopping and restarting them. When asked to explain this, one user drew upon a telephone analogy:

"...[it] is like a crackerly telephone line, if you ring again the quality is often better. " (Subject E)

Others provided more technically sophisticated rationales for their behaviour:

"... when you stop and reload, it is possible to get a better route for the data, so it may be quicker." (Subject G)

"... there is only a certain number of connections that a server can make, by reloading a page a new connection may have become free due to a time out or other users disconnecting." (Subject D)

Evaluation. The final phase allows a user to evaluate the outcome of activating the link. If the activation was successful, the referenced document can be perused and its content can be compared with the expectations generated by the assessment of value phase. Similarly, the download performance can be evaluated. The evaluation process may stimulate the formation of heuristics that will serve to make sense of the experience, and which may then come into play in future cycles of interaction. A flavour of these heuristics is evident in explanations given by our subjects such as:

"... all sites in Australia are slow." (Subject A)

"... it is night time in America so all information will be transferred very fast." (Subject I)

It was evident from subjects' comments that the URL was often an important element in the assess-download-evaluate cycle. The URL could be used to predict the content of the referenced document, the location of the Web site and the download time in the *Assessment of value* phase. In the *Evaluation* phase, in retrospectively comparing the outcome with their initial expectations, users would try to (re)-interpret the URL in a way that was consistent with the outcome, the spatial location of the site and the time taken for data transfer.

These observations led to the design of a second study to investigate what information users can extract from the URL and how this may figure in the link activation decision. Eighteen subjects with varying degrees of Web experience and expertise in computer use were recruited. Each one was shown a list of URLs and asked to predict what he or she thought the referenced document was, and to predict its download behaviour. Subjects were then shown the actual referenced document and asked to assess the accuracy of their prediction.

The results indicated that more experienced subjects were able to deconstruct a URL into its constituent parts (see Fig. 3). Generally, a 'full' prediction would include inferences about geographical location from the hostname, about content from the pathname and filename, and about download behaviour from the `file_type` and hostname parts (i.e., from document type and location). In contrast, less experienced subjects were generally unable to distinguish between the constituent parts of the URL. Inferences about location, content and download behaviour were therefore based upon a decontextualised recognition of strings within any part of the URL.

In Fig. 4 we show extracts from transcripts of session protocols. Interviewer's and subject's comments are distinguished by (I:) and (S:) respectively. The first extract shows the kind of inferences made by a typical novice subject. The next three extracts illustrate the kinds of inferences made by more experienced subjects.

Informed by the results of these investigations, we now turn to consider in more detail how the usability objectives defined earlier may be tackled. Our discussion will focus on two broad areas in turn: documents and links.

5. Documents

Despite the undoubted resourcefulness that users display for making a little information go along way, inferences based upon URLs are always likely to be an unreliable way of making judgements of document content. Instead, it is meta-data that is the

1. `http://cello.cs.uiuc.edu/cgibin/slamm/ip211`
S: "Must be a school or University or something."
I: "...you know where?"
S: "No, but it must be a university though, well maybe not, cello..."
I: "What do you think it is about?"
S: "Well it is some sort of music thing in a university."

2. `http://www.stfturist.se/engelska/index.htm`
S: "Umm looks like it is in Sweden or from a Swedish Web site."
I: "Is that from the dot se?"
S: "se yes."

3. `http://home.netscape.com/newsref/pr/newsrelease558.html`
S: "Another one that doesn't have www at the start, I guess its just the main Netscape page and newsref news information PR press release I guess, and newsrelease558 probably the five hundredth and fifty eighth press release they have had since they started cause there is no date information in the URL."

4. `http://www.lmnet.com/~bobh/shakeit.avi`
I: "Why is it going to be slow?"
S: "Ermm because dot com makes me think American company and also avi's are normally enormous."

Fig. 4. Extracts from transcripts of session protocols from the second study.

key element in improving users' capacity to make informed assessments of a document's quality and relevance [1]. Of course, the form this meta-data will take must be dependent upon the kind of document it describes. Whilst it is easy to specify that the meta-data for a scientific paper would include its authors, key words and abstract, that of many other document types is more difficult to anticipate. The issue that must be addressed then is how such meta-data may be effectively presented to the user. One possible approach is to use meta-data to provide some insights in to the document structure. A simple illustration of how this might be done is shown in Fig. 5. Here we see the document content summarised by:

(1) Author, title, keywords, abstract and section names appearing in the document. HTML already provides some support for document meta-data through the use of a number of pre-defined tags. These include author, title, and keywords. The new proposed document markup language XML goes further in that it allows the author to define new tags; in fact, every component of an XML document may be tagged in this way.
(2) The distribution of keywords through the document body. This can provide a better indication of relevance than a simple count [11].
(3) A thumbnail outline of the document. The concept of genre is one possible way in which broader notions of meta-data may be addressed. The use of genre in the print media, for example, enables people to make quick and easy distinctions between the vast variety of publications available. To help the readership distinguish between news genre, simple visual cues such as the tabloid and broad sheet format have been developed. The thumbnail is an attempt to provide similar visual cues for Web documents.
(4) The numbers and sizes of different embedded media types and the number of links. This also may provide the user with genre cues. For example, a document containing many links may appear to be a richer resource for browsing.

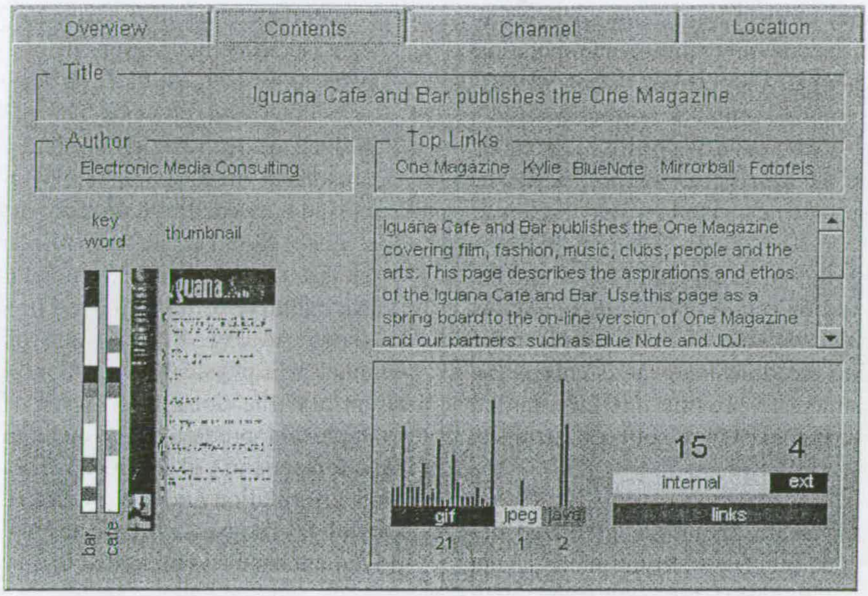


Fig. 5. An example of document meta-data: structure and content summaries.

6. The link

To provide the user with appropriate information about the link, it is necessary to identify an abstraction — or perhaps a layered set of abstractions —that can explain its behaviour [8]. Our chosen link abstraction has two elements. The first is the *channel*, a communication resource which connects two (or more) computers together, and allows information to flow between them in either direction. If we restrict our notion of channel to the overall connection between the browser client and the document server, we have a concept that is synonymous with the communication associated with a download. The other element of our link abstraction is the *site*, a resource for despatching documents provided by the server. A channel and a site can now be associated with a link, and their properties can be presented to the user as affordances for understanding link behaviour.

The key property of a link is its QoS. Definitions of link QoS for multimedia applications cite several contributing factors [10]. For the Web user, however, the key QoS criteria is time. In simple terms, what the user wants to know before initiating a download is the delay before the download is completed. For the link, this can be defined as a function of site and channel latencies and channel bandwidth. Latencies reflect server and network congestion causing competition for these resources.

A full link time affordance, however, must also provide the user with information about the task in progress. The concept of time affordance is informed by users' subjective desire for closure, the knowledge that a system task has been successfully completed. Until this is evident, users will tend to monitor the task, assessing its progress and preparing contingency plans should they determine that it has failed. Lack of adequate time affordances may mislead users, for example, causing them to terminate the task prematurely.

According to Conn, a complete time affordance should meet eight requirements, representing information about distinct task properties [4]. We confine ourselves here to those where the link user interface is particularly deficient, and discuss ways in which an enhanced link might satisfy them. Some of the solutions we propose raise functional requirements which are addressed later.

Acceptance:

Acknowledgement that the task has been correctly specified and is executable. Arguably, this is essentially a null requirement for browsers if the task 'specification' is defined syntactically as simply the act of pointing at any visible anchor. However, in so-called direct manipulation interfaces, syntactically acceptable action specifications have an implied semantic correctness [24]. For example, a document's visibility is proof of its existence. In the case of an anchor, however, this relationship between visibility and existence may not hold.

Improved satisfying of acceptance requirements could be achieved by server and document status information being available when the user is in the act of specifying the download task. This is the most timely moment to provide this information and resolves the inconsistency between visibility and existence.

Scope:

Measures of task scale, including predicted delay before completion. This requirement is not satisfied by the information provided through the link user interface. It may be partially satisfied by contextual information provided by the author of the referencing document, such as the referenced document's size.

Download scope is a function of the referenced document size and the link QoS. Satisfying scope requirements implies that these factors can be determined in advance of the download being initiated, In practice, however, only the document size can be accurately determined in this way — given suitable functional extensions.

Link QoS figures should take into account the effects of caching. Caches were introduced to exploit locality in users' document downloading behaviour. For example, there may be a cache on the user's (client) machine and a cache on a local proxy serving a number of users. If a document has been previously retrieved it could be available on these caches. A cached document will almost always download more quickly than the copy on its server.

Progress:

Clear feedback as execution of the task proceeds, including its rate of progress. This requirement we consider is only partially satisfied. Although a progress indicator is maintained during the download, it is often difficult to relate the information it provides of the task state to its subsequent behaviour. One reason for this is that a document which may seem to the user to be a single entity may actually be composed of many embedded parts (e.g., images) each of which will trigger its own download process. The order of these embedded downloads is non-deterministic. In most browsers [3] the progress display area is time-multiplexed between these separate downloads, making it impossible for the user to maintain a coherent picture of the overall download progress [17].

An example of an improved link time affordance is shown in Fig. 6. The map display shows the channel as a physical connection between the user's client and the document server. Its colour is used to represent predicted channel QoS (ranging from blue for poor to red for good). The graph shows site QoS information and the predicted download scope. Download progress is indicated by animation of the document media content display in Fig. 5.

7. Functional implications

The possible improvements outlined above are more than cosmetic and raise additional functional requirements of the underlying system.

If the link is to be able to provide the user with meta-data, this information must either be provided by the referencing document's author, or be downloaded from the referenced document's server. The former approach has been applied in some specialised hypermedia publications [1]. However, our goal is to make meta-data available for a wider range of documents, and this is only achievable by following the download approach. To avoid this creating its own delays, two requirements must be met. First, the meta-data must be accessible to the client without downloading the whole document. Second, some kind of meta-data pre-loading or transclusion strategy is required so that it is available in a timely way to the user.

The HTTP protocol already supports client access of document parts through the HTTP-EQUIV directive. This allows authors to specify document tags which can be downloaded separately from the rest of the document and so HTTP already has the

[3] Sun's Hotjava is an exception in providing progress indicators for embedded downloads.

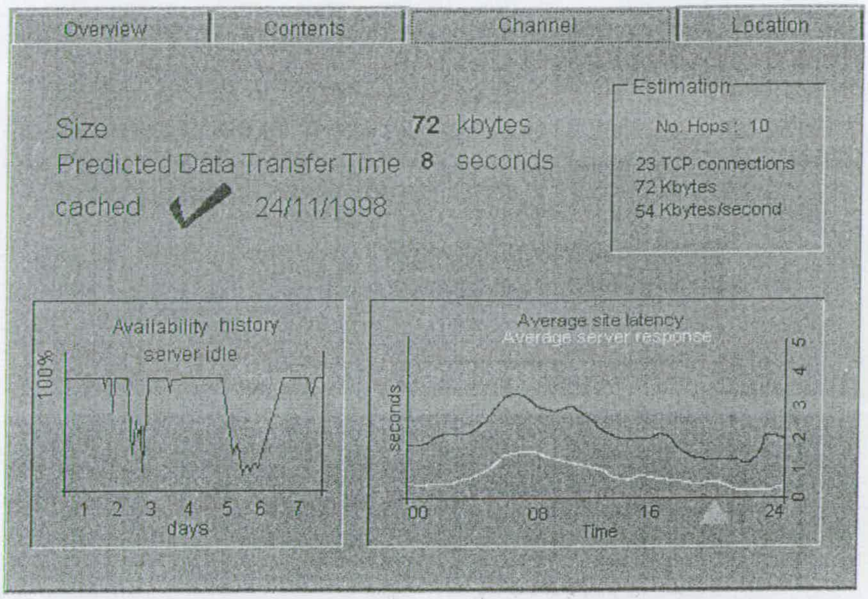


Fig. 6. An example of link QoS data: channel and site QoS summaries.

necessary hooks to meet the first requirement. The second requirement can be met by appropriate client side enhancements. There are a number of possible meta-data pre-load strategies. For example:

- All link meta-data is pre-loaded when the referencing document is opened. This is the least efficient strategy. Its drawbacks include the fact that the number of links may be large. The following three strategies adopt a progressively more conservative approach.
- Pre-loading is limited to those links that are within the currently visible area of the referencing document.
- Pre-loading is triggered when the user moves the link lens tool within a specified distance of the link anchor.
- Pre-loading is triggered when the user passes the link lens tool over the link anchor.

All the above pre-load strategies are assumed to operate automatically. It would be sensible, however, to allow user control over pre-loading, e.g., setting it on or off, and determining thresholds where appropriate. Yet another approach would be to give control of pre-loading to the referenced document's site. For example:

- Pre-loading is determined by the accumulated history of link activation by other users. This strategy would provide users with a form of collaborative filtering mechanism.

The accurate determination of download scope relies on knowing channel and site QoS. Obtaining real-time channel QoS data is problematic given current communication protocol limitations. The Hypertext Transfer Protocol (HTTP) is layered on top of the Transmission Control Protocol (TCP), which itself is layered on top of the Internet Protocol (IP). The latter does not support any QoS functions and so TCP cannot make any assumptions about factors such as latency or bandwidth. The only channel QoS related notion supported by IP is distance as measured by the number of 'hops' between the sender and receiver [3]. Support for QoS is planned as part of an enhanced HTTP protocol [10], but in the meantime we must look to other solutions.

There are now several companies offering Web server QoS data on-line. Such data typically includes the number of hits per day, the average latency, and the numbers of broken links. However, we have decided to use a different approach which, though less accurate, does not assume the availability of site QoS servers. The link lens utilises the performance of the meta-data pre-load cycle to calculate channel and site QoS for each link. In the limit case where link pre-loading has not been completed before the

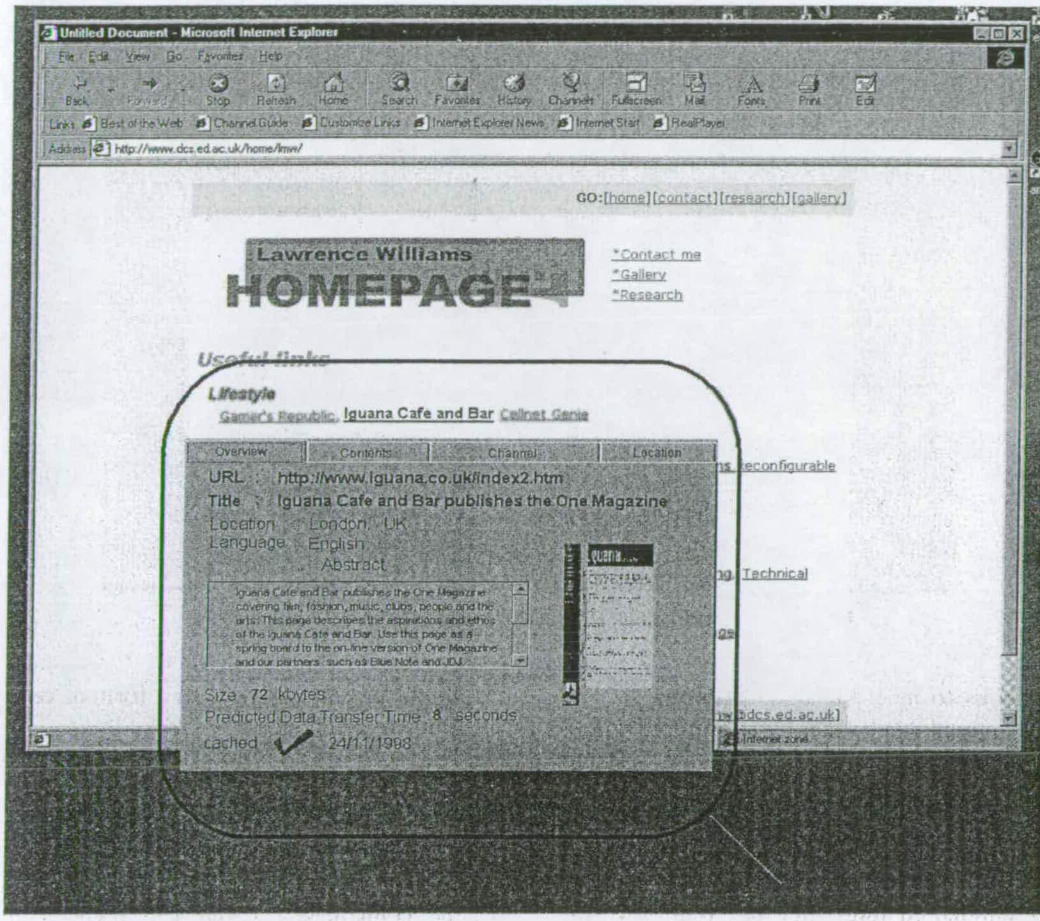


Fig. 7. The prototype link lens.

user wants to have access to it, the link can simply be assigned a low QoS, leaving the user to draw his or her own conclusions.

8. A prototype link lens

The concept upon which the link lens is based is the so-called "Magic Lens" or movable filter [2,27]. Magic lens filters combine an arbitrarily shaped visual field with an operator that changes the presentation of objects within the view field. In the link lens, the view field is bound to the cursor and is sensitive to any anchors that it is passed over. When applied in this way, its two operators, document and link, transform the anchor into views of document meta-data and link QoS. The prototype link lens is shown in Fig. 7. The design was evolved from paper-based sketches of possible lens features. These were presented to six subjects who rated them according to perceived usefulness.

Their overall opinion was that the link lens would affect their browsing strategies. With the exception of the document abstract, there was no consensus about which parts of the link lens would be the most useful, but all subjects expressed a desire to be able to choose a subset of meta-data for viewing at one time. Depending on the task being undertaken, the subjects felt that some parts of the link lens could be beneficial to the usability of the system.

We are currently carrying out further testing of several versions of the link lens with groups of Web users. The objectives include evaluating the value to users of the various different kinds of link information. We are also investigating different ways of presenting the information, including enabling users to customise or adapt the link lens to match their needs, and more economical encodings based upon the appearance of the link in the referencing document. A particular emphasis of this work will be to explore how degrees of user experience affect their ability to make use of link information.

9. Conclusions and further work

The prototype link lens presents document meta-data and time affordances in the form of link QoS information. Its goal is to provide Web users with timely and relevant information prior to their activation of links and during subsequent downloads. Underlying our work is the belief that by making a more accurate and reliable picture of Web behaviour available, users will be able to formulate more appropriate explanations of its behaviour and make more informed decisions about its use.

Future work will focus on evaluation and exploration of improved techniques for link information presentation, including the value of layered abstractions of Web behaviour. The question of how much users need to know about Web behaviour is a priority for further investigation. The evidence from our user studies confirms that users will generate explanations of Web behaviour and will act upon them. Their reasoning and the analogies they draw upon, however, range from the simple to the sophisticated. So, it is probable that novice users will need different kinds of abstractions from those for more experienced users. Our aim in developing the link lens is to provide support for all kinds of users. An important part of our strategy for achieving this is to recognise that in the hands of a less experienced user, the link lens may serve as a tool for learning about how the Web behaves.

As part of a broader programme of work for improving the value of the Web, we have recently begun to investigate the use of collaborative filtering — document usage and user ratings — as additional document quality and relevance information [12, 22,26]. As HTTP protocol enhancements become available, we will also examine how they may be applied in this context.

Finally, any mechanism for improving Web users' capacity to find relevant information must ultimately rely upon the willingness of content providers to recognise users' needs. In particular, the need for content providers to include meta-data is paramount. The link lens is designed so that the user will find it easy to distinguish between content providers who are conscientious in this regard and those who are not. In this way, we believe that putting tools like the link lens in the hands of users can provide an incentive for content providers to formulate and adhere to 'good authoring practice'.

References

[1] M. Bieber, F. Vitali, H. Ashman, V. Balasubramanian and H. Oinas-Kukkonen, Fourth generation hypermedia: some missing links for the World Wide Web, International Journal of Human-Computer Studies 47 (1997) 31–65.

[2] E. Bier, M. Stone, K. Fiskin, W. Buxton and T. Baudel, A taxonomy of see-through tools, in: Proc. CHI '94, ACM Press, 1994, pp. 358–364.

[3] G. Brebner, Computers in Communication, McGraw-Hill, New York, 1997.

[4] A. Conn, Time affordances: the time factor in diagnostic usability heuristics, in: Proc. CHI '95, ACM Press, 1995, pp. 186–913.

[5] A. Dix, The myth of the infinitely fast machine, in: D. Diaper and R. Winder (Eds.), People and Computers III, Cambridge University Press, Cambridge, 1987, pp. 215–228.

[6] A. Dix, Seven years on, the myth continues, Research Report RR9405, University of Huddersfield, 1994, http://www.soc.staffs.ac.uk/~cmtajd/papers/myth95/7-years-on.html

[7] P. O'Donnell and S. Draper, How machine delays change user strategies, in: M. Kirby, A. Dix and J. Finlay (Eds.), People and Computers X: HCI'95, Cambridge University Press, Cambridge, 1995.

[8] P. Dourish, Accounting for system behaviour: representation, reflection and resourceful action, in: Proc. 3rd Decennial Conf. on Computers in Context, 1995, pp. 147–156.

[9] Graphics, visualisation and usability center, 8th WWW user survey, Georgia Tech, 1998, http://www.gvu.gatech.edu/user_surveys/survey-1997-10/graphs/use/Problems_Using_the_Web.html

[10] X. Guo and C. Pattinson, Quality of service requirements for multimedia communications, in: Workshop on Time and the Web, Staffordshire University, June, 1997.

[11] M. Hearst, TileBars: visualisation of term distribution information in full text information access, in: Proc. CHI '95, ACM Press, 1995, pp. 59–66.

[12] W. Hill, L. Stead, M. Rosenstein and G. Furnas, Recommending and evaluating choices in a virtual community of use, in: Proc. CHI '95, ACM Press, 1995, pp. 194–201.

[13] C. Johnson, Time and the Web: representing temporal properties of interaction with distributed systems, in: M. Kirby, A. Dix and J. Finlay (Eds.), People and Computers X: HCI '95, Cambridge University Press, Cambridge, 1995, pp. 39–50.
[14] C. Johnson, The challenge of time, in: P. Palanque and R. Bastide (Eds.), The Design, Specification and Verification of Interactive Systems, Springer, Berlin, 1995, pp. 345–357.
[15] C. Johnson, What's the Web worth? the impact of retrieval delays on the value of distributed information, Workshop on Time and the Web, Staffordshire University, June, 1997.
[16] D. Kieras and S. Bovair, The role of a mental model in learning to operate a device, in: J. Preece and L. Keller (Eds.), Human–Computer Interaction, Prentice-Hall, Englewood Cliffs, NJ, 1990, pp. 205–221.
[17] B. McManus, Compensatory actions for time delays, in: Workshop on Time and the Web, Staffordshire University, June, 1997.
[18] J. Nielsen, Hypertext and Hypermedia, Academic Press, New York, 1990,
[19] J. Nielsen, Using links to help users predict where they are going, Jakob Nielsen's Alertbox, January, 1998, http://www.useit.com/alertbox/980111.html
[20] H. Parker, Temporal usability and disturbance management in interaction, in: Workshop on Time and the Web, Staffordshire University, June, 1997.
[21] M. Planas and W. Treurniet, The effects of feedback during delays in simulated teletext reception, Behaviour and Information Technology 7 (2) (1988) 183–191.
[22] R. Procter and A. McKinlay, Social affordances and implicit ratings for social filtering on the Web, in: L. Kovacs et al. (Eds.), Proc. 5th DELOS Workshop on Filtering and Collaborative Filtering, Budapest, Hungary, November 10–12, 1997, ERCIM Press.
[23] J. Ramsay, A. Barbesi and J. Preece, A psychological investigation of long retrieval times on the World Wide Web, Interacting with Computers 10 (1998) 77–86.
[24] B. Schneidermann, Designing the User Interface, Addison-Wesley, Reading, MA, 1986.
[25] A. Sears, J. Jacko and M. Borella, Internet delay effects: how users perceive quality, organization, and ease of use of information, in: Proc. CHI '97, ACM Press, 1997, pp. 353–354.
[26] Select project, http://www.omega.it/research/select
[27] M. Stone, K. Fiskin and E. Bier, The movable filter as a user interface tool, in:Proc. of CHI'94, ACM Press, 1997, pp. 306–312.
[28] P. Wright and A. Monk, The use of think-aloud evaluation methods in design, SIGCHI Bulletin 23 (1) (1991) 55–57.
[29] Extensible Markup Language, World Wide Web Consortium, http://www.w3.org/XML/

Dominic S.J. Stanyer was born in London town, England and received a B.Eng. in software engineering from the University of Sheffield, and an M.Sc. degree in computation from the University of Oxford. He is presently a Ph.D. student in the Division of Informatics at the University of Edinburgh, where he is researching human interaction within distributed information spaces.

Rob Procter is a senior lecturer in the Division of Informatics, Edinburgh University. His research interests lie principally in the field of human factors, particularly computer-supported cooperative work (CSCW), and methodologies for the design and implementation of IT systems. His current research includes computer-aided mammography, collaborative environments for digital libraries, social and collaborative filtering techniques, and social learning processes in IT systems implementation.

Link Lens: An Enhanced Link User Interface for Web Broswers

Dominic Stanyer and Rob Procter
School of Computer Science
Division of Informatics
Edinburgh University
Scotland
Tel: +44 131 650 5165|5177
Fax: +44 131 667 7209
Email: dsjs|rnp@dcs.ed.ac.uk
URL: http://www.dcs.ed.ac.uk/home/dsjs|rnp

Abstract. A number of factors may influence Web users' choice of which links to follow. These include assumptions about document quality and anticipated retrieval times. The present generation of World Wide Web browsers, however, provide only minimal support to assist users in making informed decisions. Web browser 'link user interfaces' typically only display a document's Universal Resource Identifier (URL), whilst a simple binary colour change in the URL's anchor is used to indicate its activation history. The question then is, how do users deal with the problem of having to make such decisions when the information at hand is insufficient?

We have been conducting an investigation of how users make link selections. The results show users often are forced to fall back on heuristics and improvising strategies drawn from past experience. Based upon these results, we present a prototype of the 'link lens', an enhanced link designed to make such decisions easier and more productive for all users and help less experienced ones gain a better understanding of Web behaviour.

Introduction

The present generation of World Wide Web browsers provide users with too little information to help them decide whether to choose a link. Whilst there may be a number of factors that may influence this decision, the only information that the user can rely on being provided by a browser user interface is the document's URL, and an indication of its activation history. Sometimes, the user may be able to determine from contextual information whether to activate it or not (see Figure 1). For example, by choosing an appropriate anchor, the author of the document in which the link is embedded can help the user to determine quality and/or relevance [Nielsen 98]; similarly, a description of the document's size and type may help the user to assess how long it will take to download, and whether he or she will be able to view it locally. Since this kind of information is entirely discretionary, however, authors often do not go to the trouble of including it.

Where such explicit information is lacking, users may still be able to infer useful information about the document and its download overheads. For example, the URL may contain implicit information such as the document's type and the physical location of its server. Such resourcefulness suggests that users would benefit if this kind of information were made more explicit, perhaps through the agency of an enhanced link. The selection and activation of a link is not the end of the user's dilemma, however. Download times are often long and unpredictable and may leave users uncertain about when — or if — it will be completed. Yet, browsers (see Figure 1) generally provide only limited feedback of download progress; even its placement seems to suggest that it is an afterthought.

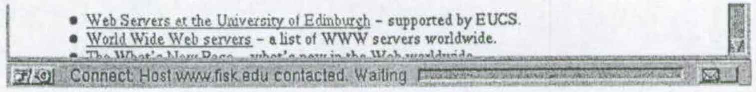


Figure 1: Netscape's 'link user interface'.

We have been conducting an investigation of how users make link selections and of their behaviour subsequent to link activation. The results show how users fall back on heuristics and improvising strategies drawn from past experience. One important issue that emerges is that users' heuristics are often flawed simply because the picture of Web behaviour that they get from the user interface is too abstract. Based upon these results, we discuss specific ways in which links might be enhanced to improve users' understanding of Web behaviour and provide more reliable information about its contents. The concept of enhanced or 'rich' links has been the subject of some discussion in the hypermedia literature (see e.g., [Bieber 97]). Our objective here is to investigate what these might entail from a user perspective.

We begin by discussing some critical determinants of World Wide Web usability. We then present a model of link selection decision-making behaviour and relate this to our empirical observations of users. We then review relevant usability principles and finally, we propose ways in which users' decisions could be better supported through the 'link lens', a prototype of an enhanced link based upon Magic Lens[1] concepts [Bier 94, Stone 94].

Usability and the World Wide Web

The Web raises many usability issues [GVU 98]. In this paper, however, we focus on two specific ones which are fundamental to understanding how to enhance the link. The first of these is content: how users determine that the document referenced by a link meets their requirements. The second is the temporal behaviour of the link: how long the selected document will take to download. These two factors together: quality of the document and quality of service are key determinants of users' selection and use of links. These factors interact, however, so their combined influence cannot be determined a priori. For example, a user may be prepared to tolerate slow download where the document quality is known to be high. Conversely, if time is pressing, then a lower quality — but more speedily downloaded — document may be more acceptable.

Neither document quality, nor its corollary, relevance, can be properly assessed except in the context of use. The only certain way that a user has of determining these properties

[1]Magic Lens is a trade mark of the Xerox Corporation.

is to download the document. However, following this strategy blindly may require the user to invest a lot of time and effort. This is simply impractical when the user has many documents from which to choose, such as when picking from a list of documents generated by a search engine. One solution is to make use of whatever contextual information is available. For example, the choice of anchor may provide a meaningful summary of the content of the document which it references. The wider context in which the anchor appears may also provide further information.

The problems of relying on contextual information partly explain the current wide interest in developing a standardised document markup language such as W3C's Extensible Markup Language (XML) [XML] so that content providers can add meta-data — information about information — to Web documents. One goal of adding meta-data is to facilitate the automatic processing of documents by Web applications such as search engines. We argue, however, that meta-data may also be an important resource for the design of an enhanced link. The first practical objective of our work is therefore to investigate how meta-data might be presented through an enhanced link.

As any Web user soon learns, the simple underlying abstraction of a seamless, distributed information space is rarely sustained in practice. Instead of an instant and deterministic response to link activation, the user typically must cope with unpredictable download times arising from variable latencies and network bandwidths. System response times in general are recognised as a major factor in determining the usability of an interactive system [Conn 95, Dix 87, Dix 94, Johnson 95b, Planas 88] and investigating how Web delays affect user behaviour is an area of growing interest [Guo 97, McManus 97, Johnson 97, Sears 97].

Studies have looked at how users' strategies change to cope with extreme delays or system failures [O'Donnell 95, McManus 97]. The unpredictable nature of Web download times means that users are unable to apply consistent coping strategies which leads to frustration and error. Attempts have been made to formalise this behaviour with an aim to helping developers design for download latencies [Johnson 95a, Johnson 97]. Other studies suggest that users' perceptions of document quality on the one hand, and their tolerance of download times on the other, are not independent. In particular, they indicate that download times may influence users' perception of document quality and relevance. For example, users' ratings of *how interesting* a document have been shown to decrease for longer download times [Ramsay 98, Sears 97]. Conversely, users' tolerance of download delays is influenced by their expectations of document quality [Johnson 97]

Despite the well-known detrimental effects of unpredictable delays on users, browser user interfaces all but ignore the problem. Whilst abstractions are useful tools for protecting users from unnecessary details of system behaviour, rigid adherence to a single abstraction may deny them the information they need to act effectively [Dourish 95, Kieras 90]. The link user interface, we argue, is a good example of this problem. The account it offers of the behaviour of the underlying system is too abstract and so fails to provide users with the resources for informed decisions and effective improvisation when potential problems are detected. Our second practical objective, therefore, is to investigate how to provide all users with a more informative account of download behaviour [Parker 97] which is helpful both for decision-making and for learning.

Terminology

A distributed information space such as the Web consists of documents connected by links. A document may represent different types of content, it may be referenced by links, and also contain links that reference other documents within the information space. A document can therefore be referenced by a number of links each situated in a different document.

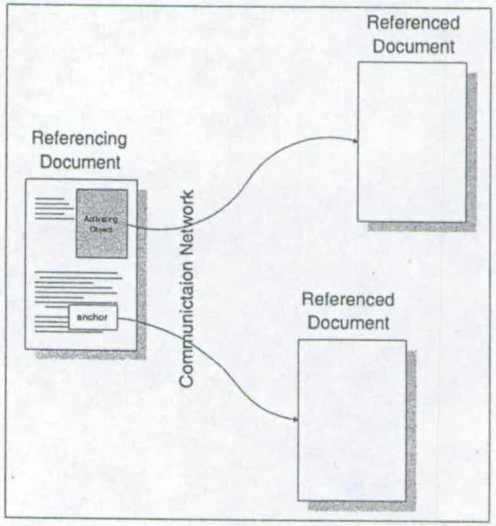


Figure 2: The link.

A link is embedded within a document and attached to part of latter's content by its *anchor* (see Figure 2) [Neilsen 90]. With respect to a particular link, we will define the object in which the link is embedded as the *referencing document*. Similarly, the object that the link points to will be defined as the *referenced document*.

As well as being identified with its anchor, referencing, and referenced documents, a link also has behavioural properties which become evident when it is activated and the referenced document is downloaded from the server to the user's client. This behaviour reflects the communication process which is implemented on the underlying network between the two locations.

Understanding Web User Behaviour

A series of studies of Web user behaviour were undertaken in order to learn more about how users made link activation decisions. A number of subjects were recruited with varying degrees of Web experience and expertise in computer use, ranging from novices to experts. Data about subject browsing behaviour was automatically logged by using a proxy server connected to the WWW client. This parsed the HTML pages that the subject downloaded and recorded the URL, number of images, number of links, number of applets and the download time from the proxy request to the proxy receiving the data. Each browsing session was also recorded using video and audio.

At the end of the session, all the collected data — video, audio and proxy logs — was reviewed and a detailed transcript of the session was produced. A re-enactment protocol was then employed to elicit from subjects the reasoning behind their actions [Wright 91]. All the utterances of the subject and the investigator were recorded on audio cassette and later transcribed.

Analysis of this data leads us to propose a three phase model of user link activation decision-making behaviour:

Assessment of value The user must decide prior to activating a link whether the referenced document contains information that is useful with respect to the user's overall objective. Reaching a positive assessment will lead to the link being activated and the download of the referenced document will commence; a negative assessment will result in the link being ignored. In addition to the assessment of the referenced document's content, the user also makes an assessment of the likely download quality of service (QoS). The potential of a long wait (e.g., because of server or network congestion, low network bandwidth) for the referenced document may persuade a user to ignore the link, or at least ignore the link for the present until a more favourable time. Our subjects often tried to infer by the URL information about the probable length of the wait. Here, heuristics and rules of thumb prompted by previous experience often come into play. For example:

> "This page is on a University server and that means it should be quick because they have fastish servers."

The assessment of value process that the user undertakes prior to activating a link can therefore, be split into two parts:

1. an assessment of content, and
2. an assessment of download QoS, particularly the anticipated delay in downloading the document.

Download Once the link has been activated, the document download begins. During this phase the user is exposed to the underlying behaviour of the communication network. As we noted above, users of interactive systems are often very sensitive to delays, especially where these are of an unpredictable extent. The problem takes on even greater significance for the Web, however, because browser user interfaces allow users to interrupt and abandon the download at any time. In our studies, observations of users choosing this option were very common.

Johnson has employed micro-economic cost-utility models and the notion of marginal utility to analyse Web user responses to download delays [Johnson 97]. He demonstrated that users' tolerance of delays could be manipulated by increasing their expectations of document quality. Johnson's findings are evidence for the close relationship between the assessment of value phase and the download phase; by increasing the information available to the user before they activate the link, the user is not only better informed about its nominal value, but is also better able to make an informed decision about whether the document is worth waiting for.

Some of our subjects revealed interesting beliefs about the extent to which their actions could influence the download time. For example, several subjects articulated the view that slow downloads could be speeded up by stopping and restarting them. When asked to explain this, one user drew upon a telephone analogy:

> "...[it] is like a crackerly telephone line, if you ring again the quality is often better."

Others provided more technically sophisticated rationales for their behaviour:

> "...when you stop and reload, it is possible to get a better route for the data, so it may be quicker."
>
> "...there is only a certain number of connections that a server can make, by reloading a page a new connection may have become free due to a time out or other users disconnecting."

Evaluation The final phase allows a user to evaluate the outcome of activating the link. If the activation was successful, the referenced document can be perused and its content can be compared with the expectations generated by the assessment of value phase. Similarly, the download performance can be evaluated. The evaluation process may stimulate the formation of heuristics that will serve to make sense of the experience, and which may then come into play in future cycles of interaction. A flavour of these heuristics is evident in explanations given by our subjects such as:

> "...all sites in Australia are slow."

> "...it is night time in America so all information will be transferred very fast."

We can now consider in more detail how the usability objectives defined earlier may be tackled. Our discussion will focus on two broad areas in turn: documents and links.

Documents

Our subjects sometimes used the URL to draw conclusions about the content of a document. For example, in the case of URL `www.dcs.cam.ac.uk/smith/publications.html`, one user was able to conclude that it referenced a list of publications (`publications.html`) by the author Smith (`smith`). Other parts of the URL may also yield useful pointers to document content, quality and relevance. For example, understanding how to decode `dcs.cam.ac` may lead a user to infer that the author's publications are in the field of computer science (`dcs`) and that they are probably academic papers (`ac.uk`). The user might even infer that, as a member of a prestigious university (`cam`), the quality of the papers is likely to be good.

Despite such user resourcefulness, it is evident that meta-data is the key element in improving users' capacity to make informed assessments of a document's quality and relevance [Bieber 97]. Of course, the form this meta-data will take must be dependent upon the kind of document it describes. Whilst it is easy to specify that the meta-data for a scientific paper would include its authors, key words and abstract, that of many other document types is more difficult to anticipate. The issue that must be addressed then is how such meta-data may be effectively presented to the user. One possible approach is to use meta-data to provide some insights in to the document structure. A simple illustration of how this might be done is shown in Figure 3. Here we see the document content summarised by:

1. Author, title, keywords, abstract and section names appearing in the document. HTML already provides some support for document meta-data through the use of a number of pre-defined tags. These include author, title, and keywords. The new proposed document markup language XML goes further in that it allows the author to define new tags; in fact, every component of an XML document may be tagged in this way.

2. The distribution of keywords through the document body. This can provide a better indication of relevance than a simple count [Hearst 95].

3. A thumbnail outline of the document. The concept of genre is one possible way in which broader notions of meta-data may be addressed. The use of genre in the print media, for example, enables people to make quick and easy distinctions between the vast variety of publications available. To help the readership distinguish between these news genre, simple visual cues such as the tabloid and broad sheet format have been developed. The thumbnail is an attempt to provide similar visual cues for Web documents.

4. The numbers and sizes of different embedded media types and the number of links. This also may provide the user with genre cues. For example, a document containing many links may appear to be a richer resource for browsing.

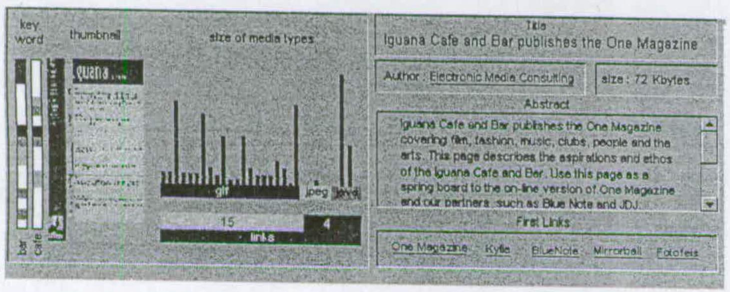


Figure 3: An example of document meta-data: structure and content summaries.

The Link

To provide the user with appropriate information about the link, it is necessary to identify an abstraction — or perhaps a layered set of abstractions — that can explain its behaviour [Dourish 95]. Our chosen link abstraction has two elements. The first is the *channel*, a communication resource which connects two (or more) computers together, and allows information to flow between them in either direction. If we restrict our notion of channel to the overall connection between the browser client and the document server, we have a concept that is synonymous with the communication associated with a download. The other element of our link abstraction is the *site*, a resource for despatching documents provided by the server. A channel and a site can now be associated with a link, and their properties can be presented to the user as affordances for understanding link behaviour.

The key property of a link is its QoS. Definitions of link QoS for multimedia applications cite several contributing factors [Guo 97]. For the Web user, however, the key QoS criteria is time. In simple terms, what the user wants to know before initiating a download is the delay before the download is completed. For the link, this can be defined as a function of site and channel latencies and channel bandwidth. Latencies reflect server and network congestion causing competition for these resources.

A full link time affordance, however, must also provide the user with information about the task in progress. The concept of time affordance is informed by users' subjective desire for closure, the knowledge that a system task has been successfully completed. Until this is evident, users will tend to monitor the task, assessing its progress and preparing contingency plans should they determine that it has failed. Lack of adequate time affordances may mislead users, for example, causing them to terminate the task erroneously.

According to Conn, a complete time affordance should meet eight requirements, representing information about distinct task properties [Conn 95]. We confine ourselves here to those where the link user interface is particularly deficient, and discuss ways in which an enhanced link might satisfy them. Some of the solutions we propose raise functional requirements which are addressed later.

Acceptance: Acknowledgement that the task has been correctly specified and is executable. Arguably, this is essentially a null requirement for browsers if the task 'specification' is defined syntactically as simply the act of pointing at any visible anchor. However, in so-called direct manipulation interfaces, syntactically acceptable action specifications have an implied semantic correctness [Schneidermann 86]. For example, a document's visibility is proof of its existence. In the case of an anchor, however, this relationship between visibility and existence may not hold.

Improved satisfying of acceptance requirements could be achieved by server and document status information being available when the user is in the act of specifying the download task. This is the most timely moment to provide this information and resolves the inconsistency between visibility and existence.

Scope: Measures of task scale, including predicted delay before completion. This requirement is not satisfied by the information provided through the link user interface. It may be partially satisfied by contextual information provided by the author of the referencing document, such as the referenced document's size.

Download scope is a function of the referenced document size and the link QoS. Satisfying scope requirements implies that these factors can be determined in advance of the download being initiated, In practice, however, only the document size can be accurately determined in this way — given suitable functional extensions.

Link QoS figures should take into account the effects of caching. Caches were introduced to exploit locality in users' document downloading behaviour. For example, there may be a cache on the user's (client) machine and a cache on a local proxy serving a number of users. If a document has been previously retrieved it could be available on these caches. A cached document will almost always download more quickly than the copy on its server.

Progress: Clear feedback as execution of the task proceeds, including its rate of progress. This requirement we consider is only partially satisfied. Although a progress indicator is maintained during the download, it is often difficult to relate the information it provides of the task state to its subsequent behaviour. One reason for this is that a document which may seem to the user to be a single entity may actually be composed of many embedded parts (e.g., images) each of which will trigger its own download process. The order of these embedded downloads is non-deterministic. In most browsers[2] the progress display area is time-multiplexed between these separate downloads, making it impossible for the user to maintain a coherent picture of the overall download progress [McManus 97].

An example of an improved link time affordance is shown in Figure 4. The map display shows the channel as a physical connection between the user's client and the document server. Its colour is used to represent predicted channel QoS (ranging from blue for poor to red for good). The graph shows site QoS information and the predicted download scope. Download progress is indicated by animation of the document media content display in Figure 3.

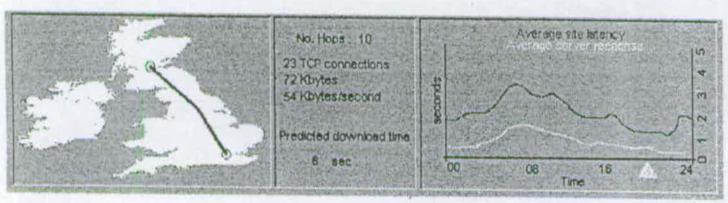


Figure 4: An example of link QoS data: channel and site QoS summaries.

[2] Sun's Hotjava is an exception in providing progress indicators for embedded downloads.

Functional Implications

The possible improvements outlined above are more than cosmetic and raise additional functional requirements of the underlying system.

If the link is to be able to provide the user with meta-data, this information must either be provided by the referencing document's author, or be downloaded from the referenced document's server. The former approach has been applied in some specialised hypermedia publications [Bieber 97]. However, our goal is to make meta-data available for a wider range of documents, and this is only achievable by following the download approach. To avoid this creating its own delays, two requirements must be met. First, the meta-data must be accessible to the client without downloading the whole document. Second, some kind of meta-data pre-loading or transclusion strategy is required so that it is available in a timely way to the user.

The HTTP protocol already supports client access of document parts through the HTTP-EQUIV directive. This allows authors to specify document tags which can be downloaded separately from the rest of the document and so HTTP already has the necessary hooks to meet the first requirement. The second requirement can be met by appropriate client side enhancements. There are a number of possible meta-data pre-load strategies. For example:

- All link meta-data is pre-loaded when the referencing document is opened. This is the least efficient strategy. Its drawbacks include the fact that the number of links may be large. The following three strategies adopt a progressively more conservative approach.
- Pre-loading is limited to those links that are within the currently visible area of the referencing document.
- Pre-loading is triggered when the user moves the link lens tool within a specified distance of the link anchor.
- Pre-loading is triggered when the user passes the link lens tool over the link anchor.

All the above pre-load strategies are assumed to operate automatically. It would be sensible, however, to allow user control over pre-loading, e.g., setting it on or off, and determining thresholds where appropriate. Yet another approach would be to give control of pre-loading to the referenced document's site. For example:

- Pre-loading is determined by the accumulated history of link activation by other users. This strategy would provide user with a form of collaborative filtering mechanism.

The accurate determination of download scope relies on knowing channel and site QoS. Obtaining real-time channel QoS data is problematic given current communication protocol limitations. The Hypertext Transfer Protocol (HTTP) is layered on top of the Transmission Control Protocol (TCP), which itself is layered on top of the Internet Protocol (IP). The latter does not support any QoS functions and so TCP cannot make any assumptions about factors such as latency or bandwidth. The only channel QoS related notion supported by IP is distance as measured by the number of 'hops' between the sender and receiver [Brebner 97]. Support for QoS is planned as part of an enhanced HTTP protocol [Guo 97], but in the meantime we must look to other solutions.

There are now several companies offering Web server QoS data on-line. Such data typically includes the number of hits per day, the average latency, and the numbers of broken links. However, we have decided to use a different approach which, though less accurate, does not assume the availability of site QoS servers. The link lens utilises the performance of the meta-data pre-load cycle to calculate channel and site QoS for each link. In the limit case where link pre-loading has not been completed before the user wants to have access to it, the link can simply be assigned a low QoS, leaving the user to draw his or her own conclusions.

A Prototype Link Lens

The concept upon which the link lens is based is the so-called Magic Lens or movable filter [Bier 94, Stone 94]. Magic lens filters combine arbitrarily shaped visual field with an operator that changes the presentation of objects within the view field. In the link lens, the view field is bound to the cursor and is sensitive to any anchors that it is passed over. When applied in this way, its two operators, document and link, transform the anchor into views of document meta-data and link QoS. The prototype link lens is shown in Figure 5. The design was evolved from paper-based sketches of possible lens features. These were presented to six subjects who rated them according to perceived usefulness.

Their overall opinion was that the link lens would affect their browsing strategies. With the exception of the document abstract, there was no consensus about which parts of the link lens would be the most useful, but all subjects expressed a desire to be able to choose a subset of meta-data for viewing at one time. Depending on the task being undertaken, the subjects felt that some parts of the link lens could be beneficial to the usability of the system.

We are currently carrying out further testing of several versions of the link lens with groups of Web users. The objectives include evaluating the value to users of the various different kinds of link information. We are also investigating different ways of presenting the information, including enabling users to customise or adapt the link lens to match their needs, and more economical encodings based upon the appearance of the link in the referencing document. A particular emphasis of this work will be to explore how degrees of user experience affect their ability to make use of link information.

Conclusions and Further Work

The prototype link lens presents document meta-data and time affordances in the form of link QoS information. Its goal is to provide Web users with timely and relevant information prior to their activation of links and during subsequent downloads. Underlying our work is the belief that by making a more "truthful" picture of Web behaviour available, users will be able to formulate more appropriate explanations of its behaviour and make more informed decisions about its use.

Future work will focus on evaluation and exploration of improved techniques for link information presentation, including the value of layered abstractions of Web behaviour. The question of how much users need to know about Web behaviour is a priority for further investigation. The evidence from our user studies confirms that users will generate explanations of Web behaviour and will act upon them. Their reasoning and the analogies they draw upon, however, range from the simple to the sophisticated. So, it is probable that novice users will need different kinds of abstractions from those for more experienced users. Our aim in developing the link lens is to provide support for all kinds of users. An important part of our strategy for achieving this is to recognise that in the hands of a less experienced user, the link lens may serve as a tool for learning about how the Web behaves.

As part of a broader programme of work for improving the value of the Web, we also plan to investigate the use of collaborative filtering — document usage and user ratings [Hill 95, Procter 97] — as additional document quality and relevance information. Finally, as HTTP protocol enhancements become available, we will examine how they may be applied in this context.

Finally, any mechanism for improving Web users' capacity to find relevant information must ultimately rely upon the willingness of content providers to recognise users' needs. In particular, the need for content providers to include meta-data is paramount. The link lens is designed so that the user will find it easy to distinguish between content providers who are conscientious in this regard and those who are not. In this way, we believe that putting tools like the link lens in the hands of users can provide an incentive for content providers to formulate and adhere to "good authoring practice".

References

[Bier 94] E. Bier, M. Stone, K. Fiskin, W. Buxton and T. Baudel, A Taxonomy of See-Through Tools. In *Proceedings of CHI'94*, ACM Press, 1994, pp 358-364.

[Bieber 97] M. Bieber, F. Vitali, H. Ashman, V. Balasubramanian and H. Oinas-Kukkonen, H. Fourth generation hypermedia: some missing links for the World Wide Web, *International Journal of Human-Computer Studies*, 47, 1997, pp 31-65.

[Brebner 97] G. Brebner, Computers in Communication, McGraw-Hill, 1997.

[Conn 95] A. Conn, Time Affordances: The Time Factor in Diagnostic Usability Heuristics. In *Proceedings of CHI'95*, ACM Press, 1995, pp 186-913.

[Dix 87] A. Dix, The Myth of the Infinitely Fast Machine. In D. Diaper and R. Winder (Eds.), *People and Computers III*, Cambridge University Press, 1987, pp 215-228.

[Dix 94] A. Dix, Seven Years on, the Myth Continues. Research Report RR9405, University of Huddersfield, 1994. `http://www.soc.staffs.ac.uk/~cmtajd/papers/myth95/7-years-on.html`

[O'Donnell 95] P. O'Donnell and S. Draper, How machine delays change user strategies. In M. Kirby, A. Dix and J. Finlay (Eds.), *People and Computers X: HCI'95*, Cambridge University Press, 1995.

[Dourish 95] P. Dourish, Accounting for System Behaviour: Representation, Reflection and Resourceful Action. In Proceedings of *The Third Decennial Conference on Computers in Context*, 1995, pp 147-156.

[GVU 98] Graphics, Visualisation and Usability Center, 8th WWW user survey, Georgia Tech, 1998. `http://www.gvu.gatech.edu/user_surveys/survey-1997-10/graphs/use/Problems_Using_the_Web.html`

[Guo 97] X. Guo and C. Pattinson, Quality of Service Requirements for Multimedia Communications. Workshop on Time and the Web, Staffordshire University, June, 1997.

[Hearst 95] M. Hearst, TileBars: Visualisation of Term Distribution Information in Full Text Information Access. In *Proceedings of CHI'95*, ACM Press, 1995, pp 59-66.

[Hill 95] W. Hill, L. Stead, M. Rosenstein and G. Furnas, Recommending and Evaluating Choices in a Virtual Community of Use. In *Proceedings of CHI'95*, ACM Press, 1995, pp 194-201.

[Johnson 95a] C. Johnson, Time and the Web: Representing Temporal Properties of Interaction with Distributed Systems. In M. Kirby, A. Dix and J. Finlay (Eds.), *People and Computers X: HCI'95*, Cambridge University Press, 1995, pp 39-50.

[Johnson 95b] C. Johnson, The Challenge of Time. In P. Palanque and R. Bastide (Eds.) The Design, Specification and Verification of Interactive Systems, Springer Verlag, 1995, pp 345-357.

[Johnson 97] C. Johnson, What's the Web Worth? The Impact of Retrieval Delays on the Value of Distributed Information. Workshop on Time and the Web, Staffordshire University, June, 1997.

[Kieras 90] D. Kieras and S. Bovair, The Role of a Mental Model in Learning to Operate a Device. In J. Preece and L. Keller (Eds.), Human-Computer Interaction, Prentice Hall, 1990, pp 205-221.

[McManus 97] B. McManus, Compensatory Actions for Time Delays. Workshop on Time and the Web, Staffordshire University, June, 1997.

[Neilsen 90] J. Nielsen, Hypertext and Hypermedia, Academic Press, 1990,

[Nielsen 98] J. Nielsen, Using Links to Help Users Predict Where They are Going. Jakob Nielsen's Alertbox, January, 1998. `http://www.useit.com/alertbox/980111.html`

[Parker 97] H. Parker, Temporal Usability and Disturbance Management in Interaction. Workshop on Time and the Web, Staffordshire University, June, 1997.

[Planas 88] M. Planas and W. Treurniet, The effects of feedback during delays in simulated teletext reception, *Behaviour and Information Technology*, 7(2), 1988, pp 183-191.

[Procter 97] R. Procter and A. McKinlay, Social Affordances and Implicit Ratings for Social Filtering on the Web. In L. Kovacs et al. (Eds.) Proceedings of the Fifth DELOS Workshop on Filtering and Collaborative Filtering, Budapest, Hungary, November 10th-12th, 1997. ERCIM Press.

[Ramsay 98] J. Ramsay, A. Barbesi and J. Preece, A psychological investigation of long retrieval times on the World Wide Web, *Interacting with Computers*, 10, 1998, pp 77-86.

[Schneidermann 86] B. Schneidermann, Designing the User Interface, Addison Wesley, 1986.

[Sears 97] A. Sears, J. Jacko and M. Borella, Internet Delay Effects: How Users Perceive Quality, Organization, and Ease of Use of Information. In *Proceedings of CHI'97*, ACM Press, 1997, pp 353-354.

[Stone 94] M. Stone, K. Fiskin and E. Bier, The Movable Filter as a User Interface Tool. In *Proceedings of CHI'94*, ACM Press, 1997, pp 306-312.

[Wright 91] P. Wright and A. Monk, The Use of Think-Aloud Evaluation methods in Design, *SIGCHI Bulletin*, 23(1), ACM Press, 1991, pp 55-57.

[XML] Extensible Markup Language. W3C. `http://www.w3.org/XML/`

Human-Computer Interaction - INTERACT'99 (Volume II)
S. Brewster, A. Cawsey & G. Cockton (Editors)
Published by The British Computer Society © IFIP TC.13, 1999

Human Factors and the WWW: Making sense of URLs

Dominic Stanyer and Rob Procter

Institute for Communicating and Collaborating Systems
Division of Informatics, University of Edinburgh
Scotland
Email: dsjs|rnp@dcs.ed.ac.uk

Abstract: We present a study of how WWW users 'make sense' of URLs. Experiments were used to investigate users' capacity to employ the URL as a surrogate for the resource to which it refers. The results show that users can infer useful information from URLs, but that such improvisation has shortcomings as a navigation aid.

Keywords: Universal Resource Locator (URL), usability, World Wide Web, navigation

1 Motivation

The typical user interface is an abstraction over the behaviour of the underlying system. The user interacts with system functionality through this abstraction which hides the complex realities of the underlying mechanisms (Dourish and Button 1998). We argue that the typical WWW browser hides too much. Users often find themselves having to navigate without a clear idea of either 'where' they are going, or of 'how long' it will take to get there.

Such deficiencies force users to improvise, to use any means to hand that might afford them a window onto a richer picture of the content, form and behaviour of the underlying system. As Nielsen (1999) observes, the Universal Resource Locator (URL) is one such window. In earlier experiments, ve explored ways in which URLs may afford useful lues concerning referenced resource content, ocation and even the behaviour of the download rocess (Stanyer and Procter 1999). We report here reliminary results from a follow-up study. This is art of on-going research to develop an nderstanding of how users use the World Wide /eb. The goal is to improve WWW usability rough better browser interfaces.

2 The URL

Every WWW resource has a unique reference, or URL, by which it can be retrieved. The three major parts of a URL are `site`, `path` and `resource` name, where the latter two parts being optional. Breaking down the URL into its components may enable users to extract useful information about content, location and download behaviour of the referenced resource. An experiment was undertaken to explore users' ability to make sense of URLs in this way.

3 Experimental Method

The subjects were eighteen Edinburgh University students with varying degrees of WWW experience. Each was presented with ten URLs in random order. They consisted of either: `site` name (short), `site` and `path` name, or `site`, `path` and `resource` name (long), and were chosen to represent a range of inference affordability. Subjects were asked to rate their confidence in predicting content, location, and download time on a scale 1 to 6. Downloading was then initiated, and subjects asked to relate the observed behaviour to the URL. On download completion, subjects were asked to rate the accuracy of their predictions. Subjects' remarks were recorded.

www.dcs.shef.ac.uk/~wmlh/ipcat97book.htm
"It's at Sheffield University, and it's in someone's personal home directory because it has a tilde for a UNIX account, home directory, and there is some acronym 97 which presumably means that it's the proceedings or something to do with the year 1997, that's what I would assume, I don't know that it will be. Book indicates that it might be an on-line version of a document ... like a paper document maybe and htm indicates that it is hypertext. My intuition would say that book htm could be fairly fast because it could be a big text file representation of a book. Again ... it is just me implying that."

Figure 1: Extract from subject transcript.

4 Results and Discussion

Averages of subjects' confidence and accuracy in location, content and download time prediction are shown in Table 1. They suggest that subjects were more confident at inferring location (e.g., see Figure 1) than content or download time. Subjects also rated their accuracy of location prediction higher than content or download time. Accurarcy is rated quite highly for all predictions. We suspect that this may be an effect of self-assessment. Performance in both tasks was better for long URLs.

The transcripts showed that subjects made sense of URLs in ways that were broadly consistent with URL format and constituent parts, but that they were easily mislead. E.g.: location inferences might be driven by recognition of a site name sub-domain, rather than a 'parsing' of the whole name. E.g.: most subjects based their inference of the location of `uk.samba.org` on its first sub-domain. There was a correlation at the 0.95 level between location and download predictions, suggesting that subjects often based their estimate of the latter on the former. This was confirmed by the transcripts. Other bases for download time inference included inferences about site popularity. E.g.: "Microsoft is always slow because so many people use it". Once download had started, subjects were able to rationalise observed download behaviour by reference to the URL. Subjects' strategy for inferring content had to be that URL parts were meaningfully named. It may be that lower confidence in content predictions reflected awareness of the unreliability of such an assumption. Content inferences were often conditioned by 'common sense'. E.g.: subjects were very confident about `www.scottishpower.plc.uk`, but much less so about `uk.samba.org`.

	Confidence	Accuracy
Location	4.9	5.2
Content	4.2	4.6
Time	3.9	4.0

Table 1: Average subject prediction confidence and accuracy.

5 Summary and Conclusions

The URLs were chosen because they exemplified different forms, not because they were typical of what users may encounter 'in the wild'. Nevertheless, the results suggest that though users can make quite artful use of URLs as surrogates for WWW resources, such improvisation has its limitations. It would seem that users are best at inferring location, yet location may be less valuable for navigation than either content or download time (Johnson 1997).

We conclude that browsers should afford users better means for navigating and understanding WWW behaviour. In particular, techniques such as link traversal preview and download time affordances deserve investigation (Kopetzky and Mulhauser 1999, Stanyer and Procter 1999). Such techniques can provide with browsable levels of abstraction without making too much of a commitment as to how these abstractions may actually be used.

References

Dourish, P. and Button, G. (1998) On Technomethodology: Foundational Relationships Between Ethnomethodology and System Design. Human Computer Interaction, Vol 13, 395-432.

Johnson, C. (1997) What's the Web Worth? The Impact of Retrieval Delays on the Value of Distributed Information. Workshop on Time and the Web, Staffordshire University, June.

Kopetzky, T. and Mulhauser, M. (1999) Visual preview for link traversal on the WWW. Proc. 8th International WWW Conference, Toronto.

Nielsen, J. (1999) URL as UI, Jacob Nielsen's Alertbox, March. http://www.useit.com/alertbox/990321.html

Stanyer, D. and Procter, R. (1999) Improving Web Usability with the Link Lens. Proc. 8th International WWW Conference, Toronto.