Information Retrieval and Human Computer Interaction

- A one day workshop in Glasgow -

Editors: Chris Johnson and Mark Dunlop,
Glasgow Interactive Systems Group,
Department of Computing Science,
University of Glasgow,
Glasgow, G12 8QQ.

GIST Technical Report G96-2, Glasgow University, 17th September 1996.

Table of Contents

Full Programme

09:45 Welcome

10:00 *Multilingual interfaces for IR systems: desparately looking for paradigms and examples*
Monnica Landoni & Forbes Gibb, Strathclyde University

10:30 *Design of an IRS for users' perception of the corpus*
Jean-Pierre Chevallet and Laurence Nigay, CLIP-IMAG Grenoble

11:00 Coffee

11:30 *Using usage data in collaborative environments*
Matthew Chalmers, UBILAB Zurich

12:00 *Use of conceptual models for a tourist information application*
Joe Bullock, Manchester University

12:30 *Information retrieval in astronomy*
Clive Davenhall, Institure for Astronomy, Edinburgh University

13:00 Lunch

14:00 *Usability inspection and evaluation of interactive IR systems*
Mark Treglown and Janet Finlay, Huddersfield University

14:30 *Cognitive aspects of information-seeking during complex learning tasks: a comparison of book and electronic versions of a medical handbook*, Joe Nicholls, Mark Howes and Rick Jones, Leeds University

15:00 *Discussion session*

15:30 Coffee

16:00 *Interfaces to support collaboration in information retrieval*
Mike Twidale and Dave Nichols, Lancaster University

16:30 *Providing expert assistance with IR in the virtual library*
Rob Procter, Lorna Banks, Kenny MacDonald, Elizabeth Davenport, Peter Burnhill & Sheila Cannell, Computer Science, Edinburgh University.

17:00 *Information retrieval: an issue of navigation in information space*
David Benyon, Napier University

17:30 Close

Introduction

Glasgow University's Department of Computer Science is fortunate in that it has research groups in both Human Computer Interaction (HCI) and Information Retrieval (IR). Over the last two or three years, the interests of these two groups have grown closer and closer together.

Research in the field of HCI has begun to focus on the use of distributed information systems within large work groups. This has increased interest about previous studies into 'hypertext' navigation and structuring. There has also been a renewed interest in the delays and timing problems that arise when users access distributed and shared resources. Finally, researchers in HCI have begun to focus upon appropriate retrieval techniques for new forms of multimedia information.

Research in IR has begun to look beyond the simplistic evaluation criteria provided by precision and recall. This has led to a renewed interest in retrieval activities and patterns of use. For example, it is of little benefit if an IR system provides relevant and focussed information long after the user has abandoned their query. Similarly, information is of little benefit if it is presented in a format that cannot easily be used to support the user's activities.

With so many areas for inter-disciplinary work, it is surprising that there have been relatively few collaborative research projects or joint workshops in HCI and IR. Recent interest in the World Wide Web has gone some way to correcting this. There is, however, a danger that Web technology will obscure our view of more fundamental research issues. Too many papers focus upon the surface details of particular browsers and the limited presentation strategies that they support. This workshop aims to look at the deeper problems facing HCI and IR. In particular, we have selected papers that focus upon:

- novel interaction techniques for information retrieval;
- the evaluation of IR systems;
- conversational modelling of information retrieval tasks;
- understanding the nature of relevance;
- networked information retrieval;
- browsing based information retrieval (inc. hypermedia);
- multimedia information retrieval.

The workshop is jointly organised by the British HCI Group and the British Computer Society's Information Retrieval Specialist Group and is hosted jointly by GIST and Computing Science at The University of Glasgow.

Multilingual Interfaces for IR systems: desperately looking for paradigms and examples.

Monica Landoni and Forbes Gibb
Department of Information Science, University of Strathclyde
Livingstone Tower, Glasgow G1 1XH.

Abstract
This paper is an overview of the state of the art in Multilingual Information Retrieval (MIR) interfaces. The main finding is that it is relatively unexplored area of research and while the need for these kinds of interfaces is growing guidelines are not yet available to assist designers in building them. In this paper we give some suggestions based on the experience with STAMP a European Commission funded project where our contribution has resulted in the production of a multilingual IR system for indexing and retrieval of text from Catalan, Spanish and English collections.

1. Introduction
IR is no longer concentrated on work with collections of English text. In the last few years there has been a growing interest and requirement for IR systems which are able to work in other languages, mainly European and Asiatic, (see TREC 94, 95). Until now the focus has been mainly on the construction of filters and the creation of stemming algorithms to handle non-English text collections, and modules which deal mainly with the functional, syntactic, and possibly semantic aspects of the problem. However there is another essential component of the IR system, the interface, which requires more attention if these multilingual systems are to be accepted and used in an operational context. Paying more attention to the language used by the user is a natural way to give greater consideration to his/her requirements. It is a way to move the focus from the functional computer-based aspects towards user needs and skills, by taking into account his/her cultural background by starting with one of the basic elements: language.

This paper contains a brief review of the state of the art in the design and production of multilingual interfaces, the result of a search for a set of guidelines for driving the design of a multilingual interface. It is also an overview of various issues related to the word "multilingual" which in this context would be better interpreted as multicultural. We have been working with two European languages, Spanish and Catalan, for which we had language experts on the team. As a result of this ongoing research a prototype of a multilingual interface, in Spanish, Catalan and English, has been produced and its design and implementation are described in this paper. Currently the prototype is almost complete and in autumn 1996 an evaluation of the interface will be undertaken. This evaluation methodology will have to be particularly sensitive to cultural issues and must place the user firmly at the centre of the system.

2. Multilingual IR systems
In order to make clear the purpose and content of this paper it is necessary to start with a definition of what a multilingual IR system is and does. In this context we will define a Multilingual Information Retrieval (MIR) system as any information retrieval system which deals with documents in more than one language. In order to make the system as simple as possible we consider as multilingual a system which can query collections in different languages while the documents inside each collection are all in the same language. Examples of such kind of systems include those involved in the TREC [1] experiments, where in addition to the classic English collection, TIPSTER, Spanish and a Chinese collections have recently been introduced. Outside the TREC experiment a small group of European projects are also studying specific examples of multilingual systems for information retrieval such as:

EMIR [2], European Multilingual Information retrieval system, a European project involving French, Belgian and German partners, for the development of a system based on both linguistic and statistical techniques for indexing multilingual text collections. The user can retrieve documents writing queries in his/her language while the result of the query can be in different European languages. The system has been developed for English , French and German.

EUROSpider [3] is a project at ETH, Zurich, in Switzerland. The main aim of this project is to build a collection of comparative multilingual documents, in Italian, German and French (being the three official languages in Switzerland) to provide linguistic tools for stemming and parsing of each language and to query the collection with any of the languages. At the moment the results obtained when retrieving Italian documents while using a German query demonstrate the validity of the chosen approach. This approach is thesaurus based and uses relevance feedback. A future development of this multilingual IR system will cover English, French, German and Italian.

EURODICAUTOM [4] is a multilingual database maintained by the European Union under the project ECHO. There is a Mosaic interface to access it from the net and it is possible to formulate queries using any of the European languages.

The MultiLingual Theory and Technology group [5] at the XEROX laboratory in Grenoble has been working on building linguistic tools such as morphological analysers, parsing and generation platforms and corpus analysis tools. Currently experiments on English and French collections have shown the feasibility of cross language multilingual IR.

While this is a recently new area of research it is clear that there is significant commercial potential as it becomes more evident that there is a need for multilingual interfaces in order to offer a natural query environment to users who speak different languages.

3. Multilingual Interfaces: the state of the art

The state of the art for multilingual interfaces is somewhat underdeveloped. In general the electronic world has chosen English as the common language, and more specifically technical English. People from different countries are obliged to be able to read in English, even if they are not able to speak it correctly, in order to have access to technical information. However information retrieval need not be purely for technical issues. In an ideal IR environment the user should be left free to express his/her needs in the most intuitive and personal way, and system would have to be able to handle his/her request and answer in an appropriate way. It is clear that under these assumptions technical English is unlikely to be the first choice as a query language for most people. However most interfaces available at the moment use English because of historical reasons and the lack of alternatives. Attempts to translate these interfaces into different languages without a proper process of interpretation can lead to ambiguous situations where the user will still prefer to use the English version in order to avoid confusion. So adaptation and interpretation, more than translation, are the keywords when proposing guidelines for the design of multilingual interfaces.

4. Proposed guidelines for multilingual interfaces

In order to define guidelines for the design of multilingual interfaces for IR systems the first consideration is to try to use what has been accomplished in defining paradigms for IR interfaces and to see how this can be modified to answer the specific needs of a multilingual situation.

The easiest solution would be simply to literally translate from English a classic and suitable version of an IR interface. Even if this sounds quite easy, rapid and logical, it implies a lot of study into what the English interface really means and what its various elements represent. This is why before suggesting a set of guidelines, it is important to examine the various components of an IR interface and see how they are influenced by the linguistic/cultural component. Speaking in very broad terms it is possible to identify some basic elements which, when properly combined, result in an interface. Text, icons, sounds and colours are the most elementary components. Their use is then influenced by user factors and machine factors in different ways. The designer has the role of considering all these factors and proposing their best combinations in order to produce the interface, which has to satisfy the user and also be efficient and suitable for implementation on the chosen platform. Thus the designer has to consider on the user side at least :

- needs and skills;
- taste and education;
- cultural influences.

On the machine side factors for consideration include:

- technical limitations;
- existing paradigms of interaction;
- consistency with established interface principles which depend on the kind of machine (e.g. Mac vs. PC)

After this preliminary analysis, which is common to the design of a generic interface, the focus for multilingual ones moves naturally to language and its crucial role. The meaning of a language goes beyond its syntax and vocabulary; it is the use of the language which indicates trends and interests in a society. Paying more attention to the language used by the user is a natural way to give him/her more attention. It is a way to move the focus from the computer towards the user needs and skills, by paying attention to his/her cultural background by starting with one of its basic elements: language.

Adaptation is the process of translation and interpretation of some text in order to perform the same role within a different cultural context. Therefore it is the critical part where there is a need for linguistic support in order to capture the real meaning and essence of the technical English and to produce a proper interpretation of the interface in a different language.

In a second step the same kind of attention will be paid to the appropriate choice of sounds and colours which will have to respond to cultural trends and accurate interpretations (e.g. colours have different meaning in various part of the word for instance white, which stands for purity in our culture, stands for sorrow in the Chinese one).

5. The STAMP experiment

STAMP is a European project involved in the design and development of a tool set for the integration of distributed multimedia systems across different platforms.

The contribution of our group to the project has been the design and implementation of an IR system based on linguistic and statistical techniques for the indexing and retrieval of documents belonging to different collections, in Catalan, Spanish and English. At this stage each collection contains documents all in the same language but future developments contemplate the possibility of dealing with multilingual collections as well.

Particular attention has been paid while designing the interface for the retrieval module to making this consistent with the chosen platform, PC with Windows 95. At the same time other examples of IR interfaces based on relevance feedback have been studied in order to achieve an intuitive and self explicative interface able to drive the user through his/her retrieval task. Currently the prototype for the retrieval interface, originally in English, is ready and the adaptations for Catalan and Spanish have been developed with the help of linguistic experts. However given the lack of paradigms available our system is likely to undergo substantial evolution. While the team of linguists has made substantial progress in adapting the interface there is still room for improvement especially after the evaluation stage when we will have the opportunity to have Catalan and Spanish speakers as users.

6. Conclusions

The main conclusion is that this is an extremely wide and interesting area of research still to be investigated. We think that language and its cultural value is a crucial key in the design of a multilingual interface and that the distinction between translation and adaptation can be a starting point for the production of design guidelines for MIR interfaces. Much work has still to be done, and first of all we will need to define a suitable evaluation strategy to test the validity of our approach on the multilingual side.

7. References

[1] TREC, [Online], (January, 1996). Available: http://potomac.ncsl.nist.gov/TREC/.

[2] EMIR, [Online], (December, 1993). Available:
http://www.newcastle.research.ec.org/esp-syn/text/5312.html.

[3] EUROspider, [Online], (May, 1996). Available: http://www-ir.inf.ethz.ch/eurospider.html.

[4] EURODICAUTOM, [Online]. Available: http://www.uni-frankfurt.de/~felix/eurodictautom.html.

[5] MultiLingual Theory and Technology, [Online], (February, 1996)
Available: http://www.xerox.fr/grenoble/mltt/home.html.

Using Usage Data in Collaborative Information Environments

Matthew Chalmers
UBILAB, Union Bank of Switzerland, Bahnhofstr. 45, CH08021 Zurich
chalmers@ubilab.ubs.ch
http://www.ubs.com/ubilab/Projects/hci.html

Abstract
Much work in virtual environments has focused on system support for supporting shared visibility, shared databases and awareness of others. Perhaps because of way that early VEs focused mostly on modelling real world environments, it is only more recently that more attention has been given to information environments. Here, the basic structures and objects shown are abstract data such as collections of documents or financial trade data. In this situation, one must artificially create the landmarks and other features which aid in navigation and which support awareness of the activities of others. This has been the subject of some recent work at Ubilab. This work has primarily focused on bibliographic data, but has also involved time series of stock values. The raw data is combined with usage information such as what documents have been searched for, what keywords have been used, and what objects have been clicked on for examination. The intention is to improve the legibility and imageability of the visualisation by making manifest more of people's interactions with the data, thereby making imageability features better conform to use, and also making such features serve as tools for mutual awareness.

1. Adaptive Information Environments
Creating effective visualisations of complex, abstract data is not a trivial task. Many of the problems come from the fact that the number of items that can be handled effectively is constrained. In everyday spreadsheets and database front-ends, users can only compare data in simple ways. Simple graphs and matrices, linear lists of database query matches: such structures are weak in conveying the richness of the stored data. The context of an item matching some query, the inter-relationships of the retrieved set, the relationship of that set to previous work, the patterns that would be apparent from an overview of the full data set - all these things are discarded in traditional database interfaces. Queries are too often a fumbling grasp into the black box of the database. Instead of an ongoing interactive process, each query essentially conforms to a batch mode of operation: the query is composed, sent off for processing, and executed in sterile isolation from the users.

Supporting the richer variety of ways in which people interact is a difficult task, but it is important that we break away from the idea that only one person uses one retrieved item of data for one instant of time. Where information comes from, who else is interested in it, what has been done with it previously, what could we potentially do with it: in our everyday lives we work with these issues as well as raw data values. These issues can change how we perceive and use information, making us adjust our perception of what is useful, significant or interesting. Usage and collaborative issues therefore should affect how we design interfaces to databases, even though at the lower data management levels we still enforce more traditional rules of control and consistency.

At Ubilab, we are exploring techniques for efficiently laying out high-dimensional data in the form of maps or landscapes which can be browsed and explored graphically [Chalmers96a]. Each data object is positioned in the modelled space, and then a triangle mesh is created to form the basic landscape model. We are trying to design easily perceivable representations of this data so that people can better maintain cognizance of - a `feel' for - the entire data set [Chalmers95]. Maps generated using these techniques are visualised using the DIVE virtual environment system [Fahlén].

Our maps show more than simple positions of objects in 2- or 3D space. Such an approach is not enough to make such graphical representations useful, legible, and imageable. One has to reveal structure and detail while avoiding occlusion and clutter. We wish to open up a database so that as people move through it and perform fine-grained tasks such as searches within it, they maintain awareness of patterns within the data and over time can build up their knowledge about it. To some extent we are motivated by work on using the information of others such as [Goldberg], where raw data is supplemented by user-derived information on quality or interest. With such approaches in mind, we have been enriching our information designs by adding static and dynamic features for imageability.

Static imageability features include landscape shorelines and local areas of surface roughness. These are based on the positions of bibliographic objects alone and currently do not move or adapt after the initial map creation. Dynamic features are further subdivided into two types: view-specific and view-independent. Specific to the current field of view are successively sampled documents, with a bias based on nearness to the eye. A few appear as the user moves through the space, and more are shown when motion stops. These 'pop-ups' are highlighted with a bright colour and have detail shown, such as titles. The orientation of text is kept orthogonal to the view of the user as he or she moves. This aids legibility for that user, but also helps support awareness of the activities of others. The text and its orientation are visible from far away, and offer extra information on a user's activity to other users in the same virtual environment.

We wish to be able to see detail about nearby parts of the landscape but also some detail about distant objects. Our interest ranges from individual documents to sets of documents in perhaps quite distant regions of the landscape. We wish to find which words dominate in such distant regions i.e. which topics predominate there, and therefore which regions are good candidates for further exploration and detailed view. Furthermore, we wish to do this in a way that is relatively passive for the user i.e. the landscape should be self-disclosing and unfolding. Rather than rely on more complex artificial intelligence or information retrieval techniques to discover topics in text, we combine layout positions with information on the work of previously active users, and with the changing field of view of each user.

Independent of the viewpoint are indicators of usage frequency which are shown only on demand. This usage information reflects past activity by all users as measured by keyword search hits, and is updated within the virtual environment as such activities progress. This data is also used in dynamically placing topic words in the scene. These are based on frequency of occurrence in the field of view and a `popularity' metric based on word usage history amongst all users. Topic and title words can be clicked on to trigger searches, which use colour to show patterns of hits. Each such click increments the word's popularity rating. At present, popularity ratings are shared by all users i.e. the rating is given by the community and not the individual. Each topic is shown at the centre of mass of those visible documents within which the word occurs.

These first steps in using usage data are described in more detail in [Chalmers 96b]. There, we also describe the use of the techniques of [Ingram] to create static clusters or regions of objects based on geometric proximity. A common colour was given to the triangles associated with the members of each cluster, producing large-scale features which were very useful for orientation and navigation. More recently we have instead been using topic and popularity data to create regions based topic words. These evolve slowly as popularity ratings change.

Although the virtual environment offers synchronously working people the opportunity to see each others' ongoing activities, making these transient activities retrospectively accessible to others may lead to greater problems of invasiveness. At this stage we have not implemented schemes to handle this, but we do acknowledge both the problem and the existence of previous work in this area such as [Bellotti]. By combining usage information with the raw data, in a way that is appropriate to the sensitivities (both perceptual and social) of the people using it, we may then regain and even enhance the rich and varied information that forms their work environment and which current mainstream systems ignore.

2. Ongoing and Future Work

The primary focus of our ongoing work is in making better and more varied use of usage data. To this end, we are working to make more on more of the document-related activities of users persistent.

We wish to implement policies which take into account privacy and invasiveness issues, especially as these issues will come more into the foreground as we add tools for filtering, combining and representing usage data. Such tools would allow users to select how items such as popularity assessments and topic words were made, using groups ranging from individuals to the whole community.

Another topic of current interest is the feedback of usage information into the layout and clustering process. An example is adjusting the landscape layout algorithm to weight words and documents in accordance with user activity, and not relying solely on its current criterion, the words contained in documents. The construction of features such as clusters should adaptively take into account user activity. The rankings used for topics may form a first step towards this.

We have also been reimplementing parts of the Bead system in Java, so as to ease software development, to allow Web-based experiments and to facilitate extension to new data types. This has allowed us to lay out and visualise financial time series data such as sets of companies' stock values over time, and also some more heterogeneously structured files of trading records. We are collaboratively constructing a simple pilot application with a section of our corporation concerned with foreign exchange trading, integrating Bead layouts into an existing commercial visualisation tool. In this way we plan to soon have a pilot running in the bank which will offer analysts a useful and novel tool, and gain for us valuable feedback and experience.

3. Conclusion

Information retrieval systems tend to underplay the context of users' documents, generally focusing on a narrow view of interaction akin to batch model computing. By opening up the 'black box' of the database to view, and thereby making it visible and explorable, we allow our perceptual skills to be brought to bear in interpreting the relevance or interest of a document, word, or user activity. By employing usage data in a visualisation we can make more use of people's own interpretations of and activities in the information. This may be done either explicitly (and usually synchronously) by having people mutually aware of each other in a shared virtual environment, or implicitly (and often asynchronously) in using usage history to gradually adapt the display. By taking advantage of time and person in this way, we can enrich document interpretation by examining a wider range of data characteristics. This affords circumvention of some difficult document analysis problems, and allows a move towards a style of system interaction more closely related to the day-to-day activities of the office or workplace.

Acknowledgements

This paper is an adapted and much reduced version of a jointly written paper [Chalmers96b]. Thanks are due to my collaborators on this work, Rob Ingram (now back at the University of Nottingham) and Christoph Pfranger.

References

[Bellotti] V. Bellotti & A. Sellen, Design for Privacy in Ubiquitous Computing Environments. In G. de Michelis, C. Simone and K. Schmidt (Eds.), Proc. 3rd European Conf. on Computer Supported Cooperative Work, (ECSCW'93), pp. 77-92. Kluwer, 1993.

[Chalmers95] M. Chalmers, Design Perspectives in Visualising Complex Information, Proc. IFIP 3rd Visual Databases Conference, Lausanne, Switzerland. Published as Visual Database Systems 3: Visual Information Management, S. Spaccapietra & R. Jain (eds.), Chapman & Hall, pp. 103-111, 1995.

[Chalmers96a] M. Chalmers, A Linear Iteration Time Layout Algorithm for Visualising High-Dimensional Data, to appear in Proc. IEEE Visualization, San Francisco, Oct. 1996.

[Chalmers96b] M. Chalmers, R. Ingram & C. Pfranger, Adding Imageability Features to Information Displays, to appear in Proc. ACM UIST '96,Seattle, Oct./Nov. 1996.

[Fahlén] L. Fahlén et al., A Space Based Model for User Interaction in Shared Synthetic Environments, Proc. ACM InterCHI'93, pp. 43-48, 1993.

[Goldberg] D. Goldberg et al., Using Collaborative Filtering to Weave an Information Tapestry, Comm. ACM 35(12), pp. 61-70, 1992.

[Ingram] R. Ingram & S. Benford, Legibility Enhancement for Information Visualisation, Proc. IEEE Visualization 95, (Atlanta, Oct 1995), pp. 209-216, 1995.

Information Retrieval in Astronomy

A C Davenhall
Institute for Astronomy, University of Edinburgh

This brief review surveys the use of information retrieval systems in astronomy. Heck and Murtagh (1993) present a more extensive discussion. Retrieval systems are used to extract information from free-text systems, numeric databases and combinations of the two. Free-text systems are used to search and access the astronomical literature. Unsurprisingly, there are no commercial abstracting services specifically devoted to astronomy, though some services, such as INSPEC, do cover astronomical journals. The use of these systems has been reviewed by Rey-Watson (1988) and Watson (1991). A number of more specialised (and limited) systems have been developed within the astronomical community (see, for example, Davenhall, 1993, Andernach *et al*, 1994[1] and Adorf, 1995 for lists). However, historically numeric and combined databases have been more important than free-text systems and most of the problems peculiar to astronomy occur in these systems.

Astronomy is an observational science. Its usual *modus operandi* is not conducting laboratory experiments, but rather the passive observation of celestial phenomena. It progresses by accumulating data on the positions and properties of celestial objects. The recording and systematising of astronomical observations is not a new activity. The oldest series of astronomical observations extant are records of the position of Venus, impressed on clay tables in the Babylonian city of Uruk (Erech) around 1650BC. Catalogues listing the position and brightness of the fixed stars first appear in classical antiquity. The first star catalogue was a small list of about a hundred objects, produced in the first half of the third century BC by two Ptolemaic Greeks, Timocharis and Aristyllus. In recent times the volume of data recorded has increased dramatically, and modern digital archives may contain gigabytes or terabytes of data.

A very wide variety of types of measurements are recorded, including images, spectra and time series. The data may be saved as raw measurements, stored in the form in the form in which they were recorded, or calibrated measurements where instrumental distortions have been removed and the values converted into physical units. Catalogues or observing logs are another important type of data set. These data sets are lists summarising the properties (position, brightness, classification etc.) of some set of objects. They may have been derived from a single set of observations or combined from different sets of observations and, indeed, instruments. The formats used to represent the data vary as widely as the data themselves, though the Flexible Image Transport System (FITS) format (Wells *et al*, 1981)[2] is a successful and widely used format for exchanging data.

Data must be preserved so that subsequently they can be retrieved and compared against new observations, compared with the predictions of theories and used to plan new observations. Often a separate archive will be constructed for a given instrument, telescope or observatory, with its own peculiar data format and *ad hoc* software to query the list of observations made or objects observed and to retrieve selected observations. The ubiquity of the internet and the availability of tools such as the World Wide Web have made the construction of such systems almost routine. The principal problems with such archives are that the user must learn to use a new piece of software for each archive he accesses, and, more fundamentally, he must have prior knowledge of which data are available.

However, two systems have been constructed to give uniform and consistent access to a variety of heterogenous data sets and archives; the Astrophysics Data System (ADS) in the United States and the European Space Information System (ESIS) in Europe. The ADS (Eichhorn, 1994)[3] was developed by NASA to provide uniform access to data archives from various NASA astronomy satellites, as well as more general astronomical catalogues. ESIS (Giommi *et al*, 1995)[4] was developed by the European Space Agency to provide access to a number of astronomy and space physics archives. Both systems were conceived in the second half of the 1980s and became available for use early in the current decade (just before the explosion of the World Wide Web). The systems are conceptually similar; each comprises a client which the user runs locally and which submits queries to, and receives information from, various remote databases and archives. The user manipulates the same interface to access a range of diverse remote archives. It is easy to establish what information is available in a variety of archives for a given object, wavelength range or region of the sky. The details of accessing the individual archives are hidden from the user. The ADS and ESIS were powerful and flexible systems, but both foundered and their funding has been discontinued or reduced. The ADS, however, has evolved into a powerful system for retrieving bibliographical information (Eichhorn *et al*, 1995).

The integrated databases SIMBAD (Egret *et al*, 1991)[5] operated by the Centre de Données astronomiques de Strasbourg (CDS) and NED (Helou *et al*, 1991)[6], the NASA/IPAC Extragalactic Database operated by the Infrared Processing and Analysis Center, California Institute of Technology since 1987, have been more successful. SIMBAD contains data for approximately one million objects (stars, other Galactic objects and external galaxies) and NED contains data for over 375,000 extragalactic objects. Both are `integrated' databases; for each object in the database they contain a variety of data obtained from various external sources. The information included for each object includes basic data (position, brightness, classification etc.) and bibliographic references (and abstracts) to scientific papers pertaining to the object. Ideally images and spectra of the objects would be included too, where they are available. Obviously, ingesting the information to create an integrated database requires a great deal of effort, particularly with regard to quality control and the standardisation of physical quantities and units.

A common feature of astronomical databases and archives is they they tend to be organised around individual objects[7] and searches of the archives are intended to select individual objects or groups of objects. Thus it is desirable that the retrieval system should support searches for named objects. Unfortunately, there is no consistent, universal naming scheme for astronomical objects. Indeed, numerous schemes are in use, some of which are at least partly contradictory. Also the brighter stars, in particular, enjoy multiple designations. For example, some of the alternative designations of the star Aldebaran are: α Tau, 87 Tau, BS 1457, HR 1457, BD +16d[1] 629, HD 29139 and SAO 094027. Much effort has gone into systematising the various naming conventions (Fernandez *et al*, 1983, Lortet and Spite, 1986, Borde and Lortet, 1990, Lortet *et al*, 1994). If an archive is to be used effectively it must have a dictionary of cross references for the various names of objects, so that a user can search it by specifying any of its names. Currently the most extensive such dictionary is the *Catalogue of Stellar Identifications*, (Ochsenbein and Bischoff, 1982). It is a fundamental part of the SIMBAD database and is also used by NED, ADS and ESIS. The software to search such dictionaries should allow for permitted variations in the name, such as the absence or inclusion of optional spaces. Dictionaries of cross references are rarely incorporated into individual archives and are a significant advantage of the integrated and distributed systems.

An alternative to searching for objects by name is to search by position. The positions of astronomical objects are normally specified using two-dimensional spherical-polar coordinates analogous to terrestrial latitude and longitude. In general the search position specified and the position of the target object in the archive will not match exactly, in part because positions are measured quantities with associated errors. The search procedure must find the object (or objects) within a given tolerance (or distance) from the specified position. If the search is to proceed reasonably efficiently it is usually necessary that the catalogue or observing log should be indexed on at least one of the two coordinates. Some large modern catalogues, such as the *Two Micron All Sky Survey*, (Kleinmann *et al*, 1994) contain millions or even hundreds of millions of objects. In this case it is often necessary to use two-dimensional indexing techniques (see, for example, Page, 1996 for a discussion).

Acknowledgments

Thanks to C.G. Page (University of Leicester) for useful comments on the draft document.

Notes

[1] See URL: `http://www.eso.org/online-resources.html`

[2] The reference given is to the original paper proposing and describing the basic FITS format. The FITS standard is now maintained and documented by the FITS Support Office of the Astrophysics Data Facility at the NASA Goddard Space Flight Center. See URL:

`http://www.gsfc.nasa.gov/astro/fits/fits_home.html`

Though basically an astronomical standard, FITS is occasionally encountered outside astronomy; see, for example, Kay and Levine (1995).

[3] See URL: `http://adswww.harvard.edu/`

[4] See URL: `http://www.esrin.esa.it/htdocs/esis/esis.html`

[5] See URL: `http://simbad.u-strasbg.fr/`

[6] See URL: `http://www.ipac.caltech.edu/ned/ned.html`

[7] Here the word `object' is being used in its astronomical sense of a heavenly body, not in any of its computer science senses.

References

H.-M. Adorf, 1995, *Vistas in Astron*, **39**, pp243-253.

[1] Editors note - this should be a degree symbol

H. Andernach, R.J. Hanisch and F. Murtagh, 1994, *Publ. Astron. Soc. Pacific*, **106**, pp1190-1216.

S. Borde and M.-C. Lortet, 1990, *Bulletin d'Information du CDS*, **38**, pp41-67.

A.C. Davenhall, 1993, *An Astronomer's Guide to On-line Bibliographic Databases and Information Services*, Starlink User Note 174.1 (Starlink: Didcot, Oxfordshire).

D. Egret, M. Wenger and P. Dubois, 1991, in *Databases and On-Line Data in Astronomy*, eds. M.A. Albrecht and D. Egret (Kluwer: Dordrecht) pp79-88.

G. Eichhorn, 1994, *Expr. Astron*, **5**, pp205-220.

G. Eichhorn, A. Accomazzi, C.S. Grant, M.J. Kurtz and S.S. Murray, 1995, *Vistas in Astron*, **39**, pp217-225.

A. Fernandez, M.-C. Lortet and F. Spite, 1983, *Astron. Astrophys. Suppl*, **52**, no. 4.

P. Giommi, S.G. Ansari, P. Donzelli and A. Micol, 1995, *Expr. Astron*, **6**, pp143-162.

A. Heck and F. Murtagh, 1993, eds. *Intelligent Information Retrieval: The Case of Astronomy and Related Space Sciences* (Kluwer: Dordrecht).

G. Helou, B.F. Madore, M. Schmitz, M.D. Bicay, X. Wu and J. Bennett, 1991, *in Databases and On-Line Data in Astronomy*, eds. M.A. Albrecht and D. Egret (Kluwer: Dordrecht) pp89-106.

D.C. Kay and J.R. Levine, 1995, *Graphics File Formats second edition* (Windcrest/McGraw-Hill: New York). See in particular chapter 18, pp235-244.

S.G. Kleinmann, M.G. Lysaght, W.L. Pughe, S.E. Schneider, M.F. Skrutskie, M.D. Weinberg, S.D. Price, K. Matthews, B.T. Soifer, J.P. Huchra, C.A. Beichman, T.J. Chester, T. Jarrett, G.L. Kopan, C.J. Lonsdale, J. Elias, J.W. Liebert and P. Seitzer, 1994, *Astrophys. Space Sci*, **217**, pp11-17.

M.-C. Lorte, S. Borde and F. Ochsenbein, 1994, *Astron. Astrophys. Suppl*, **107**, pp193-218.

M.-C. Lortet and F. Spite, 1986, *Astron. Astrophys. Suppl,* **64**, pp329-389.

F. Ochsenbein and M. Bischoff 1982, in IAU Colloq 64, *Automated Data Retrieval in Astronomy*, eds. C. Jaschek and W. Heintz (D. Reidel: Dordrecht) pp211-213.

C.G. Page, 1996, *Proc. Eighth International Conference on Scientific and Statistical Database Management*, Stockholm, (IEEE Computer Society Press), pp 44-52.

J.M. Rey-Watson, 1988, in *Astronomy from Large Databases* eds. F. Murtagh and A. Heck, ESO conference and workshop proceedings no. 28, (ESO: Garching), pp453-458

J.M. Watson, 1991, in *Databases and On-Line Data in Astronomy*, eds. M.A. Albrecht and D. Egret (Kluwer: Dordrecht), pp199-210.

D.C. Wells, E.W. Greisen, and R.H. Harten, 1981, *Astron. Astrophys. Suppl*, **44**, pp363-370.

Managing Temporality in Information Retrieval Interfaces

Helen Parker and Chris Roast,
Computing Research Centre, School of Computing and Management Science,
Sheffield Hallam University, Sheffield S11 8HD

Abstract

This paper is concerned with the design of interactive systems for retrieving information from large and complex information stores. The retrieval of relevant information from large information stores and data corpuses can incur substantial computational delays, [Roast, 95]. We assess the impact that such undesirable system behaviour can have upon effective interaction.

The paper initially posits a simple model of information retrieval in exploratory information acquisition, in which a user initiates a search for information without a highly specific goal and without a clear idea of the outcome of the search. The user develops a strategy for effective information retrieval iteratively, narrowing the goal in response to the outcomes of earlier queries in a session. Within this model of information retrieval, we observe that an understanding of user behaviour in the presence of significant system delays has not been adequately addressed. A more dynamic model of the user is required, in which subjective "user time" can be compared with the objective timing of the information retrieval system, [Dix, 90; Parker et al, 96].

We review current findings in the psychological study of time and describe a dynamic account of cognition as a tuning process, in which reminding and anticipation drive knowledge and action, [Michon, 93; Schank, 82]. The user's subjective experience of time arises from tuning to event and context changes. These may be external changes at the interface, or in the wider information retrieval environment and internal changes in the user's cognitive processing.[Estes, 85; Block, 90; Grant, 94]

In this account, the user's judgments of time, past, present and future, cannot be assumed to be veridical. User judgments do not necessarily bear a direct relation to actual durations and orders of events at the interface. They are sensitive to a range of factors, including information complexity, the degree of active engagement in a task and the relative pace of a task, [Hicks et al, 76; Vroon, 70; Pedri and Hesketh, 95]. We consider implications of the variability of users' temporal judgments for effective planning and decision making in managing interactive, exploratory information retrieval.

These findings from the psychological study of time and their contribution to the design of information retrieval interfaces are illustrated using two contrasting information retrieval interface models. We take the view that the behaviour of an interface can be interpreted as a theory of the user, that is, the interface defines who will be able to use the system effectively. An information retrieval system provides a theory of "user time" which can be contrasted with the psychological findings.

We conclude our examination of interface design by proposing that information retrieval is not only dependent upon the user having a model of the *logical* organisation of stored information. The user's ability to assess the *value* of information, relative to the time required for retrieval and subsequent processing, is also necessary for effective interaction. This, in turn, requires the user to be cognisant of aspects of the *physical* organisation of information. Such a proposal cuts across the conventional separation of concerns within information systems design.

References

Block, R A, "Models of Psychological Time", in "Cognitive Models of Psychological Time", ed Block R A, Lawrence Erlbaum Associates, 1990

Dix, A J, "Non-determinism as a Paradigm for Understanding the User Interface", in "Formal Methods in Human Computer Interaction", eds Harrison M D and Thimbleby H W, Cambridge University Press, 1990

Estes, W K, "Memory for Temporal Information", in "Time, Mind and Behaviour", Michon, J A ed, Springer Verlag, pp151-168, 1985

Grant, A S, "Modelling complex cognition: contextual modularity and transitions", in Proceedings 4th International Conference on User Modelling, Hyannis, MA, USA, 157-162, 1994

Hicks, R E, Miller, G W and Kinsbourne, M, "Prospective and retrospective judgments of time as a function of amount of information processed", American Journal of Psychology, 89, pp 719-730, 1976

Michon, J A, "Concerning the Time Sense: The Seven Pillars of Time Psychology", in

Psychologica Belgica, 33(2), pp 329-345, 1993

Parker, H E D, Roast C R and Siddiqi J I, "Towards a Framework for Investigating Temporal Properties in Interaction", UCRE Workshop paper, Glasgow, 1996

Pedri, S and Hesketh, B, "Time Perception: Effects of Task Speed and Delay", Perceptual and Motor Skills 76, 599-608, 1995

Roast, C R, "Modelling interaction with distributed information - so much information and so little time", Colloquium of the IEE (Computing and Control Division), Digest No. 95/223, 1995

Schank, R C, "Dynamic Memory: a theory of reminding and learning in computers and people", Cambridge University Press, 1982

Vroon 1970 - full reference details to be supplied - cited in Block, above

Usability Inspection and Evaluation of Interactive Information Retrieval Systems

Mark Treglown and Janet E. Finlay
hci@hud - The Human-Computer Interaction Research Centre,
School of Computing and Mathematics, University of Huddersfield,
Queensgate, Huddersfield, Yorkshire. HD1 3DH
Tel: +44 (0) 1484 472974 Fax: +44 (0) 1484 421106 Email: scommt@zeus.hud.ac.uk

1. Introduction

While the need to develop improved user interfaces to databases has been recognized by the human-computer interaction (HCI) and information retrieval (IR) communities for some time, work has tended to be conducted at the margins of the two communities. The HCI and IR research communities are now, however, taking note of the concerns of the other (as this workshop shows), and plans for productive areas of research are starting to be proposed. We report in this paper on the first fifteen months of the two-year "HIBROWSE for Bibliographic Databases" project. This project continues over five years work on the development of novel graphical user interfaces to bibliographic databases. The project seeks to explore the design and development process of user interfaces to databases and attempts to introduce appropriate methods into the usability evaluation and re-design of an existing database system.

The system we are considering is the HIBROWSE user interface (Pollitt et al, 1994) applied to the EMBASE medical bibliographic database, which grows by approximately 700,000 records each year. HIBROWSE presents the user with a number of views onto the database, each view is presented in a window and lists the contents of a single field of the database. Users can, by clicking on terms and values, restrict the contents of views to the dates of publication, journals, authors, thesaurus terms used to index and retrieve articles, and so on, to the set they are interested in. They can then create *refined views* where every view lists details about only the articles constrained by the user's previous selections. We shall not consider HIBROWSE any further except where discussion of the usability evaluation methods employed during the project requires it. During the course of the project so far, we have attempted to employ two different approaches to the development of the HIBROWSE for EMBASE system. The first approach is the highly structured MUSiC suite of tools and techniques, the second is to employ discount usability engineering techniques and to employ less formal methods for evaluating system usability when a suitable context arises.

2. Applying a Structured Approach to Usability Evaluation

It was felt in the early stages of the HIBROWSE project that a structured approach to applying usability methods in design would be appropriate; for this reason a number of tools from the MUSiC suite of tools and techniques were acquired. MUSiC is the result of the ESPRIT II project 5429 - MUSiC (Measuring Usability of Systems in Context) project, and provides a number of methods and tools for providing quantitative measurements of system usability with regard to metrics, and of subjective usability. MUSiC is strongly influenced by the ISO 9241 definitions in that usability should be thought of as "the quality of use **of** an interactive system **by** its (intended) users **for** achieving specific work goals and tasks **in** particular work environments" (Macleod, 1994).

The first tool from the MUSiC suite to be examined was the Usability Context Questionnaire. This detailed, formal questionnaire is employed to determine the context of use of a product, its intended users, the physical environment in which it will be used, the major tasks the system will be required to support and relevant aspects of the users' work processes that will affect the way in which users are able to carry out tasks when using the system. It is intended that much of the context questionnaire is completed during a meeting of relevant stakeholders in the development of the system; the designers, marketing personnel, managers, and representatives of intended user types. A completed context questionnaire is used to prepare metrics of users' expected performance when using the final system during usability testing. It is also used to define a context in which testing should be performed and to suggest tasks that representative users should perform during formal usability testing.

When considering IR systems, and medical IR in the current project, applying this tool becomes a difficult task, and one that has not been satisfactorily completed to date in the project. The user population of this IR system is any and all people who have need to seek information from articles indexed in a bibliographic database. It is clear that assumptions have to be made and imposed on the system design by the stakeholders and design team if a context questionnaire is to be completed. If not, considerable market surveying of users of the database is required. It is useful if the developers and maintainers of a bibliographic database, who will be stakeholders in the design team, have performed appropriate market research. A problem that faces the current project is that no such market research data is available. By making assumptions, however, other aspects of the questionnaire may be completed. The range of physical environments in which an IR system will be used, for example, can be described following field visits to typical locations. Such visits have been undertaken during the project.

The requirement by the questionnaire to classify user groups and describe typical attributes and abilities of users in each group again requires considerable market research. It is unclear whether constructing classes of users and describing typical users within each class is relevant in the IR domain when the users' typical tasks are considered. The simple descriptions of users' major tasks required for completion of the context questionnaire will not capture the range of information that users will require of the system, the complexity of the query that they are attempting to construct, the complexity of the view they are refining to, or the plan or method they follow to obtain the information that they require. We find ourselves in agreement with Belkin and Henniger (1996) that the variety of cognitive processes underlying the IR process should be understood in order to influence design. Such information is not obtained from a superficial reading of the maxim 'know the user'; we are currently undertaking a study of the roles of task analysis methods in IR design which may reveal more useful input to design.

Context questionnaires are used to generate metrics of required or expected user performance. It has not proved possible to generate such metrics without comparative data from studies of existing bibliographic IR systems. The issue of developing usability metrics, especially those that can be calculated using other tools from the MUSiC suite, is complicated when examining IR systems by the linking of measures of usability with measures of the effectiveness of IR systems. For example, measures such as the quality and quantity of users' work depend not only on the user interface but the retrieval mechanism if they are to be properly calculated.

We are currently planning comparative formal usability evaluation trials at a large teaching hospital from which we hope to understand how such evaluation methods can help produce data relevant to improving the design of an interactive IR system. We also expect to gain some understanding of the process of creating scenarios for evaluation, and the criteria by which the quality of users' efforts at information retrieval can be assessed. These plans depend to a considerable part on the expertise of those who train others in branches of medical practice, and who are able to provide expertise on what areas of knowledge IR systems should support, and what are the IR skills that medical practitioners should possess. In drawing up these plans, we recognise that domain knowledge is important in being able to judge the usability of even fairly general-purpose IR tools.

We have also considered other structured lifecycle models proposed for system development by HCI researchers. The "STAR" development lifecycle model, for example, proposed by Hix and Hartson (1993), stresses initial analysis of extant user's tasks and work processes leading to the development of a requirement specification given in the semi-formal User Action Notation (UAN). Development then proceeds through iterative cycles of development of prototypes based on the specification, formative evaluation of the prototypes, and re-specification based on the results of evaluation. In order that a common design artifact could be referred to by all members of the project team, a UAN specification was reverse-engineered from published descriptions of the HIBROWSE user interface (Treglown, 1996).

This effort proved useful as it revealed that many aspects of the design, and a number of interface features necessary to perform tasks that the developers intended that it support were under- or un-specified. While it was found that the UAN was capable of describing this interactive IR system, the description produced describes interaction with the system at a low-level of detail, with user interface widgets, whereas larger-grained tasks are not considered by the specification. It has not proved possible to attempt to follow the STAR lifecycle and evaluate it as a possible approach to developing IR systems due to the time required to use the software tools and database management system employed by the project team.

3. Applying Discount Usability Methods

Although MUSiC provides a suite of tools that can be applied during a notional development lifecycle, quantitative measurement of comparative and subjective usability are obtained from laboratory-based summative testing once a prototype has been developed. Such an approach could be accused of the traditional flaw of usability evaluation methods, that they offer too little too late. This is tempered within MUSiC by concentration early in design on determining the context in which a system will be used and the tasks that users will carry out using the system which form the basis of system functionality. Paton et al (1994), in their discussion of their experience of using a number of evaluation methods while designing IR systems, reject quantitative evaluation methods for reasons of cost, and the expenditure of other resources, and the perceived difficulty of employing quantitative results in redesign. It should be noted however that other authors (for example Henniger & Belkin, 1996) from the IR community value quantitative evaluation to determine measures such as effectiveness of recall and expressiveness of the querying mechanism.

Instead, Paton et al (op cit.) advocate the use of qualitative methods which, they judge, provide greater input into redesign. This assessment reflects conclusions by the HCI community that usability must overcome prejudice and prove itself cost-effective (Bias & Mayhew, 1994), and that usability inspection (Nielsen & Mack, 1994) and usability engineering (Nielsen, 1993) methods are of considerable benefit. That some of these methods can be of benefit has been supported by work conducted as part of the "HIBROWSE for Bibliographic Databases" project.

A formative, usability inspection method that proved successful in influencing redesign (Treglown & Finlay, 1996) was the cognitive walkthrough method. This method has achieved a high level of popularity, and has proved capable of being learned by software designers previously untrained in HCI and design for usability. The simplified cognitive walkthrough method (Wharton et al, 1994) was taught to the project team in a seminar lasting two hours. The designers then performed a walkthrough on a simple task (resizing an on-screen window) which lasted approximately an hour. A full walkthrough was performed which took approximately 2 $^1/2$ hours to complete. As other case studies of employing the cognitive walkthrough have noted the difficulty of generating task scenarios on which to base walkthroughs, the walkthrough reported in (Treglown & Finlay, 1996) used a scenario observed in a field study of a librarian (and frequent search intermediary) at a large university medical library. As the only usability evaluation of HIBROWSE had been conducted on a previous version employed as the user interface to a far smaller and simpler database a year before the start of the current project, many usability problems were revealed when an initial walkthrough was performed, some quite severe. Treglown and Finlay (1996) should be consulted for details of the design changes undertaken following the walkthrough.

4. Conclusions and Design Proposals

It is clear that usability input into the design and development process has been less than ideal. By applying usability methods late in design evolution, methods will necessarily be employed in an ad hoc manner. Despite this, we have found that discount usability methods can be used to provide a real influence on system design, into the design process and on designers' thinking about the needs of the user. Work to be conducted over the remaining months of the project will explore the results of conducting formal quantitative usability evaluation during system development and whether much of the information usability analysts (as well as those marketing a system) would like to have, such as comparative effectiveness and task efficiency between systems, can be obtained late in system development.

From our experiences, we suggest the following issues need to be addressed in IR system development:

- task analysis and understanding users' transient needs for information about the database and the system state while following a particular search process is required.
- the problem of collecting scenarios for design, metric creation and planning usability evaluation when considering general IR tools needs to be better understood.
- software tools that permit rapid prototyping both of complex user interfaces and database access and retrieval algorithms in an iterative development lifecycle are essential, unless resources exist to conduct the sort of market research and extant task analysis that would permit the use of a more structured development lifecycle model.

Acknowledgments

The first author and "HIBROWSE for Bibliographic Databases" project are supported by The British Library research grant RDD/G/282. We gratefully acknowledge additional sponsorship from Sun Microsystems Inc., Elsevier Science B.V., Software AG and Elonex. EMBASE and EMTREE are registered trademarks of Elsevier Science B. V. HIBROWSE is a registered trademark of The University of Huddersfield HEC.

References

Bias, R. G. & Mayhew, D. J. [eds.] (1994) *Cost-Justifying Usability*. Academic Press, London.

Henniger, S. & Belkin, N. (1996) *Interface Issues and Interaction Strategies for Information Retrieval Systems*. CHI'96 Tutorial notes.

Hix, D. & Hartson, H. R. (1993) *Developing User Interfaces: Ensuring Usability Through Product and Process.* John Wiley and Sons, Chichester.

Macleod, M. (1994) Usability: Practical Methods for Testing and Improvement. In *Proceedings of the Software'94 Conference,* (Sandvika, Norway, 1-4 February 1994), Norwegian Computer Society.

Nielsen, J. (1993) *Usability Engineering*. Academic Press, London.

Nielsen, J. & Mack, R. L. [eds.] (1994) *Usability Inspection Methods*, John Wiley & Sons, New York.

Paton, N., al-Qaimari, G., Doan, K. & Kilgour, A. C. (1994) Techniques for the Effective Evaluation of Database Interfaces. In P. Sawyer (ed.) *Interfaces to Database Systems (IDS94) Proceedings of the Second International Workshop on Interfaces to Database Systems (Lancaster, 13-15 July 1994),* Springer-Verlag, Berlin.

Pollitt, A. S., Ellis, G. P. & Smith, M. P. (1994) HIBROWSE for Bibliographic Databases. *Journal of Information Science*, 20, 413-426.

Treglown, M. (1996) *A Specification of the HIBROWSE for EMBASE System given in the User Action Notation*. University of Huddersfield, School of Computing and Mathematics Research Report RR9601.

Treglown, M. & Finlay, J. E. (1996) Driving Design with the Cognitive Walkthrough. In A. Blandford & H. Thimbleby [eds.] *Ancillary Proceedings of the 11th BCS Conference on Human-Computer Interaction - HCI'96* (London, 20-23 August 1996).

Wharton, C., Rieman, J., Lewis, C. and Polson, P. (1994) The Cognitive Walkthrough Method: A Practitioner's Guide. In J. Nielsen & R. L. Mack (1994).

Cognitive Aspects of Information-Seeking During Complex Learning Tasks: a comparison of book and CD-ROM-based versions of a medical textbook.

Joe Nicholls, Mark Howes* and Rick Jones ~
Media Resources Centre, University of Wales College of Medicine.
* *Department of Psychology, University of Leeds.*
~ *Chemical Pathology, School of Medicine, University of Leeds.*

1. Introduction

Like much computer software, full-text retrieval systems are marketed at ‘end-users’ whilst little is known about the psychology of their intended use. Major expository texts, such as the Oxford Textbook of Medicine†, are now being made available in both electronic (CD-ROM) and paper versions. Fundamental issues concern the appropriateness of on-line expository texts for the purposes of study and learning. Usability issues are particularly pertinent in the context of reading-to-learn tasks1 and in education as a whole2. Critical comparative evaluation of such interfaces is required, but existing theoretical frameworks and methodologies have been found to be lacking. Electronic texts need to be shown to be at least comparative or better than their paper counterparts if they are to be used. Although reading from paper has been found to be both faster and less fatiguing than reading from computer screens3, attention should also be focused on the ability of each medium to support the location of required text.

Earlier studies, by the authors, showed that access mechanisms and main-body text of electronic and paper books were used differently4,5: a higher proportion of time and more frequent interactions with the access structures when using the electronic version. Subsequent studies focused on cognitive activity during textbook use to solve complex learning tasks. This paper addresses these issues by describing a new information-seeking (IS) framework, and the results of its application to the analysis of verbal protocols from users of paper and electronic textbooks.

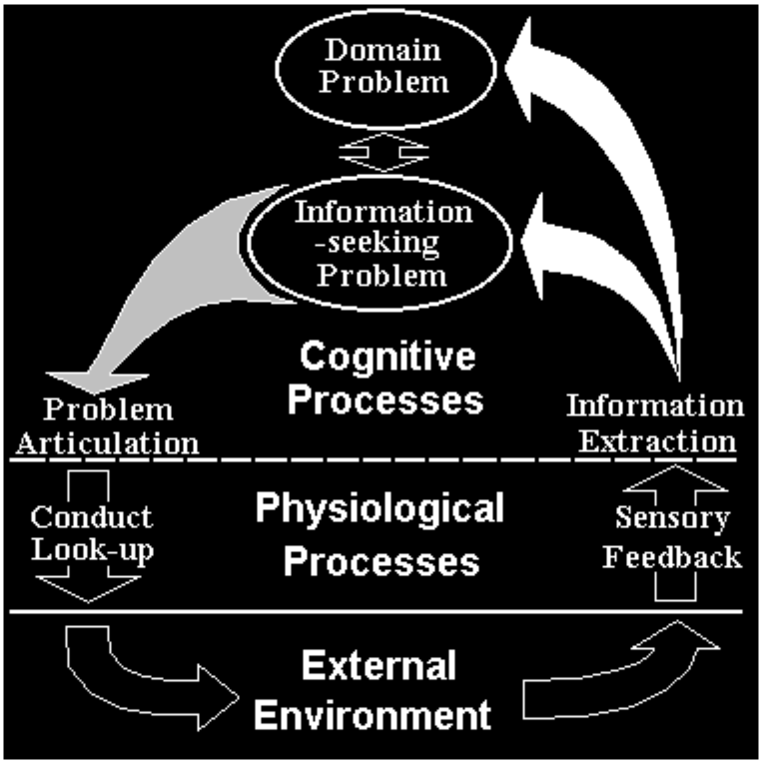


Figure 1. Information-seeking framework.

2. Cognitive Framework

† Oxford University Press

To facilitate the comparison of resource use, an IS framework was derived from a critical analysis of existing literature6-9 and subjects' verbal protocols. A generic IS framework must be able to account for control processes that influence both visual and physical search behaviour. Therefore, the framework was designed to accommodate activities associated with the use of both static and dynamic resources (see Figure 1). Of particular significance is the separation of two interrelated concurrent problems: the *Domain Problem* (DP) and the *Information-seeking Problem* (ISP). The first concerns the reason(s) for seeking information, and the latter, the means by which it may be located and retrieved. The separation allows for the representation of information and strategic activity in relation to each problem.

The whole process is instigated from consideration of the DP, which, if external information is required, gives rise to the ISP. *Problem Articulation* (PA) refers to cognitive activities concerned with the expression of information requirements to the external world, i.e., how to use access structures. Behaviour occurs as a direct result of PA. *Conduct Look-up* (CL) represents actions designed to produce alterations in perceived data; this may or may not result in changes to the external environment, e.g., directing visual gaze compared to pressing keyboard keys. As a result of CL, different data from the external environment are perceived. *Information Extraction* (IE) processes provide information applicable to each problem space. The framework was used to structure the analysis of transcripts, and was in turn modified in the light of findings.

3. Resource Description

The content of the electronic version of the Oxford Textbook of Medicine was an exact copy of the original paper book, except for the absence of all tables, figures and pictures. Instead of the book's contents list and index, the CD-ROM interface provided a Boolean-based search tool which incorporated proximity operators and index specifiers (e.g., main-body text, headings, authors, etc.). Successful queries modified the contents list with frequencies of occurrence of the search terms used, and also highlighted them in the main-body text. Access could also be achieved via an unmodified contents list with an expandable hierarchy of headings. The keyboard was used for typing text, menu commands, and function keys.

4. Methodology

Six pre-clinical medical students were asked to seek text-based information in order to answer 10 ill-defined clinical problems. Five questions were answered using the on-line CD-ROM textbook and five using the paper book equivalent. Responses were gathered from each student over a number of sessions. Audio-visual recordings were made, and for the electronic textbook a 'terminate-stay-resident' program captured keyboard use. The students were asked to think-aloud whilst performing each task. From the 60 video recordings, verbal protocols, actions and resource state changes were transcribed for further analysis.

Subsequent analysis concentrated on access structure use, as opposed to reading of the main-body text. Comments made by students were categorised according to the five components identified in the cognitive framework: DP, ISP, PA, CL, and IE (see above). Independent rater categorisation of twenty of the transcripts achieved 87% agreement. Primary interest was directed at identifying search strategies employed for each resource and their interface dependencies. Although examination of these strategies is ongoing, an interesting picture is emerging which reveals both advantages and disadvantages for each medium in the context of these tasks.

5. Results and Discussion

Domain knowledge : Task relevant vocabulary proved to be a key factor in determining usability. This was especially so when using the CD-ROM's query facility, but also featured whilst searching the book's index and contents list. Limited and poorly structured clinical knowledge meant that students struggled to generate appropriate search terms. Other than keywords derived from the question, very few novel words/phrases were used. Recognition-based approaches to search were strongly evident, demonstrated by scanning contents list and index entries. Interestingly, the identification of relevant entries often prompted reconsideration of the domain problem prior to text entry.

The respondent in Figure 2 shows a clear preference for starting with the contents list as opposed to the index. A section was immediately recognised as potentially relevant [4-6]. This was scanned [7] with the aim of orientating themselves to the topic area [8] and identifying relevant headings [10]. A rapid assessment was made of the extent of the search task [12-15] and their inability to extract useful information [13]. An alternative, more direct keyword search using the index [16] was chosen to save time [15]. The problem arising from poor domain knowledge is evident [13], in this case making even the recognition-based strategy ineffective.

Book : What are the characteristics of mitral regurgitation ?

4	I've just turned to cardiovascular disease	11	Heart failure ...
5	What section was that ?	12	it's a big section. ... uhm ...
6	Section thirteen.	13	Not really ... cause I don't really know what I'm looking at.

7	I'll just have a quick look down.	14	So, I think, to save myself from going through all that,
8	So I can get my bearings.	15	which would take ages,
9	Section thirteen ... cardiovascular disease.	16	I'll just look up mitral in the back.
10	Just seeing if there's anything relevant.		

Figure 2. Shows the use of the book contents list and redirection of search.

Strategy shifts : The transition from browse oriented search, for example of the contents list, to a more focused form of search, using the index, was often observed (see Figures 2 and 3). On occasions this occurred quite quickly: for example, the rapid scanning of sub-indices or headings following location of a relevant main entry. However, each resource differed markedly in its ability to support such strategies and transitions. Preference was shown for low cognitive load activities, such as visual scanning, and recognition. As might be expected, the book was better for visual scanning, whilst the CD-ROM was more appropriate for focused, structured searches. The ability to shift from one form of search to another was not as well supported by the CD-ROM interface either.

CD-ROM : What are the characteristics of mitral regurgitation ?

2	Right, well the first thing ...	6	So, I don't know whether going straight in at the term
3	I guess I'm gonna look for ...	7	in the title is going to be too detailed.
4	Is a section with some overview of	8	Can I go into the contents ? ...
5	heart valve defects.	9	and scan the contents ?

Figure 3. Illustrates an attempt establish overview and strategy consideration.

Goal specificity : A dilemma expressed by all participants, was the need to obtain very specific information to answer the question, whilst at the same time requiring sufficient background information to help integration with existing knowledge, and hence to direct search. As a result, search goals varied in their specificity, with attempts made to retrieve both detailed and general information for the same task (see Figure 3). Consequently, strategies were found to be both adaptive and dynamic. The CD-ROM interface did not facilitate retrieval of general information as easily as the book. Search via the query facility required the precise specification of search terms [6-7]. The results of this produced highlighted search terms in the main-body text, which were visited sequentially or selected via headings in the contents list. The common outcome was that highlighted words were frequently seen in a 'sea of text', often without any headings on screen and some distance from the start of the section. Participants directed much effort and attention at establishing context and orienting themselves, as well as seeking the introductory text to the section. In contrast, using the book, it was possible to rapidly scan over the index, contents list and text virtually all at the same time, thereby easily obtaining a range of information at different levels of detail, of relevance to both the information-seeking task and the domain problem.

Management of search goals : Of particular interest, was the participants' deliberate management and prioritisation of multiple search goals. Resource and domain knowledge were major factors influencing the order. In the first extract below (see Figure 4), the focus of the search was determined by consideration of two areas of domain task knowledge [6]. The second extract shows how a form of 'search agenda' was often created. In this example, headings in the contents list of the CD-ROM were identified as candidate text entry points [25], but rather than immediately entering the text, search of the contents list continued [26]. Resource knowledge [27] and information applicable to the ISP derived from the resource [28] were important factors influencing this decision. In addition, serendipitous encounters with unexpectedly relevant information often prompted the suspension of a line of inquiry and a change of direction. This illustrates both the active consideration of alternative approaches to the domain and information-seeking problems, and the need for interfaces that accommodate flexible management of multiple search paths.

Book : Name the common physical signs that co-occur in Ulcerative colitis and Rheumatoid arthritis.

1	Right, I think the best way to do it is to look up ... uhm ...	5	then go on to rheumatoid arthritis.
2 :	I'll look up colitis and then ...	6	Because I'm sure I know more about rheumatoi arthritis than I know about ulcerative colitis.

CD-ROM : What are the commonest causes of blindness in the western world ?

25	Uhm!...well that's...like to look at that.	27	So, I can go back to that anyway, can't I ?

26	Well, I want to just carry on looking at the others as well.	28	‘cause there aren’t that many, are there ?

Figure 4. The prioritisation and suspension of search.

Contextual orientation : A very important issue was the capability of each interface to help participants establish and maintain context. They directed considerable effort in orienting themselves with respect to the ISP and DP. Following extraction of new information from the resource, attempts were made to establish its significance in relation to both problems. For example, relationships implied by the hierarchical structure of entries in the contents list and index, and between headings and main-body text, conveyed important contextual information. The ability of the interface to convey such information to help users establish context is a crucial factor in promoting usability. In this respect, the book provided information that was more timely and appropriate, and more easily accessed, compared to the CD-ROM interface (see Figure 5).

Book : What are the commonest causes of blindness in the western world ?

11	See where chapter 23 is.	13	See what chapter 23 is actually.
12	It’ll be in this one. leaves open at index, reaches for other volume, open to contents list.	14	Oh, right, disorders of the eye ... makes sense.

Figure 5. Broadening of context using the book.

6. Conclusions

Central to improving interfaces to full-text resources is the need to increase our understanding of the cognitive processes involved. This study has shown that users employ subtle and complex methods of interacting with such resources, of which designers are generally not aware. Information-seeking must be recognised as a significant component of textbook usage, and in terms of resolving the domain problem the process of locating to text may be just as important as reading it.

Results thus far, indicate that there are features of both interfaces which foster productive, as well as poor, search strategies. The challenge is to understand how different access structures facilitate location and retrieval with the minimum of cognitive effort, whilst at the same time maximising cognitive resources and information for the DP.

Participants’ verbal protocols continue to be a source of very rich data. Ongoing analysis is aimed at elaborating and clarifying the I-S framework, and informing the design future of full-text retrieval interfaces.

References

1 Wright, P. (1968) Reading to learn. *Chemistry in Britain*, 4, 445- 450.

2 Laurillard, D. (1993) *Rethinking University Teaching: a framework for the effective use of educational technology*. Routledge.

3 Gould, J., and Grischkowsky, N. (1984) Doing the same work with hardcopy and with CRT terminals. *Human Factors*. 26 (3), pp. 323-337.

4 Nicholls, J.A. (1994) Text Search : Information-seeking strategies using paper and CD-ROM versions of a medical textbook. *PhD Thesis*, University of Leeds.

5 Nicholls, J., Howes, M., and Jones, R. (1995) Information-seeking behaviour using paper and electronic versions of a textbook. *IEE Colloquium : Human-Computer Interface Design for Multimedia Electronic Books*. Digest No. 1995/038.

6 Bates, M.J. (1991) The Berry-Picking Search: User Interface Design. In: Dillon, M. (ed.) *Interfaces for information retrieval and online systems: The state of the art*. Greenwood Press.

7 Guthrie, J.T. (1988) Locating information in documents: Examination of a cognitive model. *Reading research quarterly,* Vol. 23, pp. 178-199.

8 Marchionini, G. (1992) Interfaces for End-User Information Seeking. *Journal of the American Society for Information Science*, Vol. 43 (2), pp. 156-163.

9 Norman, D.A. (1986) Cognitive Engineering. in: Norman, D. and Draper, S. (eds.) *User Centred Systems Design: New Perspectives on Human-Computer Interaction.* Lawrence Erlbaum Associates. pp. 31-62 .

Interfaces to support collaboration in information retrieval

Michael B. Twidale & David M. Nichols
[mbt | dmn]@comp.lancs.ac.uk
Computing Department, Lancaster University
http://www.comp.lancs.ac.uk/computing/research/cseg/projects/ariadne/

1. Introduction

> … support for communication and collaboration is as important as support for information-seeking activities and that, indeed, support for the former is needed to support the latter.
>
> [Levy and Marshall, 1995]

Interfaces to information retrieval (IR) systems increasingly support progressive interactive search formulation and refinement. Part of their aim is to make information retrieval easier for the end user without the need for (or with minimal support from) an information intermediary. One approach to supporting these activities has been to design 'intelligent' interfaces that embody some of the knowledge and functions of a human intermediary (e.g. [Belkin and Marchetti, 1990]). An alternative, complementary, approach is to build an interface that actively supports interaction with an intermediary or other user: expanding the 'interactive' nature of the IR system to include other human resources [Hoppe and Zhao, 1994; Twidale et al, 1995].

This social approach, taking direct advantage of human intermediaries' knowledge, in addition to any smart underlying functionality, such as intelligent agents, is clearly more robust and adaptable than solely constructing an 'intelligent' stand-alone interface. It should be generalisable across databases, reactive to system change and likely to become more common as asynchronous interactions dominate accesses to remote digital libraries [Twidale and Nichols, 1996].

However, collaboration has its costs; someone has to pay for the interaction with an expert user, and with large numbers of searchers, help from an expert may be hard to get. Thus it is incumbent on us in advocating collaboration to investigate how it can be supported cost effectively. This will involve minimising the amount of the expert's time that a help session will take (by for example giving the expert a better understanding of the context of the query) and providing mechanisms for the more efficient delivery of the advice constructed. In addition, we wish to investigate how users can minimise their recourse to experts by facilitating the incremental learning of search skills, and by learning from their peers.

In this paper we describe the requirements of interfaces to support these collaborative interactions and our first approach to implementing a system that embodies one kind of explicit support for collaboration, a graphical representation of the search process.

2. Collaboration in Information Retrieval

Collaboration already occurs in current physical libraries. However, in addition to the much studied reference interview with a skilled intermediary, substantial informal collaboration occurs. We have observed that many students at Lancaster learn how to use the online public access catalogue (OPAC) from their peers, leaning across adjacent terminals or clustering around a single terminal, engaged in synchronous co-located collaborative learning and working [Twidale, Nichols and Paice, 1996].

This collaboration occurs despite rather than because of the systems, which are conventionally designed on the assumption that users work alone. The collaboration is also desirable, leading to effective learning and reducing the load on information intermediaries. If we are to advocate the development of systems that support collaboration rather than automation, then we must be concerned with supporting cost-effective collaboration – allowing users to be able to quickly explain to each other what they have already managed to do and what they now need help on.

The nature of collaborative IR activities will change with the trend towards remote digital libraries and interfaces should support both existing (under-acknowledged) forms of collaboration and the new kinds of collaborative activities opened up by networking.

We are studying the nature of collaboration as it currently occurs in conventional libraries to give us a starting point for developing more advanced systems. To date we have mostly looked at the collaboration that occurs between users, chiefly because this has in the past received far less attention than the expert-novice style interaction of the conventional reference interview. Future work will also look at collaboration between experts such as library staff working together on a user's problem.

3. Interface Design

A carefully designed interface offers the potential to provide its users with a visualisation of concepts that are normally only available as abstract ideas. This reification process offers great advantages in facilitating both learning about the concept and cooperative working using the concept.

We undertook a series of preliminary studies [Twidale, Nichols and Paice, 1996] of synchronous remote help giving using a combination of rudimentary tools including Unix talk. It became clear that a crucial need was the ability to share not only the search *product* (the hits obtained) but also the search *process* (how those hits were obtained). Observations of the library's help desk backed up this observation; not surprisingly when confronted by users claiming to be unable to find something, librarians frequently ask what they have tried. Equally unsurprisingly, users find this a difficult question to answer: it is very difficult to remember the search actions when one is focusing on the search goals, there is a natural human inclination to auto-correct, remembering what one intended to type rather than what one did, and naturally those most often asking for help are usually the ones who understand the system least and so are most likely to get confused and to lack a specialist vocabulary for describing their actions.

Thus even in a face-to-face help interaction, it is useful to have an interface that enables participants to share an exchangeable representation of the user’s search, and indeed perhaps of an expert's proposed solution. Where the user and intermediary are separated by distance and time, such an interface becomes even more important. Although a common representation is needed to aid communication, an expert may benefit from an additional, more powerful view of the information.

One method to achieve this commonality is to use a video-link between users – typically between end-user and library staff [Sugimoto et al, 1995]. However, video is not appropriate for many collaborative activities as it fails to make the search a discrete manipulable object. The C-TORI model of (synchronous or near-synchronous) cooperative IR provides facilities for sharing query histories, cooperatively browsing results and cooperative query formulation [Hoppe and Zhao, 1994]. Nevertheless, the costs and inconvenience of synchronising help or collaboration are likely to override the advantages of its immediacy.

As the search needs to be preserved for communication purposes it can also be used as a visualisation of activity for the end-user. Once a system supports the creation of such an interaction history, it can used in many ways [Hoppe and Zhao, 1994]. Single-user uses include: reuse, error recovery, navigation, reminding and user modelling [Lee, 1992]. Providing a complete record of a search activity frees users from having to remember low-level goalstack details and enables them to concentrate on more strategic elements of the search process. As a digital object it can also be communicated to other people - serving as the basis for a much wider range of collaborative activities [Twidale, Nichols and Paice, 1996]. Such a search object can be stored, high-lighted, edited, annotated, replayed and re-executed: typically the kinds of activities that occur in tutorial and help-giving interactions [Lemaire and Moore, 1994].

4. The Ariadne Interface

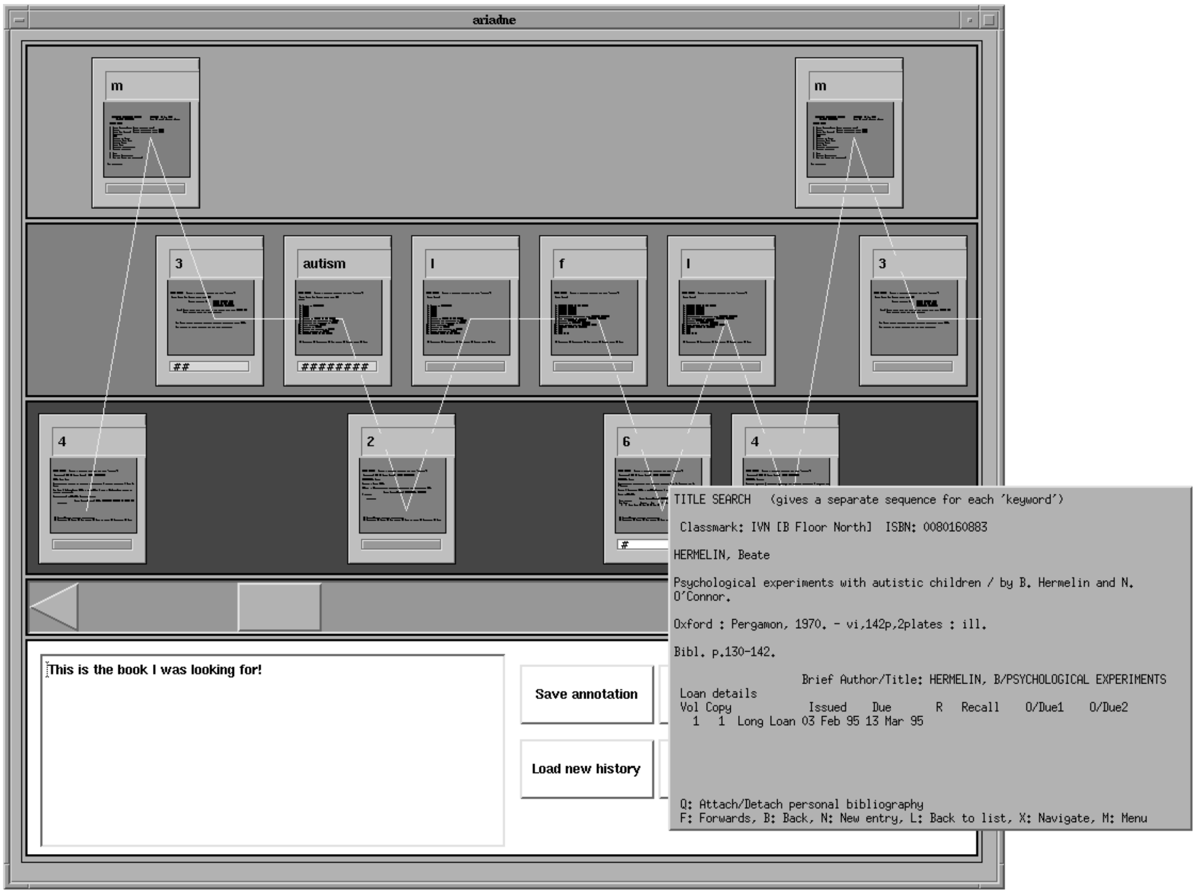


Figure 1 The Ariadne interface

We have developed a prototype interface, Ariadne, which attempts to support these activities. Ariadne runs on top of an OPAC or online database, recording the queries and their results and then producing a visualisation of the search process that can be reflected on, shared and discussed by interested parties.

A search is represented as a sequence of cards containing in thumbnail outline the input search command and the output result screen. Cards can be explained by clicking on them. The scrollable sequence of cards can be read left to right as a history of the entire search activity. Additional qualitative information is provided by the vertical position of a card in one of three levels; a high level activity (such as menu selections from the top part of a system's menu hierarchy) is placed on the top level, search queries are placed in the middle level and actions involving the viewing of results on the bottom level. The result is that it is possible to obtain at-a-glance impressions of certain characteristic patterns of behaviour such as the common novice 'error' (dependant upon circumstances) of composing a search that yields hundreds of hits and then proceeding to read through them all. As might be expected, a search session can yield a substantial trace and so facilities are provided to elide sub-sequences of cards, replacing them by a single group card. This card and the individual cards can be annotated, either by the originator of the search or by others commenting upon it.

A main advantage of such an interface is that discussions (both remote and co-located) can then be undertaken between participants about sophisticated searching techniques without the need to learn the specialised vocabulary of the information profession. By contrast, instead of using the abstract conceptual terminology, the participants can point to, or highlight, the particular instances that were of significance in the search under discussion.

5. Conclusion

We have proposed the development of interfaces for information retrieval systems that are not merely oriented to the delivery of the results of a search, but attempt to support collaboration with other users. These others may be expert intermediaries, co-workers, peer learners, etc. One way of supporting collaboration is to provide a visualisation of the search process that can be manipulated and discussed by the users. We believe that support for cooperative use of an IR system can improve the learning and understanding both of that system and of generic search skills. In addition, by acknowledging the importance of other people in the search process, it leads to a more robust search process, more able to cope with failures of whatever underlying sophisticated intelligent search mechanisms it may additionally use.

Acknowledgements

Work on the Ariadne project has been funded by the JISC New Technologies Initiative of the HEFCs, and by the British Library Research and Innovation Centre.

References

Belkin, N.J. & Marchetti, P.G. (1990), Determining the functionality and features of a intelligent interface to an information retrieval system, *Proceedings of the the 13th International Conference on Research and Development in Information Retrieval (SIGIR'90)*, Brussels, Belgium, ACM.

Hoppe, H.U. & Zhao, J. (1994), C-TORI: an interface for cooperative database retrieval, in Karagiannis, D. (Ed.), *Database and Expert Systems Applications*, Berlin: Springer-Verlag, 103-13.

Lee, A. (1992), *Investigations into History Tools for User Support*, PhD Thesis, Department of Computer Science, University of Toronto.

Lemaire, B. & Moore, J. (1994), An improved interface for tutorial dialogues: browsing a visual dialogue history, *Proceedings of the Conference on Human Factors in Computing Systems. CHI'94*, Boston, MA, ACM, 16-22.

Levy, D.M. & Marshall, C.C. (1995), Going digital: a look at assumptions underlying digital libraries, *Communications of the ACM* 38(4), 77-84.

Sommerville, I., Rodden, T., Sawyer, P., Bentley, R. & Twidale, M.B. (1993), Integrating ethnography into the requirements engineering process, *Proceedings 1st International Conference on Requirements Engineering*, San Diego, USA, IEEE Press, 165-74.

Sugimoto, S., Gotou, S., Zhao, Y., Sakaguchi, T. & Tabata, K. (1995), Enhancing usability of network-based library information system – experimental studies of a user interface for OPAC and of a collaboration tool for library services, *Proceedings of Digital Libraries '95*, Austin, TX,

Twidale, M.B. & Nichols, D.M. (1996 to appear), Collaboration and visualisation of the search process, *Aslib Proceedings*.

Twidale, M.B., Nichols, D.M. & Paice, C.D. (1996), *Browsing is a Collaborative Process*, Technical Report CSEG/1/96, Computing Department, Lancaster University.

Twidale, M.B., Nichols, D.M., Smith, G. & Trevor, J. (1995), Supporting collaborative learning during information searching, *Proceedings of Computer Support for Collaborative Learning (CSCL 95)*, Bloomington, Indiana, 367-74.

Providing Expert Assistance with Information Retrieval in the Virtual Library

Rob Procter, Lorna Banks and Kenny MacDonald,
Department of Computer Science, Edinburgh University

Elisabeth Davenport,
Department of Communication and Information Studies, Queen Margaret College, Edinburgh

Peter Burnhill and Sheila Cannell,
Edinburgh University Library

1. Introduction

The increasingly widespread use of on-line information retrieval (IR) systems represents for many the first tentative steps in the realisation of the much heralded 'virtual library'. However, whilst people gain the convenience of IR from their desktops, at the same time their opportunities to consult with experts in IR to help resolve problems may be seriously impaired; if library users no longer visit libraries, the question is how best to bring librarians' expertise to users. We describe a project to develop a pilot multimedia service to support consultation between reference librarians and remotely-located users of bibliographic databases.

On the evidence of various studies, bibliographic database users presently face significant difficulties; for example, it is reported that between 30 and 50% of searches result in no 'hits' (Marchionini, 1995). These poor results may be attributable to various causes, ranging from poor user interfaces, to users' lack of knowledge of the domain and poor search strategies (Marchionini, 1995). Whilst better user interfaces for search engines will undoubtedly alleviate some of these problems, concentrating on this factor alone ignores the benefits of restoring and promoting a collaborative approach to IR.

2. Investigating Collaboration in IR

Historically, the reference consultation has been described as a question and answer session: "one of the most complex acts of human communication ... (in which) one person tries to describe for another person, not something he knows, but rather something he does not know." (Taylor, 1968). We have begun to investigate how support remote collaboration in IR by studying consultations between bibliographic database users and reference librarians in both conventional face-to-face circumstances and under conditions of computer-mediated communication (CMC). In classifying these, we have drawn on a range of prior studies which have analysed reference consultations within a framework of conventions that are observed 'inside' the reference consultation (Spink and Saracevic, 1993). Several researchers have described the on-line search process in terms of phases on the basis of analyses of transaction logs, verbal protocols, and qualitative studies of sense-making (Kuhlthau et al, 1992). Advice seeking and giving is manifest in some of these phases (induction; breakdown) more than others, as users derive their own heuristics to cope with the search space, which will conform, more or less, with the practices recommended by experts associated with the design and exploitation of that space.

Gaining a more complete picture of the nature of reference consultations involves:

- devising a taxonomy of the problems users typically encounter,
- analysing how problems and solutions are articulated, including the rules which operate within consultations, and
- the factors which circumscribe consultations or shape their practices.

To date, our data has come from reference librarians within the Edinburgh academic community and members of the EDINA user group, who have been beta testing the user interface for the BIOSIS service. The group is distributed across the UK, and requests for help come mostly by e-mail and telephone. The problems users pose can be broadly classified as practical questions such as:

- how do I do xy?
- is it possible to have the BIOSIS search guide on-line?

and questions of a more conceptual nature such as:

- is xy available -- can I do xy
- I couldn't get xy to work -- did I do something wrong?
- why is feature x on that part of the screen?

Our investigations indicate that where sustained consultations with a user occurs, 'rules' are 'followed' on the basis of the librarian's perceptions of the user's expertise and from the stage of the search, and that these are derived from the former's professional training. In many cases, however, the demands made of librarians are not amenable to treatment by means of rule-based consultations that relate to a subject area; clients' demands are not concerned with content, but with technical or spatial problems. Diaries kept by the librarians indicate that they are subject to a barrage of miscellaneous and heterogeneous questions, from users whom they do not know, in a pattern of interaction which is characterised by high volume, impersonal, uninformed, short exchanges. The dearth of sustained searches (and thus of sustained search consultations) reflects that fewer demands are made of librarians in a world where end-users tend to serve themselves. Two librarians stated that after induction or training sessions an a given database, users just disappear: "we don't know where they go". The content of consultations involves multiple 'starting overs' rather than sustained 'problem solving' reference work (Taylor, 1968).

Interviews with librarians also addressed the context, from the librarian's perspective, in which these consultations take place. All described full schedules (confirmed under observation and by the diaries), to have work distributed in space, and to parallel process at least some of the time (answering a phone query, for example, while working with paper, or dealing with e-mail; cataloguing, on-line searching and dealing with verbal enquiries at the reference desk). As we indicate above, few of them carried out extensive IR consultations of the sort described in the LIS literature, with face-to-face interaction over a workstation and finessing of feedback. Most of the collaborations logged in diary records were short (3 to 5 minutes) interactions of the 'help desk' kind, rather than subject searches, and followed a pattern of multiple tracking of user needs.

In order to elicit stories about their work and its context, we asked the librarians to describe how they handled multiple concurrent demands on their time: how did they multi-task? how might new media help with problem areas? All of the librarians did appear to apply criteria to structure their handling of complex demands. For example, some librarians appear to have established rules for priority: in a mediated interaction, a hierarchy appears to determine that visual demands have precedence over audio, and audio has precedence over text. One librarian observed that video, as a near equivalent of face-to-face, would have priority over phone and e-mail. Three participants expressed concern about exposure: they would condone the use of video (a 'demanding' medium) only where access to their expertise was managed by means of a timetable of 'office hours' or 'desk hours' arrangements, with notices at user workstations.

Other rules to accommodate multiple tasks in limited time appear to be based on a hierarchy of perceived user needs, and the level of skill required of an expert who attempts to satisfy them. The emergence of rules to manage the situation in which reference consultation is embedded is of particular importance for the design of remote consultation support systems. For example, the rules set by practitioners may allow for the matching of a 'demanding' medium with tasks where a high degree of specialist skill is required.

3. The Pilot System

A pilot system using various forms of multimedia tools, including videoconferencing, WWW, email and FAQ lists, has been implemented in order to investigate their roles in multimedia-based IR consultation[2].

We may take the face-to-face meeting to be the 'gold standard' for IR consultations. Face-to-face interaction is able to draw upon a rich variety of verbal and non-verbal contextual cues which make a significant contribution to communication structure and management. In addition, a shared, physical work context is important for the way in which it helps to sustain the common ground -- i.e. mutual understanding -- with the minimum of effort (Monk et al, 1994). For the design of a remote IR

[2]URL: http://www.remote.lib.ed.ac.uk

consultation service, however, the question is to what extent the full circumstances of face-to-face interaction need be recreated for it to be effective. Also, our study to date suggests that the design of a system to support remote consultation between librarian and user must acknowledge:

- the 'rules' which operate within consultations, and
- the circumstances which condition them or shape the practices within which collaborations occur.

The kinds of support provided should therefore reflect the variety of forms and styles of collaboration which arise as a result of the interplay of the above factors. In addition, given the limited availability of librarians and the various demands on their time, it is important to help users to help themselves by encouraging collaboration within the user community e.g. by making their knowledge and experience gained more easily available to others (Twidale and Nichols, 1996). In this context, resources such as FAQ lists are a particularly valuable low-cost means to accumulate information 'capital' and make it accessible on a user community-wide basis.

Videoconferencing represents the closest practical recreation of synchronous, face-to-face interaction currently available. Initial trials indicate, however, that for untutored users, the cheaper LAN-based systems (e.g. CUSeeMe) do not provide the quality of conferencing required[3]. In the pilot service, synchronous consultation is supported through audio and text-based conferencing with workspace sharing to enable user and librarian to jointly observe the results of searches, thereby provide the common ground for the consultation. This has the advantage of accessibility, requiring in its simplest form only a WWW browser. Given the patterns of interaction revealed in the investigation, it is clearly important that the service also provide various asynchronous forms of consultation. In this context, we are currently developing WWW-based tools for sharing bibliographic search histories asynchronously and for maintaining archives to support the automatic generation of FAQ-like search history lists from the flow of queries and answers.

These various forms of collaboration support have different characteristics and it is important for both overall service quality and cost-effectiveness to determine the circumstances in which the various options would be most appropriate -- for both librarian and user. Three service characteristics are especially relevant here: 'problem-resolving bandwidth' (capacity to articulate and sustain common ground), availability constraints (limits on initiation), and responsiveness (speed of handling once initiated). We can match the media against these characteristics.

It has been suggested that certain literary genres are deployed on the basis of 'rules' for interaction that are established within given working groups: a memo, for example, is used at junctures where time is short and familiarity is absent (Orlikowski and Yates, 1994). From this perspective, the face-to-face interaction may be seen as one of several genres which constitute the repertoire of 'reference work', each defined by appropriate rules which emerge in the working practice of a given institution.

4. Future Work

As experience is gained from running the pilot consultation service, we intend to elicit more information on how the demands of a given medium are perceived by both librarians and users, how judgements are made by the former on issues such as the prioritisation of tasks, and how the latter perceive its effectiveness for solving their problems. At the beginning, when understanding of the relevance and consequences of media characteristics is limited, patterns of use may be unpredictable, but we expect to see an 'ecology' of service use evolving over time as both librarians and users learn to match forms of consultation against their particular constraints and requirements.

Acknowledgements

We would like to thank Irene McGowan and David Carroll (Edinburgh University Library), Chris Atton (Napier University Library) and Catherine Fleming (Heriot-Watt University Library) for their participation, expertise and patience.

References

1. Kuhlthau, C., Spink, A and Cool, C. (1992) Exploration into stages in the information search process in online information retrieval: communication between users and intermediaries. *Proceedings of the 55th Annual meeting of the American Society for Information Science*, Learned Information, Medford, NJ, 67-71.

2. Marchionini, G. (1995) *Information Seeking in Electronic Environments*, Cambridge University Press, Cambridge.

[3]We plan to investigate the role of videoconferencing for peer-to-peer collaboration between librarians at the various participating libraries.

3.Orlikowski, W. J. and Yates, J. (1994) *Genre repertoire: norms and forms for work and interaction.* MIT Sloan School Working Paper 3671-94; Centre for Coordination Science Technical Report 166, March 1994. MIT, Cambridge MA.

4. Spink, A. and Saracevic, T. (1993) Dynamics of search term selection process in mediated online searching. *Proceedings of the 56th Annual Meeting of the American Society for Information Science*, 63-72, Learned Information, NJ.

5. Taylor, R. S. (1968) Question-negotiation and information seeking in libraries. *College and Research Libraries*, 178-194.

6. Twidale, M. and Nichols, D. (1996) Collaborative browsing and visualisation of the search process, *Proceedings of ELVIRA-96*.

Information Retrieval: an issue of navigation in Information Space

David Benyon,
Dept. Computer Studies, Napier University, 219, Colinton Road, Edinburgh, EH14 1JD, UK
D.Benyon@dcs.napier.ac.uk

Abstract
The issue of how users can navigate their way through large information spaces is one that is crucial to the ever expanding and interlinking of computer systems. There are many ways of dealing with the issue of navigation. The use of appropriate metaphors is one, virtual reality and 3D interfaces another. A third is to provide adaptive interfaces based on individual differences in users navigational ability. This paper takes a critical look at the alternatives for assisting users to navigate information spaces and concludes by suggesting how a theory of navigation may be developed.

1. Introduction
Information retrieval can come in many different forms. A user writing an SQL query needs to know about the structure of the underlying database. A user trying to find out about a command in a software package might browse a hypertext help system. A person looking for the time of a train might listen to a talking timetable. To find a file which I have saved somewhere I will use some operating system utility. To find one's way through the countryside, a person will use a map. All these (and many more) are examples of information spaces. A database is an information space, hypertext help is an information space, a talking timetable is an information space as is a disc or a map.

An information space is best understood by appealing to the notion of an information system, or information artefact. An information artefact is 'any artefact whose purpose is to allow information to be stored, retrieved, and possibly transformed' (Green and Benyon, 1996). Interactive devices such as spreadsheets, word processors, and operating systems are examples of information artefacts, but so are non-interactive devices like tables, documents and musical scores. An information artefact consists of two levels of description; a conceptual level provides some abstraction of the experienced world and a perceptual level provides a view, or viewport onto that structure. Thus a paper train timetable is one information artefact providing information about train journeys and a talking timetable is another (Bental, Benyon and Green, in press).

All information artefacts employ various symbols, structured in some fashion, and provide functions to manipulate the symbols (whether conceptually or physically). From these symbols people are able to derive information. Thus every information artefact defines an information space - the symbols, structure and functions which allows information to be stored, retrieved and transformed.

As humans we live, work and relax in information spaces. At one level of description all our multifarious interactions with the experienced world are effected through the discovery, exchange, organisation and manipulation of information. Information spaces are not the province of computers. They are central to our everyday experience. Finding your way through an airport, a hotel or a city involves travelling in an information space. Paper documents represent another type of information space. Users will get quite different information from books, from newspapers and from magazines. Similarly they will find various information in timetables, guides and maps.

In computer systems there is a range of information spaces which demonstrate different characteristics. First we can identify software systems. A software system such as a database, spreadsheet or word processor is characterised by some structure; linked files in a relational database, hierarchical structures in browsers, network structures as in hypertext systems, or single file structures as in text processors. These information spaces are primarily concerned with a few kinds of information objects and the attributes which they possess, but typically they provide a wide variety of ways to manipulate them. A mailer systems deals with electronic messages which have attributes such as sender, date, time, subject and so on. A spreadsheet deals with numbers, formulae and graphs. Knowing the type of software that you are using defines the types of object you are dealing with.

Physical media such as discs and CDROMs are information spaces which are primarily user structured. Users are responsible for deciding how to organise their information, how to name things, when to duplicate items and where to store them. Whereas a particular software system such as a database or spreadsheet will have only a few types of object, physical media contain a wide variety of types of object such as applications and application specific objects (mail files, spreadsheet documents, HTML documents etc.), textual data, pictures, video and so on. In contrast to software systems, the viewport onto these physical structures typically provide relatively few methods of manipulation (finding, copying, naming, and so on).

A third type of information space is the network of computers. Here issues such as authorisation and access become more important. A user will typically have different levels and types of access privileges to different nodes on the network. The user may need to consider different protocols, operating systems and the ease of translating between heterogeneous systems. The speed of access is significant. On a single machine the user is rarely too worried about where he or she is in the system. In contrast location is a vital piece of information on a network. Besides the variety of objects, the user now has to know about links and communication.

Of course there are many hybrid systems which demonstrate the characteristics of these types to a greater or lesser degree. Issues concerning the number of types of object and their attributes, the range of manipulative functions and the organisation and distribution of the information characterise different sorts of computer-based information space. It is also important to recognise that these spaces exist within time. A well-structured information space can soon become a tangled mess if someone or something is not there to organise it. Authorisations change, passwords expire and computers are replaced. Information space/time may be a better conceptualisation.

2. Moving around in information space

Humans spend their lives in spaces; whether social, geographical, personal or informational City planners (e.g. Lynch, 1977), geographers, anthropologists and historians deal with spaces, though from a social rather than psychological perspective. Architectural design (Passini, 1992) has considered issues of navigation in relation to the construction of buildings, shopping malls and theatres. Post-modern analysis of spaces takes a critical view of the 'traditional' psychological approach to space and navigation. (See Liggett and Parry, 1995 for a variety of perspectives on spatial theory and urban studies). It is perhaps surprising that these disciplines have commented so little on information spaces *per se.* For example, Shields (1991) is quick to dismiss the objectivist view of traditional geography, preferring instead a semiotic, culturally-based approach. The importance of socially-constructed ideas and values is indeed important, but by emphasising that people derive meaning from such interactions, he shifts us from a geographical to an informational space. The issue of how to travel in that space remains (O' Shay and Jeffries, 1993; Shum, 1990).

Given that humans spend so much time in information spaces it is not surprising to find that psychologists have devoted their attentions to the topic of moving around in information spaces; navigation. Neisser (1976) deals with navigation in terms of cognitive maps; the mental representations which people have of their environment. However, these are not static models. Ecological considerations are important as people draw important cues from the immediate environment and develop knowledge of the space over time and through the experience of interacting with it. Issues of navigation are central to Hutchins recent discussion of distributed cognition (Hutchins, 1995) and cultural differences have been explored with respect to navigation (Hutchins, 1983). A considerable amount has been published on Polynesian navigators in particular (see Hutchins, 1995 p 65 - 93).

Much of the psychology of navigation is couched in terms of *wayfinding*. How do people work out how to reach their destination? The answers to this question are many and, often, unsurprising. People use maps and guides. They exploit landmarks in order to have something to aim for and to recognise a place when they arrive or use 'dead reckoning' at sea when there are no landmarks. Novel techniques can be provided to give the user different views of a real or virtual landscape - fish-eye views focus attention on one area but provide a wide angle to overall view. Three dimensional 'fly overs' of virtual spaces are made possible with computer technology.

For the interests of developers of large scale computer-based information spaces, there may be much which can be learned from the psychological experiences of wayfinding in general. However, wayfinding is primarily concerned trying to reach a destination. Whilst trying to obtain some specific information is certainly one thing which users of computer systems do it is not the only thing. Users do not simply make plans and carry out appropriate actions. They browse, scan and wander through information space.

We might oppose the idea of wayfinding with the notion of *exploration*. With exploration, people are not trying to get anywhere, they are not trying to find their way. Instead they are just interested to have a look around. On many occasions they will not want to return to this particular space. Exploration is concerned with finding out about a local environment and perhaps how that environment relates to other environments. The user is interested in the nature of the terrain, with seeing new vistas. The recognition of place can be important, but much knowledge in exploration is knowledge in the world, and need not be made knowledge in the explorer's head.

A third activity which is undertaken in information space is identifying objects. Here the user is less interested in the location of objects, nor is the user interested in finding a path or reaching a goal. Although object identification is somewhat akin to exploration, the purpose of the activity is different. Exploration focuses on understand what exists in an environment and how the things are related. Object identification is concerned with finding categories and clusters of objects spread across environments, with finding interesting configurations of objects and finding out information about the objects.

3. Metaphors for thinking about information spaces

The use of metaphor in computer-based information spaces has had a long and rather checkered history. Alan Kay has criticised the use of metaphor as a design approach because the computer allows us to escape from the limitations of erstwhile devices and we do not want to import the restrictions inherent in a metaphor into novel systems (Kay, 1990). However the use of extended and complex metaphors continues to be popular with designers and there have been many examples of designers employing rich metaphors in their systems. For example, Card's Rooms metaphor (Card, et al., 1991), the Information City (Dieberger, 1994), Guides (Oren, et al., 1990) and so on. General advice on selecting a metaphor (such as considering the nature of the user population and the correspondence between the structure and functioning of the metaphor and the underlying system) has been published (e.g. Erickson, 1990; Carroll et al., 1988), but little has been said about metaphor and navigation in general. As we have seen, users will want to do different things in the information spaces, and metaphors will help or hinder them in those activities. The different types of information space will demand different types of metaphor.

One recent conference on spatial metaphors and hypertext (Dieberger, 1994) offers some useful ideas. For example, Benking and Judge (1994) identify types of metaphor: 'geometric forms (cube, sphere, polyhedra in general), artificial forms (townscapes, house, room), natural forms (landscape, trees, etc.), systemic structures (highway systems, pathways, flow systems), dynamic systems (atomic, molecular, planetary, galactic systems), traditional symbol systems (mandalas, sand paintings, etc.)'. Besides these overall metaphors, 3D style interfaces can be effective in helping users to visualise the information space, providing a simulated 'fly-over' effect. Given the right software (such as Hyper-G and the Harmony browser), the visualisation can be generated automatically with the attributes of the information objects being used to structure the display (Maurer, 1996 gives a full description of the Hyper-G server and its browsers).

Spatial metaphors come in a wide variety of forms. A city metaphor is one popular example. A city metaphor employs the familiar constructs of roads, subways, houses, shops and a variety of buildings. For example, Dieberger (1994) states 'We are used to navigate cities using very vague descriptions and using various navigational tools like maps, verbal descriptions, sketches, navigational infrastructure (signposts), personal inquiry, transport systems(taxis, subways) and collaboration. Cities are large structured spaces and they can easily be connected to systems containing several cities by using some simple transport systems between cities. Cities are dynamic spaces and people cope well with limited change in those spaces whereas they get quite confused when there is some change in the spatial structure of - say - a desktop. There is a rich body of knowledge about the organization of cities and navigation in cities that can provide a starting point how to make the city metaphor usable as organising principle in a virtual world.' He also points out that users can be transported quickly 'by magic' from one part of the city to another without feeling disorientated as the concept maps well onto using the subway (which counters some of Kay's criticism). There is no doubt that cities provide a rich metaphor, suitable for exploring and locating specific objects.

Other metaphors which might be useful for thinking about navigation include a wilderness (frightening, confusing, enchanting) or a desert (intimidating, beautiful, with no landmarks). Such metaphors might encourage forging a path, enjoying the scenery and getting out. Wilderness and desert can be included in an overall countryside metaphor where different types of terrain represent different types of information. These metaphors encourage exploration by the user; the system provides a high level metaphor, but the user provides a more detailed structure

The night sky offers a different sort of space. It represents objects, clusters, patterns and it is very big. It supports the activities of mapping and identifying objects. However, there are relatively few types of object in the night sky (galaxies, stars, planets). It is the configurations and sub-types of these objects which is of interest. Science fiction concepts such as warp drives and worm holes can be sued for quick and magic transportation to distant parts of the space.

The open sea is another metaphor. It encourages a distinction between the surface and depth. Thus it is natural to think of a lot of information as being hidden beneath the surface and only available for viewing if the user dives down. Currents can link continents and islands and take people to unexpected places. People can look for islands of information; archipelagos provide clusters. Waterworth (1996) combines an island metaphor with a personal 'vehicle' space.

A 3D or 4D object offers a different perspective. Here the user can see the whole of the space and can manipulate it, rotate it. It is suited to users who want to see the whole space and become oriented. A library is suitable for finding information. It is well-organised and structured. The librarian has already organised the material with like subjects kept together. A breakfast bar (so popular in holiday hotels) offers another type of metaphor. Here a whole range of choices are laid out for the users. The user can see how far the space extends and can be offered choice. Users are encouraged to feed a need, to take what they want and to come back for more. It's a constrained, help yourself environment.

This is not intended to be an exhaustive list, nor even to provide an example of each possible type of metaphor. Metaphors work by providing correspondences between one domain and another (Lakoff, 1988). In using them as the basis of a physical design, it is important that these correspondences are consistent and meet with users expectations. Moreover, it is important that the metaphors support what users want to do in that space - to find a way, to identify, cluster and classify, or simply to wander around and explore.

4. A way forward for Navigation Support

We live in a world of information spaces. Increasingly these are very large information spaces. Whether the information space is an application, a hard disc, a CDROM, a local area network or the internet, the problem of how to help the user to navigate, explore and identify the objects of interest is critical to the success of the system.

There is much that computer system designers can learn from other disciplines which deal with space. Lynch (1977) identifies techniques based on a number of characteristics such as; spatial form (different types of models, maps, etc.), temporal form (showing the history of development), sequencing to show paths through a space, visibility to show alternative vistas and ambient qualities of different parts of the space. Architecture (Passini, 1992) provides another perspective. No doubt the information spaces can benefit from the new geography analysis as well (Shields, 1991). However, we have, in addition, our own problems and our own opportunities.

The abstract nature of information space means that designers have to import the concrete from elsewhere. Metaphors for conceptualising information space can exploit common well-grounded experience from the users world. But care needs to be taken to provide consistency and completeness. Metaphors must also be chosen which allow the users to find their way, to wander, scan the space and identify what they want. Information spaces also make different demands on different users owing to their different cognitive capabilities - in particular some measure of their spatial ability.

Computer-based information spaces, however, have a great opportunity. They do not have to be single, stable spaces; they can be adaptive. In a review of adaptive hypermedia systems, (Brusilovsky, 1996) identifies the following types of navigational support which may benefit from some adaptiveness. Direct guidance systems decide which is the next most appropriate node for users. Adaptive ordering adjusts the order of links on a page - for example by making more relevant links nearer to the top of a display. Hiding constrains the size of the information space by automatically hiding pages from users and adaptive annotation adds additional information to links in the form of textual annotation, icons, colour and so on. To be effective these adaptations need knowledge of the individual user; the cognitive characteristics, purpose and domain knowledge which the individual has. Maes (1994) has shown that relevant knowledge about user preferences can be quickly inferred from their early choices and (Benyon and Murray, 1993) show how cognitive characteristics can be inferred from error rate. Other studies support the notion that users differ greatly in their navigational ability (Dahlbëck, et al., 1996; Borgman, 1989).

Thus we have at least the beginnings of a way forward for the problems of navigation in information space. We have a good understanding of the nature of the domain (in terms of the objects, attributes, and relationships which exist) and of the generic tasks which are undertaken in information space (locating an item, browsing, exploring). We have a good understanding of the user characteristics which are significant in navigation (some measure of spatial ability, learning style, experience). We also have a good understanding of the range of metaphors which can be used in navigation and on the impact which these can have on users conceptualisations of the domain.

We now need to bring these features together into a theory of navigation in information space. Hutchins (1995) reminds us that knowledge is distributed between the various representations and functions of the whole cognitive system. In creating information space we, as system designers, have the opportunity to provide effective representations for effective navigation.

References

Benking, H. and Judge, A. J. N. (1994) Design Considerations for Spatial Metaphors. Presentation at *ECHT'94 - the European Conference on Hypermedia technology*, Sept. 1994, Edinburgh. Available from
http://www.lcc.gatech.edu/faculty/dieberger/ECHT94.WS.toc.html

Bental, D. Benyon D. R. and Green, T. R. G. (in press) *Understanding and Building Entity-relationship Models of Information Artefacts*.

Benyon, D. R. and Murray, D. M. (1993) Adaptive Systems; from intelligent tutoring to autonomous agents. *Knowledge-based Systems,* 6 (3)

Borgman, C. L. (1989) All Users of Information Retrieval Systems are not Created Equal: an exploration into individual differences, *Information processing and management* vol. 25, no. 3, pp. 237-251.

Brusilovsky, P. (1996) Adaptive Hypermedia: An Attempt to Analyse and Generalise. In Brusilovsky, P., Kommers, P. and Streitz, N. (eds.) *Multimedia, Hypermedia and Virtual Reality* LNCS 1077 Springer Verlag

Card, S. K., Robertson, G. C. and Mackinlay, J. D. (1991) The Information Visualizer, an Information Workspace. In Proceedings of CHI '91 Conference on Human factors in Computing Systems, ACM publications, New York pp 181 - 188.

Carroll J.M., Mack R.L., Kellogg W.A. (1988) Interface Metaphors and User Interface Design, in Helander M. (Ed.): *Handbook of Human-Computer Interaction*, Elsevier pp. 67-85

Dahlbëck, N., Höök, K and Sjölinder, M. (1996) Spatial Cognition in the Mind and in the World - the case of hypermedia navigation, *The Eighteenth Annual Meeting of the Cognitive Science Society*, University of California, San Diego, July, 1996.

Erickson, T. D. (1990) Working with Interface metaphors. In Laurel, B. (ed.) *The Art of Human-Computer Interface Design* Addison-Wesley, Cambridge, Ma.

Esichaikul, V., Smith,, R. D. and Madely, G. R. (1994) The impact of learning style on problem-solving performance in a hypertext environment *Hypermedia* 6 (2) 101 - 111

Green, T. R. G. and Benyon, D. R. (1996) The skull beneath the skin; Entity-relationship modelling of Information Artefacts. *International Journal of Human-Computer Studies* 44(6) 801-828

Hutchins, E. (1983) Understanding Micronesian Navigation. In Gentner, D. and Stevens, A. L. (eds) *Mental Models.* Erlbaum

Hutchins, E. (1995) *Cognition in the Wild* MIT press, Cambridge Ma.

Kay, A. (1990) User Interface: A Personal View. In Laurel, B. (ed.) *The Art of Human-Computer Interface Design* Addison-Wesley, Cambridge, Ma.

Lakoff, G. (1988). Cognitive Semantics. In Eco, U., Santambrogio, M. & Violi, P. (eds.). *Meaning and Mental Representations*. Indiana University Press.

Liggett, H. and Parry, D. C. (eds.) (1995) *Spatial Practices* SAGE publications Thousand Oaks, Ca.

Lynch K (1977) *Managing the sense of a Region* MIT press, Cambridge, Ma.

Maes P. (1994) Learning Agents in *Communications of the ACM* . August

Maurer, H. (1996) *HyperWave. The next generation web solution* Addison-Wesley Longman, Harlow, UK

Neisser U. (1976) *Cognition and Reality*

O'Day, V. and Jeffries, R. (1993) Orienteering in an Information Landscape: How Information Seekers Get from Here to There, *Proc.. ACM InterCHI'93*, pp. 438-445.

Oren, T. Salomon, G., Kreitman, K. and Don, A. (1990) Guides: Characterising the Interface. In Laurel, B. (ed.) *The Art of Human-Computer Interface Design* Addison-Wesley, Cambridge, Ma.

Passini, R. (1992) *Wayfinding in Architecture* Van Nostrand Reinhold

Shields, R. (1991) *Places on the Margin* Routledge, London, UK

Shum S.B. (1990) "Real and Virtual Spaces: Mapping from spatial cognition to Hypertext", Hypermedia, Vol .2, No.2 pp. 133-158

Walker, J. (1990) Through the Looking Glass In Laurel, B. (ed.) *The Art of Human-Computer Interface Design* Addison-Wesley, Cambridge, Ma.

Waterworth, J., (1996). A Pattern of Islands: Exploring Public information Space in a Private Vehicle. In Brusilovsky, P., Kommers, P. and Streitz, N. (eds.) *Multimedia, Hypermedia and Virtual Reality* LNCS 1077 Springer Verlag

The Value of Information for Distributed Systems

Chris Johnson
Department of Computer Science, University of Glasgow, Glasgow, United Kingdom, G12 8QQ.
http://www.dcs.gla.ac.uk/~johnson E-mail: johnson@dcs.gla.ac.uk

1. Introduction

Recent estimates place the number of Internet users at almost seventy million (Johnson and Kavanagh, 1996). The very success of distributed computing has, however, exacerbated a range of existing usability problems. For example, retrieval delays have reached a stage where many sites suffer from a form of 'electronic gridlock' (Johnson, 1995). At peak times it can take thirty or forty minutes to download relatively small files. Further problems occur because users, typically, cannot predict the time taken to retrieve a file. Server performance varies according to the time of day, local network capacity, the availability of mirror sites etc.

A number of other usability problems arise from the sheer volume of information that can be sent from remote sites. It has been estimated that the number of electronic mail accounts is doubling every twelve months (Schofield, 1994). Many users now find it difficult to plough through the mass of new messages that arrive into their accounts every day (Hiltz and Turoff, 1985). Similarly, the number of daily contributions to many of the most popular newsgroups has led to a form of 'information saturation'.

1.1 Electronic Gridlock

The term 'electronic gridlock' is used to refer to situations in which networks and servers cannot cope with the amount of information that is being requested from them. Recent innovations in caching, including the widespread use of mirror sites, have done much to increase the efficiency of network communications. However, the majority of Internet users still rely upon modem and ethernet connections. As more and more people access a greater and greater volume of data over these low bandwidth systems, there has been an increase in the amount of time that is required to download information (Johnson, 1995). Kuhmann (1989) argues that these retrieval delays have a profound impact upon the quality of interaction with distributed systems. The difficulty of predicting the amount of time taken to retrieve a file frequently leads to frustration and error.

1.2 Information Saturation

The term 'information saturation' refers to situations in which users simply cannot cope with the volume of information that is being received from remote sites. Hiltz and Turoff (1985) argue that this problem stems from 'information entropy'; electronic documents are insufficiently well structured for users to select the items that they need. This problem affects a wide range of applications. For example, it can often take ten or twenty minutes to extract critical items from the morning's electronic mail messages. The use of threads in newsgroups and the introduction of dedicated mailboxes have reduced some of the problems posed by information saturation. However, the low marginal cost associated with electronic communication and the large number of potential senders helps to offset any benefits that these systems might offer.

1.3 The Utility of Information

This paper presents techniques that can be used to reduce the frustration and delays that are caused by information saturation and electronic gridlock. In particular, the relevance or utility of information must be considered if designers are to improve the usability of distributed systems. Section 2 argues that users must be able to assess the quality of remote information before they download it. Even if users cannot directly avoid the delays created by electronic gridlock then it should be possible for them to ensure that the information will be worth waiting for.

A weakness of the previous approach is that users must manually assess the value of each document before they download it. This increases the burdens associated with information saturation. In contrast, we have also built a number of systems that automatically infer measures of utility. Section 3 describes an information retrieval system that exploits user-defined measures of relevance to match natural language queries with large numbers of documents. Section 4 describes an email filtering tool that exploits a user-defined lexicon to rank the priority of incoming messages. Both of these systems provide filtering mechanisms that reduce information saturation. Both exploit measure of utility and value to control the filtering process.

2. User Defined Utility

Designers can help users to combat information saturation and electronic gridlock by carefully selecting the information that is used to index remote resources. For example, many links on web pages provide no information about the size of the resource that they reference This exacerbates the problems created by electronic gridlock because users cannot even predict the retrieval delay under 'perfect' conditions. In contrast, Figure 1 illustrates a web page that avoids this problem. The length of each video clip is indicated next to the link that is used to retrieve the film. Running time was chosen rather than file size because this page was designed for school children in the remoter areas of Scotland. Preliminary evaluations suggested that file sizes had little or no meaning for the intended audience. Classes could, however, use clip length together with a knowledge of previous delays to estimate the period of time required to download subsequent videos.

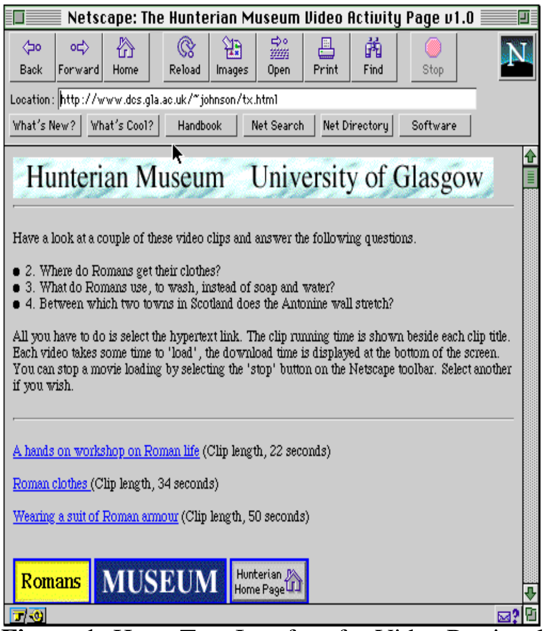


Figure 1: HyperText Interface for Video Retrieval

Although the information in Figure 1 helps users to estimate the costs of access a video file, it provides little indication of the benefits or utility of the films. In consequence, 30-40% of requests were abandoned before they were completed. The actual percentage varied with the length of the clip. These results were extremely depressing for the designers, including the author of this paper, because a great deal of time and effort had gone into the development of the web site. The videos were produced by a professional crew working for a national broadcaster. They were processed using 'state of the art' compression software. Usability studies were conducted into optimal file lengths for web-based retrieval (Johnson and Kavanagh, 1996). In spite of all of these precautions, users were not prepared to wait for the films to be downloaded.

As mentioned, the main limitation with the interface shown in Figure 1 was that it did not present users with any indication of the quality or value of the material that it provided. It was similar in style to the many web pages that offer low quality videos with poor direction and inadequate compression techniques. Usability trials confirmed our analysis (Johnson and Kavanagh, 1996). Users were more likely to complete a retrieval request if they had already downloaded one of our videos and hence were aware of the production standards. We, therefore, re-designed the page to provide users with additional cues about the quality of the underlying material Figure 2 illustrates the new page design; thumb-nail icons and descriptive paragraphs are used to indicate the production and compression quality of the associated material.

Figure 2: Indicating the Quality of Remote Information

This revision had a profound impact upon user attitudes towards the retrieval delays. There was a significant reduction in the number of abandoned requests, even for relatively long clips over 70 seconds. More information on the implementation and evaluation of these pages can be found in Johnson and Kavanagh (1996). The important point here is that even if users cannot predict the delays imposed by network loading, our experience with this application shows that designers can affect attitudes to retrieval delays by providing a clear indication of the quality of the resources that users request.

The revised interface, shown in Figure 2, represents a relatively simple solution to the problem of electronic gridlock. By providing users with more information about the quality and content of remote resources, they are better able to judge the potential rewards from downloading those resources. Unfortunately, this approach exacerbates the problem of information saturation. Users must browse through additional summary information, such as the video descriptions shown in Figure 2. In contrast, the following section describes an information retrieval tool that automatically generates a ranked list of resources. Rather than the user manually searching for the document that provides the greatest benefits, this system automatically reduces information saturation by filtering out irrelevant documents according to pre-defined measures of utility.

3. System Defined Utility

Information saturation occurs when users are over-whelmed by the volume of data that is provided by remote sites. This is not a novel problem. Information retrieval systems have long been faced with the task of helping users to extract relevant information from a mass of background 'noise' (van Rijsbergen, 1979). A number of web-based search engines have been developed to address the increase in scale that exacerbates information saturation for distributed systems. For instance, we have recently extended the Inquery tool to support free-text retrieval for multimedia information. Figure 3 shows the results of a query about wooden artefacts in an ethnographic collection. The HyperText links can be selected to retrieve photographs, videos or text about the objects in question. In terms of this paper, the significant point about this tool is that it processes the users' query to *rank* the documents that are extracted from a catalogue or database.

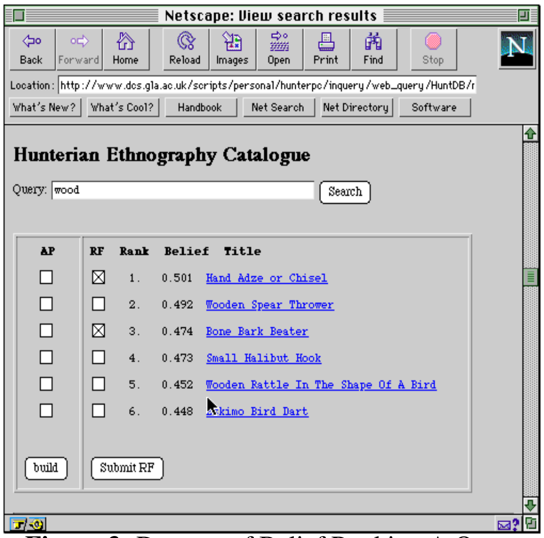


Figure 3: Degrees of Belief Ranking A Query

The rankings shown in Figure 3 are automatically calculated from the relevance of the document to the query. In order to do this, the system exploits a thesaurus of pre-defined synonyms. Associated with each synonym is a degree of belief. If for example, a user entered a query about a canoe then the system might retrieve any document which contained the word kayak. Associated with that document might be a 70% belief that any query relating to canoes would be satisfied by a document referring to kayaks. In Figure 3, this is illustrated by a belief of 0.501 (50%) that a query about wood will be satisfied by a document on a 'Hand Adze or Chisel'. This systems, therefore, automates the filtering process that would otherwise have to be performed by the manual inspection of distributed resources.

A limitation with the approach described above is that different tasks and contexts determine the utility or value that a user might have for a particular resource. For example, a document about a kayak might be completely irrelevant if the user was searching a boat building database for a canoe design. This context dependency creates problems for interface designers because the degrees of belief in particular synonyms tend to be derived from a standard thesaurus (Lalmas, 1996). In practical terms, the context specific associations between documents means that general purpose information retrieval systems may only provide limited benefits for the problems of information saturation. Users still have to extract relevant documents from a large number of false 'hits'. Fortunately, most tools enable their users to tailor the degrees of belief that are embedded within information retrieval applications. They can exploit relevance feedback techniques to indicate which documents were relevant to a particular query. In the tool shown in Figure 3, the user simply checks the boxes labelled RF to indicate that a document was relevant to their query. The system then uses this information to increases the degree of belief that the document matches the query term for any subsequent retrieval. The selected boxes in Figure 3 make it more likely that the hand adze and the bone bark beater will be recalled fro any subsequent query about wood. The critical point here is that information retrieval engines are moving beyond pre-specified measures of relevance. Current tools now exploit mixed initiative approaches where both the user and their system co-operate to reduce the problems of information saturation. They do this by exchanging the actual and assumed *value* of a document for a particular query.

4. Co-operative Utility

Information saturation not only affects web based retrieval. The number of daily contributions to newsgroups and mail accounts can quickly overwhelm even the most enthusiastic reader. This makes it difficult for users to identify critical items from the mass of more trivial messages. Fortunately, a number of filtering tools can now be used to separate threads or themes in electronic correspondence. These systems are, however, relatively unsophisticated. They assign weightings to mail from particular users or to messages with particular sequences of characters in their subject line. Correspondence containing the line Re: Tomorrow's Meeting and Re: Tomorrow's Appointment would, therefore, be assigned different priorities even if they referred to the same event.

Existing filtering systems can be enhanced through the introduction of a synonym-based thesaurus, as described in the previous section. This would enable the application to identify links between terms such as Appointment and Meeting. The degree of belief in these synonyms can then be used to rank mail messages according to their relevance for a particular query. This is similar to the way in which the information retrieval system ranked multimedia documents in the previous section. For example, the user might request correspondence about some books that they had borrowed. Such a tool might then use the degrees of belief in the thesaurus to rank mail about outstanding library fines above messages concerning swimming lessons. This ranking would then be used to filter the messages that were presented in response to the query.

A number of problems limit the utility of information retrieval techniques for mail filtering. The value of any synonym is highly dependent upon the relationship between the sender and the recipient. This, in turn, effects the utility of the synonym as means of filtering previous correspondence. For example, if a user worked in a library then it would be extremely irritating if the system retrieved all correspondence about that library in response to a query about book fines. On the other hand, if the user worked in another domain then it would be equally irritating if the mail system filtered out previous correspondence about the library when prompted for information about book fines. A number of further problems limit the utility of information retrieval techniques for mail filtering. In particular, mail messages cannot simply be ranked by examining synonyms within the body of the text. For instance, the value or relevance of correspondence can also be determined by the identity of the sender, the date on which it was received, the date on which it was sent, the length of the message and so on.

We are developing a system that addresses the limitations of existing filtering tools by analysing both the content and the context of mail messages. In order to do this, the user provides the set of criteria that guide the automatic ranking of mail messages. This is a co-operative approach in that the system applies the users's criteria when assessing the priority of each individual message (Mitchell, 1996). At present, the implementation relies upon Prolog clauses. For example, the following relations represent synonyms within either the subject line or the body of a mail message:

synonym('joke', 'humour'). (1)
synonym('mirth', 'humour'). (2)

These clauses can be used to rank incoming mail. For example, the following clause states that a message should be assigned a relatively low ranking if it contains a synonym for the word 'joke'. Here we use the numeric value, 10, to indicate a particular ranking, this could easily be changed to indicate a mail box where the message might be stored:

filter(Message, 10) :-
re(Message, Subject),
synonym(Subject, 'joke'). (3)

This offers a number of advantages over the standard thesaurus, described in the previous section. The ranking is related to the situation in which the synonym is used rather than to the synonym itself. A library worker need not, therefore, assign a high priority to messages that contain a synonym for the word 'book'. A further benefit of this approach is that users can specify a range of other factors that might be considered during the filtering process. For example, the following clauses relate positions within an organisation to particular individuals and their accounts. Clause (4) states that ray@dcs is the name of an account which is currently operated by the Head of Department. Clause (5) states that the name of the Head of Department is Ray Welland. Clause (6) then associates the highest priority, *1*, with any message that is received from a user who is synonymous with the Head of Department:

synonym('ray@dcs', 'Head of Dept'). (4)

synonym('Ray Welland', 'Head of Dept'). (5)

filter(Message, 1) :-
from(Message, Sender),
synonym(Sender, 'Head of Dept'). (6)

A number of issues remain to be addressed. For instance, clauses such as (4), (5) and (6) are not in a form that can easily be manipulated by the users of mass-market mail tools. This is an important limitation if advanced filtering techniques are to applied to reduce the problems created by information saturation. Similarly, previous examples have not considered how the results of the filtering process might be displayed to particular users. Figure 4 illustrates how the rankings derived from clauses such as (3) and (6) can be translated into the clusterings of a 'starfield display' (Ahlberg and Schneiderman, 1994). In this prototype, mail with the highest value or ranking is clustered at the centre of the screen. In Figure 4, it is assumed that all mail about the *Project* has a higher priority than *Miscellaneous* mail or even mail that is classified as *Important/Urgent.* Lower priority information is placed on the edges of the display. For example, mail about the *Comms* course has been assigned a relatively low value. Within each cluster, users can specify that the messages are ordered by date of arrival, by date of dispatch, or by sender's address.

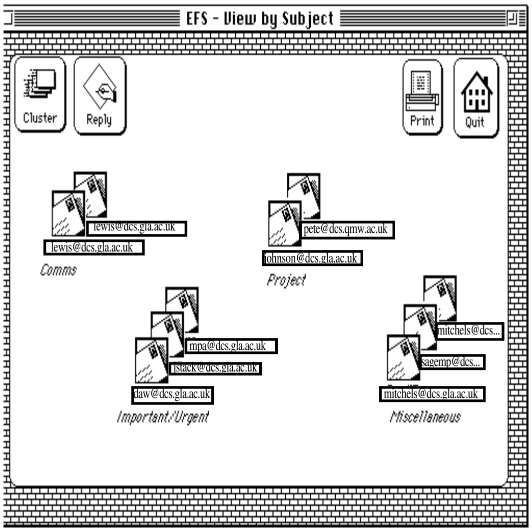


Figure 4: Starfield Display For Mail Filtering

Further work intends to address a number of remaining problems with this application. For example, if users receive large amounts of mail then less important messages may simply be lost on the edges of the screen. Other problems relate to the calculation process that is used to derive the rankings. Mail about two different topics will be merged together if both clusters are given a similar weighting. This reduces the effectiveness of filtering as a means of combating information saturation. Users must manually extract individual items from the merged clusters. Similarly, it is difficult for the system to assign a weighting to messages that do not fit into any recognised categories. Again, users must manually extract important messages from the mass of less relevant items.

Current research is investigating the problems described above. For instance, we intend to integrate the individual ranking mechanisms defined by clauses, such as (3) and (6), together with the automated relevance calculations described in Section 3. This would enable users to issue queries that would filter any unclassified mail against the standard thesaurus in the information retrieval system. Classified mail could be presented using the starfield display in Figure 4. Further research will also extend relevance feedback from the information retrieval system into the mail filter. It is perfectly possible for mail to be categorised in the wrong cluster. This occurs if a synonym appears in more than one context. Future research will help the user to inform the system when an incorrect value has been assigned to a particular item of mail.

5. Conclusion

This paper has argued that the increasing importance of distributed systems is creating new challenges for the design of interactive systems. In particular, the limitations of network and server capacity have introduced significant delays when sending or receiving information over the Internet. This problem has been termed 'electronic gridlock'. We have also identified a dual of this problem that arises when remote sites offer more information than a user can cope with. This problem has been termed 'information saturation'. For example, it is not uncommon for users to be over-whelmed by the amount of email that is forwarded by popular mail servers. Similarly, it difficult to extract relevant information from the rich mass of data that is provided by many web sites.

Several techniques have been proposed to reduce the impact of electronic gridlock and information saturation. In particular, this paper has argued that interface designers must help users to assess the value of remote resources. For instance, it must be possible for individuals to determine the content and production quality of multimedia documents if they are to sacrifice their machines, and phone lines, for long periods of time. This helps to reduce the frustration and error that often result from the delays that are associated with caused by electronic gridlock. Unfortunately, this approach also increases the problem of information saturation. Users must browse through additional information about the content and production quality of remote resources.

A number of information retrieval systems have been developed to reduce the problems created by information saturation. For example, we have extended the Inquery tool to extract multimedia documents from Museum catalogues. This system extracts a number of terms from each of the user's queries. These are then used to identify a set of synonyms from a standard thesaurus. Associated with each synonym is a degree of belief that is then used to rank the relevance of a document which contains the synonym. The closer a synonym is to the query then the higher the document's ranking will be. This tool has the advantage that it automates the mechanism by which values are associated with remote resources. It can, therefore, helps users to search the mass of information that is being provided by distributed systems.

Unfortunately, the utility or value of remote information cannot simply be measured in terms of degrees of belief between synonyms. For instance, the importance of an email message may depend on the sender, the date it was received, its length and so on. We have, therefore, developed an email filtering system that exploits all of this information to analyse the priority of incoming messages. An important feature of this application is that the user supplies the criteria for the automated ranking process. This is intended to ensure that both the user and their system agree over the value of information from remote sources.

Acknowledgements
Thanks go to the members of the Glasgow Interactive Systems Group (GIST). Particular thanks go to Paul Hunter, Jim Kavanagh and Susan Mitchell who drove the implementation work behind the video pages, the mail filtering system and the catalogue retrieval application. Chris Johnson's work is supported by the UK Engineering and Physical Sciences Research Council, grants GR/JO7686, GR/K69148 and GR/K55040.

References

C. Ahlberg and B. Schneiderman (1994), Visual Information Seeking: Tight Coupling of Dynamic Query Filters with Starfield Displays. *CHI'94*, 313-317, ACM Press, New York, United States of America.

S.R. Hiltz and M. Turoff (1985). Structuring Computer-Mediated Communication Systems to Avoid Information Overload. *Communications of the ACM*, 28: 680-689.

C.W. Johnson (1995), Time and the Web: Representing and Reasoning about Temporal Properties of Interaction with Distributed Systems. In M. Kirby, A. Dix and J. Finlay (eds.), *People and Computers X,* 39-50, Cambridge University Press, Cambridge, United Kingdom.

C.W. Johnson and P.D. Gray (1996), Error Driven Design. In M. Harrison and J. Vanderdonk (eds.), *Design, Specification and Verification of Interactive Systems'96*. Springer Verlag, Berlin, Germany.

C.W. Johnson and J. Kavanagh (1996), Electronic Gridlock. In H. Thimbleby and A. Blandford (eds.) *Adjunct Proceedings of HCI'96*. British Computer Society, London, United Kingdom.

W. Kuhmann (1989). Stress Inducing Properties of System Response Times. *Ergonomics*, 32(3):271-280.

M. Lalmas (1996). Modelling Information Retrieval With The Dempster-Shafer Theory of Evidence. In *ECAI Workshop on Uncertainty in Information Systems: Questions of Viability. Budapest, 1996 - to appear*.

S. Mitchell (1996). Novel Techniques for Automatic Mail Filtering. Technical Report. Department of Computer Science, University of Glasgow, Scotland.

S. Schofield (1994). *The UK Internet Book*. Addison-Wesley, Wokingham, United Kingdom.

C.J. van Rijsbergen (1979) *Information Retrieval.* Butterworths, London, United Kingdom, 2nd edition.

Using People As An Index To Information

Clare F. Harvey,
University of Sunderland, School of Computing and Information Systems, St. Peter's Campus, St. Peter's Way, Sunderland, SR6 ODD .
clare.harvey@sunderland.ac.uk

Abstract
In information-seeking situations, people often ask each other who is the best person to help them get to grips with a particular topic. InfoVine is a piece of software which mimics this informal word-of-mouth system. By employing user modelling techniques it collects a database of information about people. This is then available for interrogation. It forms a "human resource layer" of organisational memory.

1 Introduction
Information overload has become a great problem. A good strategy for avoiding information overload is to ask the advice of colleagues on where and what the most useful information is. This strategy is successful because it works not only when information is over-plentiful, but also when it is scarce. In either case it makes use of the fact that somebody else has already undertaken pioneering work in making initial searches and in filtering information. Within an organisation, the projects that an individual is involved with will determine the areas in which they are knowledgeable. They will have been able to amass expertise and interests in some areas and not in others. To find the right colleague to talk to on any given subject, a person will ask around; they will use the informal word-of-mouth system which operates within an organisation.

2 The Human Resource Layer of Organisational Memory
The knowledge which employees have about various topic areas may be seen as one layer of organisational memory. The knowledge about **who** knows about **what** is a further over-lying layer. It is this "human resource layer" of organisational memory which is in jeopardy whenever there is a period of great mobility in the workforce due either to times of prosperity or to a period of hardship (for instance, companies downsizing during the recession). When these human resource layers of memory are damaged people start to repeat work, to reinvent the wheel, to fail to make connections with those who could advise and collaborate, to take a long time to get started in a project area which is new to them. The informal word-of-mouth system, while very effective, is fragile and susceptible to organisational change.

3 The InfoVine Software Prototype
InfoVine is a prototype software system which automates this word-of-mouth process by which colleagues find out who can help them (Refs. 1,2,3,4,5). I t was built in Visual Basic Ver. 3. It's first implementation has been in the context of laboratory notebooks at Courtaulds Coatings Holdings Ltd. at their research and development site in Gateshead. The key to the system's design is that people are used as an index to information. The user can search for entries in a given topic, or, using exactly the same user interface and query form they can retrieve information on who within the organisation has either expertise or an interest in the given topic area. Thus, they can retrieve both the raw material and information about somebody who should be able to interpret it for them.

There has been much interest lately in skills registers. However a major problem with these is the need for updating. Considerable energy and organisational power needs to be devoted to their upkeep. If the individual who is enthusiastic about keeping a skills register is diverted away from the project, it soon deteriorates. Similar problems have been reported concerning the keeping of profiles for current awareness or SDI (Selective Dissemination of Information) (Refs. 6,7). The InfoVine approach differs from skills registers/databases in that the information on individuals' interests and expertise' is built up organically using user modelling techniques. Parallel to the original internal database (in this case, laboratory notebooks), InfoVine constructs a database of user models which is made available to other users. It does this by inferring information about their interests and expertise form their usage of the database.

4 Evaluation of InfoVine
The prototype has already undergone preliminary evaluation at Courtaulds Coatings Holdings. Because it was built to contain knowledge specific to one particular subsection, the number of people who could undertake hands-on evaluation was limited. The array of strange terms was a matter for confusion when people who were not from that subsection tried to use it. As a result, it was evaluated by 3 Courtaulds information scientists, 4 first-line managers and 6 scientist/technologists (Ref. 8). The evaluation was very positive. Both the implementation and the design of the evaluation form took care to separate out the InfoVine word-of-mouth features from the information retrieval features. People

felt that the features were "credible" "interesting" and "not threatening" and seemed quite prepared to use them.

5 Future Work

The original implementation of InfoVine has been made at Courtaulds Coatings Holdings' Felling Site, where paint technology is researched and developed. The implementation of InfoVine which took place there was based on a database index of laboratory notebooks. The interactions of scientists with this database were used as a basis for building user models. This project is continuing with implementation now occurring over a larger number of people so that more evaluation can be performed.

The author and her collaborators are convinced of the genericity of the ideas illustrated by InfoVine and are currently starting implementations in their own "back yard", in both Sunderland University Library and the School of Computing and Information Systems. In both cases, the aim is to increase the flow of information concerning the research areas and specialisms of staff.

Future implementations will include a Visual Basic front end on a Microsoft Access back end. They will seek to make the software more generic. For instance, it will be possible for the system manager to adjust the measures by which interests and expertise are diagnosed to allow the system to be better tailored to patterns of work in different environments. The new software will also seek to extend the metaphor of people as an index to information. It will capture and hold information about the composition of teams. Thus if the person you seek is no longer available it may be possible to trace somebody who was working closely with them at the time when they developed a particular piece of expertise, and can act as a guide to their work.

6 Conclusions

The principle of finding out who can help you rather than going straight to the underlying data, is practised by many people in information-seeking situations, thus the model of the system should be one which is familiar to the user. Positive evaluation results showed that it was an easy idea for users to grasp.

InfoVine's database of information about the interests and expertise of people grows organically. This helps to build the human resource layers of organisational memory and to contribute to organisational learning.

InfoVine will not replace the informal word-of-mouth system. Where the informal word-of-mouth system is working well, InfoVine will only be a second resort, but in contexts where the human resource layers of organisational memory are damaged, or simply stretched by time passing or by the distance between two sites, InfoVine will be an invaluable guide. It will operate in the interstices where the informal system does not work well.

The contexts where InfoVine operates well are becoming increasingly important as the Internet and email encourage and facilitate interaction with individuals whom distance makes it hard for us to know well. InfoVine should give an indication of who might be the right person on another site or at headquarters to contact.

The author would like to thank the DFE and the British Library R&DD for financing and administering her Information Science studentship and fellowship which allowed this work to take place.

REFERENCES

Harvey, C. F., Smith, P., Lund, P. (1993a) "Emulating the Word-of-Mouth System in a Research and Development Environment" *Online Information 93 Proceedings* Eds. Raitt D.I. & Jeapes B., Learned Information, Oxford,UK.

Harvey, C. F., Smith, P., Tan, T. C. & Pegman M (1993b) RADA: an Information System to emulate and enhance corporate communication structures, *BCS 14th IRSG Colloquium, 1993.*

Harvey, C. F., Smith, P., Lund, P. & Bell G. (1994) The Word-of-Mouth Emulator - Representing Expertise and Interests in a Research and Development Environment. *Online Information 94 Proceedings* Ed. Raitt D.I. Oxford,UK: Learned Information.

Harvey, C.F. & Smith, P. (1995) A New Paradigm for Circulating Proprietary Research and Development Information. *ET95 Proc.s on the Conference on Emerging Technologies, Jan. 5-6, 1995, London.* CCTA.

Harvey, C.F. & Smith, P. (1996) An Information System to Improve Organisational Memory. ECIS, '96, Proc.s of the European Conference of Information Systems, Lisbon 2-4th July, 1996.

Weaving, E. (1991) Current Awareness Service and the Information Centre, *Aslib Proceedings,* **Vol. 43, No. 10** (pp 299-303).

Hawkins, D.T. (1990) Customised Information, *Online*, Vol. 14, No. 5, (pp 117 -120).

Harvey, C.F. (1995) The Development of a Computerised Word-of-Mouth Emulator , PhD thesis, University of Sunderland, UK.

A Cognitive Model of Search Behaviour: a first step towards improving interface design for IR

A.G.Sutcliffe and M.Ennis
Centre for Human Computer Interface Design, School of Informatics, City University, Northampton Square, London EC1V OHB, U.K.
Tel: 0171 477 8994
e-mail:M.Ennis@city.ac.uk4

The paper outlines theoretical work on the cognitive activity underlying searchers behaviour in information retrieval. This is to help address the mismatch between information retrieval interfaces, and their associated functionality, and users' mental models of the search process. A cognitive model was developed based on existing theories of search behaviour and factors known to effect the retrieval process. The purpose of our work is to minimise the 'cognitive effort' required to perform an information seeking task. This is to be performed by linking the model of human action to what users need to know and the task support provided by an IR system (design requirements).

1. Introduction

Several models offer some insight into users' information retrieval behaviour in various contexts, but are unable to predict it: Bates (79, 89); Markey and Atherton (78); Marchionini (95) and Ingwersen (82). For instance, we have very little understanding about the differences between goal related searches and more serendipitous browsing; and how these may be supported by query languages, concept maps, pre-formed queries, etc. The demand for intelligent user interfaces for IR has caused a need for a predictive model of search behaviour. We propose a cognitive model to provide a basis for predicting user behaviour in the search process. Guidelines for interface design will be developed from the model to improve the match between the cognitive activities of a searcher and systems' facilities.

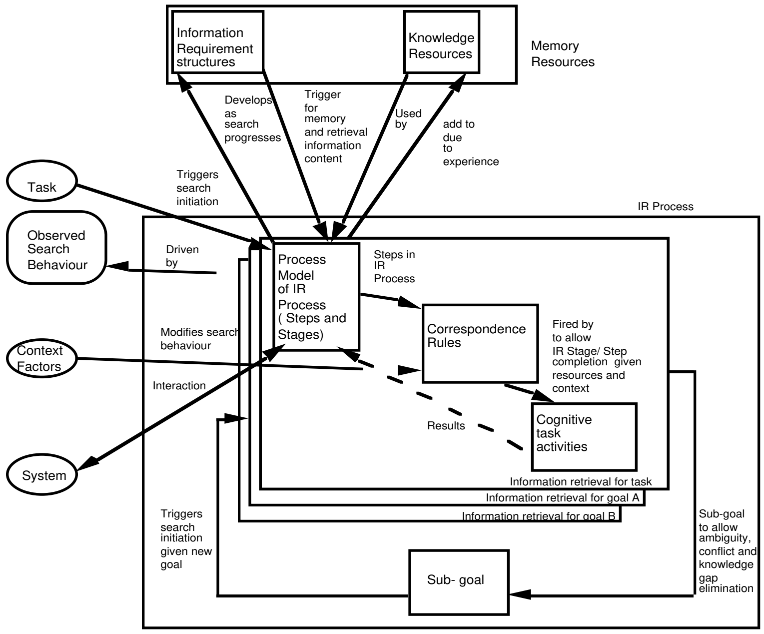
2. Cognitive model of search behaviour

The model accounts for the different types of behaviour observed in search activity: goal directed browsing; exploratory browsing; recursive sub-goaling; rapid prototyping and direct query composition. Figure 1 shows how a multi-layered approach to modelling searcher behaviour including multiple retrieval cycles tailored to different information problems can explain sub-goaling and browsing, in which searchers are seen to attack the information need as a series of information problems from which they collate the solution. The cognitive model was motivated by the desire to incorporate factors known to effect the search process: task complexity, Large et al (94); user intention, Solomon (92); information need type, Elkerton et al (84); value of the information required, Irwin et al (57) and Wendt (69); knowledge sources, Michel (94) and Allen (91); information source selection, Bassili (77) and motivation, Jacobson et al (92). The main aim of the model is to provide a flexible structure into which context descriptions and searching strategies can be processed to forecast a searcher's behaviour.

There are static and dynamic features in the search process. Static elements occur irrespective of the task performed or the searchers' knowledge profile, while dynamic features are driven by the situation; task context and searcher abilities.

4 URL: http://www.soi.city.ac.uk/hcid/index.html

Figure 1: *Architecture*

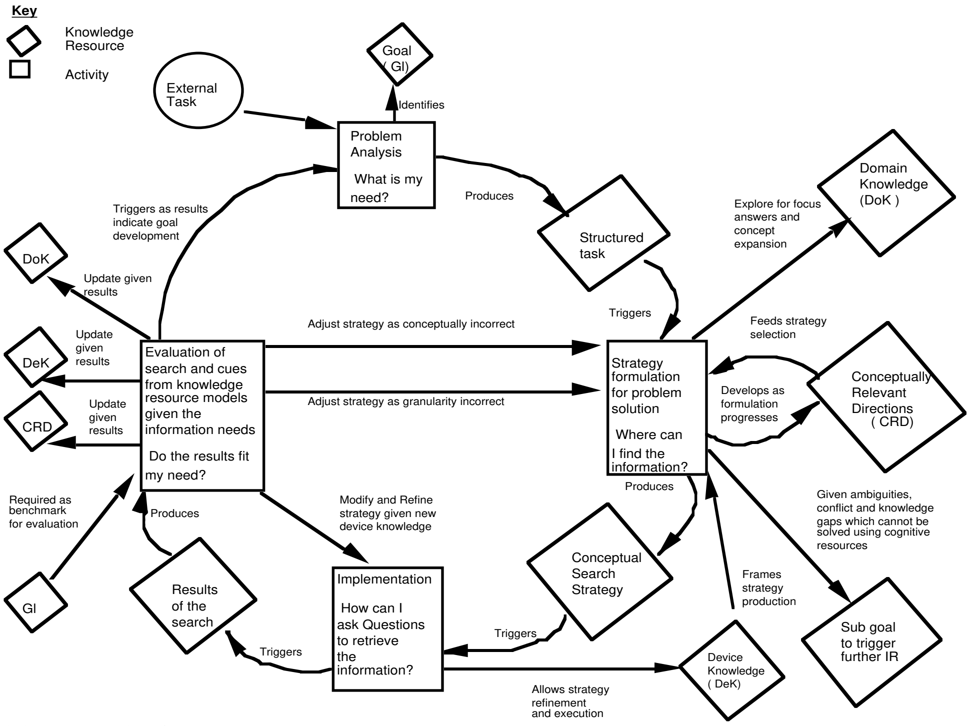


The architecture (figure 1) illustrates the interaction between the main components within the IR process: the process model of IR; correspondence rules and cognitive task activities.

Process model of IR

The process model of IR includes four main activities (problem analysis; strategy formulation; search implementation and evaluation); transition paths between these activities and knowledge sources which effect their behaviour (see Figure 2). These activities are the core processes from which all tasks, cognitive and physical, in the search process radiate. The cognitive tasks required to complete the core activities in the process model are a function of a searchers knowledge profile. A knowledge profile is based on the level of granularity of users knowledge resources required in the different stages of the IR process: Michel (94) and Allen (91).

Figure 2 : *Process model of IR*



Problem analysis: This transforms an unstructured need, or physical stimuli, into a structured and targeted need. The factors affecting this activity are task characteristics and complexity (Large et al (94), Elkerton et al (84), Irwin et al (57), Wendt et al (69)); and the searchers' knowledge (Michel (94) and Allen (91)). The activities performed to achieve this are: recognition of the existence of a information need; identification of concepts and their inter-relationships in the task; infer problem space elements based on domain knowledge; filter out irrelevant information; structure and prioritise problem space elements and trigger either the formulation of a strategy based on the current problem space or a new IR cycle to solve a part of this problem space.

Strategy formulation: This explores the problem space and formulates a plan of action to satisfy the information need. The activities performed to achieve this are: assess methods of locating the information required and select the path of least cognitive resistance. The search methods for locating information are: explore concepts in knowledge schema's of the domain to complete the mental 'image' of the information required (i.e. reason about the problem space and find lexical/synonym variations for concepts); execute query based on current concepts; and explore concepts in the thesaurus etc. The choice of which method to use is dependent upon the searchers' knowledge levels, and the trade-off between search efficiency and cognitive effort.

Search Implementation: Translate the information requirement, search plan, for the information need into physical actions and execute it. The ability to perform this effectively is dependent on the users knowledge and past search success (Large et al 94). The activities performed are: the refinement of the current search strategy into the systems' language and facilities; and the execution of this refined strategy through physical action. Refinement is broken down into: select, or create, a plan for search interaction; select term representations for concepts and appropriate relationship operators.

Evaluation: Results retrieved by the system are compared with the target information need. The evaluation breaks down into: evaluate the number of records retrieved; evaluate records retrieved on their information content; evaluate search success or failure; and plan future activity given the search results.

Correspondence Rules

Correspondence rules dictate the composition and sequence of cognitive activities in the four main stages for a search given a searcher's context. Rules allow the model to account for the diversity of tasks in information retrieval; differing levels of detail in users' knowledge and a searchers completeness requirement. The searcher uses a sub-set of the cognitive processes available as dictated by correspondence rules which specify the cognitive task activities to be performed given the presence of specific trigger conditions.

e.g. The following two simple rules demonstrate the cognitive tasks required to identify concepts from an external stimuli given different domain knowledge levels. The rules have been developed to show the effect of different knowledge levels on the configuration of the cognitive activities performed in the search process.

```
if      Domain_knowledge_high,
        Identify_concept_focus_points (Perceptual_buffer, Wm[Goal_list], LTM)
then
        Scan_perceptual_buffer (Perceptual_buffer, LTM, Wm[Goal_list]),
        Importance ( Wm[Goal_A | Goal_List], LTM),
        Identify_concept_focus_points ( Perceptual_buffer, Wm[Goal_A | Goal_List], LTM).

if      Domain_knowledge_low,
        Identify_concept_focus_points (Perceptual_buffer, Wm [Goal_list], LTM)
then
        scan_perceptual_buffer (Perceptual_buffer, Wm[Goal_list ], LTM),
        Identify_concept_focus_points ( Perceptual_buffer, Wm [Goal_A | Goal_List], LTM).
```

Cognitive task activities

Cognitive task activities are the actions and decision processes underlying the four main activities in the process model. Thus these cognitive tasks provide the functionality to deal with the information content in the external tasks and allow decision judgements to be made. The cognitive tasks perform operations which manipulate; extract and compare memory sources allowing search behaviour to solve the problem.

3. Evaluation and further work

The model is being evaluated using the MEDLINES database through verbal protocol analysis, and standard performance and query generation metrics. MEDLINE is the copyright of the US National Library of Medecine and SilverPlatter. These will allow us to make any adjustments required to the model and to test our guidelines for IR interface design through re-engineering the interface. The model is being implemented in the COGENT computational modelling environment (Cooper 96) owing to the complexity of explaining the effects of the correspondence rules on behaviour without animating them.

4. Conclusion

Our tentative conclusions from cognitive modelling of IR are that representations of the search space should incorporate methods to enhance and improve the query building process and thus help user articulation in an attempt to reduce cognitive load. The interface should include automatic term suggestions to show synonyms. Due to cognitive laziness searchers often don't assess alternative terms for concepts. Users require guidance in the production of their searches if they are to be more optimal and are to reduce the amount of 'cognitive effort' associated with using the retrieval system, but the level of this guidance must be tailored to their knowledge and experience in the domain and device. This suggests that an embedded, software version of our model may be useful as a basis to develop an intelligent advisor for IR strategy assistance as well as a mechanism for tailoring appropriate IR support facilities for users' needs.

Acknowledgements

Mark Ennis is funded by the EPSRC under grant number: 95306422

References

B.L. Allen, Cognitive research in information science: Implications for design, Annual Review of Information Science and Technology, 1991, 26, pp 3-37

J.N. Bassili and D.T. Regan, Attributional focus as a determinant of information selection, Journal of Social Psychology, 1977, 101, pp.113-121

M.J. Bates, Information search tactics, Journal of the American Society for Information Science, 1979, Vol. 30, No. 4, pp. 280-289

M.J. Bates The design of browsing and berrypicking techniques for the on-line interface, On-line Review, 1989, Vol. 13, No. 5, pp. 407-424

R. Cooper, Perseverative subgoaling in production system models of problem solving, In Cottrell G. W.(ed), Proceedings of the 18th Annual Conference of the Cognitive Science Society, 1996, pp. 397-402, San Diego, CA

J. Elkerton and R.C. Williges, Information retrieval strategies in a file search environment, Human Factors, 1984, 26(2), pp171-184

P. Ingwersen, Search procedures in the library- Analysed from the cognitive point of view, Journal of Documentation, 1982, Vol. 38, pp. 165-191

F.W. Irwin and W.A.S Smith, Value, cost and information as determiners of decisions, Journal of Experimental Psychology, 1957, 54 (3), pp 229-232

T. Jacobson and D. Fusani, Computer, system and subject knowledge in novice searching of a full text multifile database, Library and Information Science Research, 1992, Vol. 14, No. 1, pp. 97-106

A. Large, J. Beshesshti, A. Breuleux and A. Renaud, A comparison of information retrieval from print and CD-ROM versions of an encyclopaedia by elementary school students, Information Processing and Management, 1994, Vol. 30, No.4, pp.499-513

G. Marchionini, Information seeking in electronic environments, Cambridge University Press, 1995

K. Markey and P. Atherton, ONTAP: Online training and practice manual for ERIC database searchers, Syracuse University, ERIC Clearing House on Information Research, Syracuse, N.Y, 1978 (ED106109)

D.A. Michel, What is used during cognitive processing in information retrieval and library searching? Eleven sources of search information, Journal of the American Society for Information Science, 1994, Vol. 45, No.7, pp.498-514

P. Solomon, Children's information retrieval behaviour: A case analysis of an OPAC, Journal of the Society of Information Science, 1988, Vol. 44, pp. 245-264

D. Wendt Value of information for decisions, Journal of Mathematical Psychology, 1969, 6, pp 430-443

A semantic network-based phrase-prediction system for aphasic individuals

Nik Whitehead & Annalu Waller,
ACSD, Dept. of Maths & Comp. Sci., University of Dundee, Dundee DD1 4HN.

Abstract
Previous studies at Dundee have shown that people re-use phrases regularly in day-to-day conversation and that some of this conversation can be predicted. Our aim, therefore, was to design and evaluate a predictive computer system that can provide people who have communication difficulties with access to appropriate commonly-used phrases for conversational use.

The TalksBac system has been developed as a semantic net-based communication system for adults with dysphasia, a communication disorder resulting from brain damage, e.g. after a stroke. The system uses frequency and recency of selected phrases and topics, together with a series of weightings reflecting the relative closeness of topics and phrases to each other, to predict probable communication topics and phrases which the user can then select to be outputted through a voice synthesis system.

Two separate programs were required to use and maintain the semantic net. The TalksBac User program was designed to allow the dysphasic clients to communicate using phrases stored within the semantic net, while the TalksBac Carer program was designed to allow the clients' carers to input data to and edit the data within the semantic net. As both the user and carer groups are composed entirely of naive computer users, the design of the human-computer interfaces had to be designed with this in mind. The machines are currently with four clients, and preliminary results show that the semantic net-based system is working well.

1. Background
Analysis of language use has shown that large amounts of conversation are reused. It is therefore appropriate to design communication systems which use pre-stored narratives as well as simple phrases and phrases [Waller 1992]. The TalksBac system is designed as a communication system for adults with aphasia to allow them to communicate more naturally by exploiting this premise.

The current project, TalksBac 3 continues to refine both the interfaces and the retrieval system that were used in TalksBac 2 [Waller et al., 1994]. The programs are written in C++ for an Apple Macintosh laptop computer.

2. The TalksBac Interfaces
The purpose of the TalksBac software is to make the storage and retrieval of reusable phrases and stories as simple and as intuitive as possible. The system consists of a user and a carer interface, a retrieval system and a database of phrases and stories.

The carer interface is used to add, modify or delete phrases and stories from the database, while the user interface allows the user to retrieve and "speak" these pre-stored conversational items. A seperate carer interface is required as the users in this study have severe expressive problems, although they are able to recognise familiar words and simple phrases. These interfaces are intended to be as simple as possible, as they are intended for users with a) little or no prior computer knowledge and/or b) impaired reading abilities. The retrieval system employs a semantic network to identify which data is presented to the user for selection.

3. The User Interface
Figure 1 shows the main user screen from the TalksBac User Interface. The three large buttons bear topic titles, which the user may click upon to retrieve a list of phrases (phrases) and story titles associated with that topic. The four phrase boxes in the lower portion of the screen contain phrases associated with the current selected topic. Further buttons are for navigation about the screen, restarting and exiting the program.

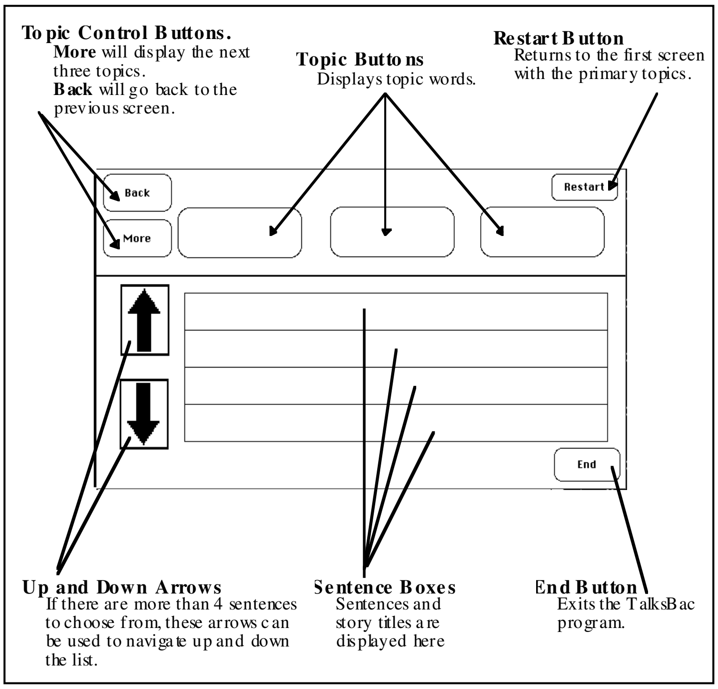


Figure 1: Main Screen Display

The user is initially presented with a set of primary topics (predicted from the user's previous use of the system) and selects one of these as the main area of conversation. Secondary topics associated with the primary topic are then predicted, and once one of these has been selected, the system predicts further topics and associated phrases from the user's personalised database. These phrases may then be 'spoken' using a speech synthesiser. Clicking upon a phrase will either speak the phrase using the computer's voice synthesis system or, in the case of a story title (shown by an ellipsis after the phrase), retrieve the phrases which comprise the story and can then be selected and spoken.

Figure 2 shows an example of the TalksBac 3 User Interface in operation.

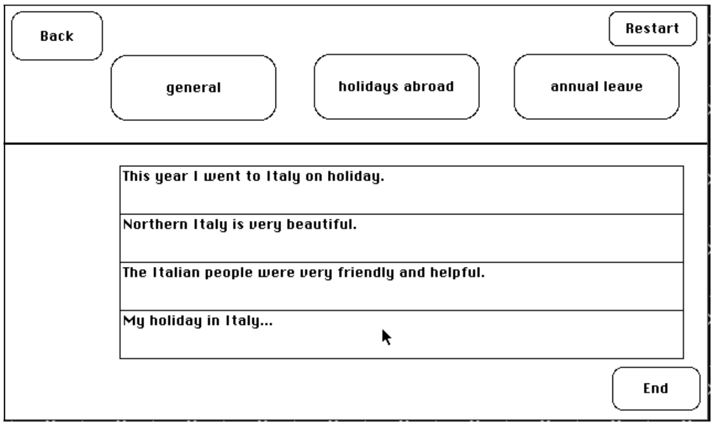


Figure 2: An example screen showing the user interface in operation.

4. The Carer Interface

The carer interface has two main sections. The first is a simple word processor allowing the carer to type phrases and stories for entry into the system, and the second section prompts the carer to enter the links between topics, phrases and stories.

Once the carer has entered the text of the phrase or story, he or she is then prompted to select a topic or topics with which the phrase or story should be associated. The phrase or story can be associated with more than one topic. All that the carer need do is click upon a topic to select it. It is these connection which provide the links of the semantic net.

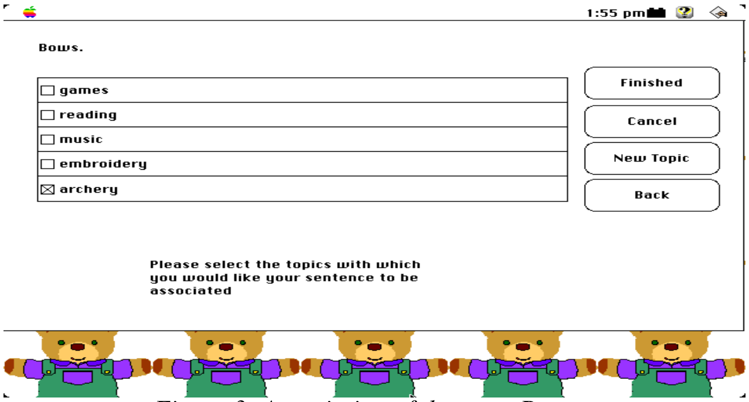


Figure 3: Association of the story Bows... with the topic archery using the carer interface.

5. The Retrieval System

The semantic net used within the TalksBac 3 package is a network of linked objects: Primary topics Px, Secondary topics Sx, Phrases Fx and Story Titles Tx. The prediction engine accepts the selection of either a topic or a phrase and produces the appropriate output of topics and/or phrases and story titles. The output topics are all secondary topics. The output of the selection of a phrase is that phrase spoken by the voice synthesiser.

IN	OUT
P_x	n x S_x
S_x	n x S_x and m x F_x or T_x
T_x	m x associated F_x
F_x	Spoken output of F_x

Figure 4: Summary of Input and Output requirements

6. Data Structures

There are two basic data structures: the topic and the phrase. The topic structure stores the name of the topic together with the ID, frequency, recency and weight assigned to all phrases or other topics associated with it. Flags mark the topic as being primary (associated only with other topics) or secondary (associated with topics, phrases and story titles).

The phrase structure can be used in one of two ways, to store either a single phrase, or to store a story title. In the case of a single phrase, the structure contains the phrase content, while for a story title both the phrase content and a list of associated phrases are stored. A flag is used to distinguish between these two uses of the structure.

Every phrase and story title is accessed by selecting topics. The way in which objects are linked is not static but changes over time. As each topic or phrase is selected its frquency, recency and weight are ajusted, thus allowing the semantic net to predict topics and phrases more effectively with increased use of the system. The procedure for determining the order in which topics and phrases are presented to the user is controlled by the prediction algorithm.

7. The Prediction Algorithm

The task of the prediction algorithm is to anticipate which items the user may wish to retrieve. The factors used to determine which items need be displayed to the user are: frequency (how often an item is used), recency (the current topic of conversation), and the association between topics (topic shift). These are updated every time a phrase or topic is selected.

The selection of a topic by the user results in a series of calculations which assign a value to each phrase and story title in the semantic network. The calculations are performed over three levels of topics, i.e. items may be

1. linked directly to the selected topic (Level 1)
2. linked to a level 1 topic (Level 2)
3. linked to a level 2 topic (Level 3)

Phrases are presented to the user in order of descending V, the value for which is given by:

$$V = \sum_{n=1}^{3} f_n r_n w_n \kappa_n$$

where
fn = the frequency of the item,
rn = the recency of the item,
wn = the weight (strength of association) of the item,
n = the level at which the calculation is being performed, and
(n = a value depending upon the level of the link. For level 1 links (n = 0.5, for level 2 links (n = 0.25, and for level 3 links (n = 0.05.

8. Current Status

At present both the user and carer interfaces are being evaluated by volunteer users and carers. Each of these users has a severe expressive communicative disability, and we hope to show that the TalksBac system improves not only the frequency of their communication, but also its quality.

The carers are drawn from the user's family and friends and have had training in the use of the carer interface and also in the selection and framing of appropriate conversational topics and phrases.

Although the evaluation period is only half-completed, it has already become apparent that one of the project's aims has been met. The current carer interface has led to the input of more data by the carers and a general improvement in database maintenance. Whereas at the end of the six-month evaluation period of TalksBac 2 none of the user databases contained more than 100 phrases, by the mid-point of the evaluation of TalksBac 3 the clients had over 200 phrases.

The use of the system by the users is being logged and will be analysed once the evaluation period is complete in the same way as way the TalksBac 2 software, in order to compare the two systems [Dennis et al. 1996, Waller et al 1996].

References

Dennis, F., Waller, A., Cairns, A.Y. & Brodie, J.K. (1996). *Evaluating TalksBac, A Predictive Conversation Aid, with Adults with Aphasia.* Communication Matters, In Press.

Waller, A. (1992). *Providing Narratives in an Augmentative Communication System*. Ph.D. Thesis, University of Dundee.

Waller, A., Miller, D., Dennis, F., Cairns, A.Y., Brodie, J.K., Gregor, P., Newell, A.F. & Morrison, K. (1994). *A Predictive Adaptive Conversation System for Non-Fluent Dysphasics.* RESNA '94, Nashville, Tennessee, 17-22 June 1994.

Waller, A., Dennis, F., Whitehead, N.J., Cairns, A.Y., Brodie, J.K., Gregor, P., Newell, A.F. & Morrison, K. (1996). *Evaluating TalksBac with Dysphasic Adults.* ISAAC '96, Vancouver, 7-10 August 1996.

How Many Relevances In IR?

Stefano Mizzaro
Department of Mathematics and Computer Science – University of Udine, Via delle Scienze, 206 - Loc. Rizzi, 33100 - Udine - Italy
E-mail: mizzaro@dimi.uniud.it   WWW: http://www.dimi.uniud.it/~mizzaro

1. Introduction

It is widely recognised [1] that relevance is a (if not 'the') central concept of IR. I show that: (i) there are *many* kinds of relevance, not just *one*; (ii) these kinds can be classified in a four dimensional space; and (iii) such classification helps us to understand the nature of relevance judgement.

2. Kinds of relevance

2.1. First dimension

It is commonly accepted that relevance is a relation between two entities of two groups. In the first group, we have one of the following three entities:

• *Document*, the physical entity that the user of an IR system will obtain after his seeking of information;
• *Surrogate*, a representation of a document, consisting of one or more of the following: title, list of keywords, author(s) name(s), bibliographic data, abstract, and so on;
• *Information*, the (not physical) entity that the user receives/creates when reading a document.

2.2. Second dimension

In the second group, we have one of the following four entities:

•*Problem*, that a human being is facing and that requires information for being solved;
•*Information need*, a representation of the problem (implicit) in the mind of the user. It is different from the problem because the user might not perceive in the correct way his problem [2,3];
•*Request*, a representation of the information need of the user in a 'human' language, usually in natural language;
•*Query*, a representation of the information need in a 'system' language, e.g. boolean.

These entities appear if one analyses the interaction between a user and an IR system (see Fig. 1): the user has a problem, *perceives* it and builds the information need, *expresses* the information need in a request, and *formalises* (perhaps with the help of an intermediary) the request in a query. Perception, expression, and formalisation are not so simple as it might seem at first glance, since some well-known problems (ASK, label effect, vocabulary problem, and so on, see [4]) appear.

2.3.Third dimension

On this basis, *a* relevance can be seen as a relation between two entities, one from each group: the relevance of a surrogate to a query, or the relevance of the information received by the user to the information need, and so on. Therefore, a relevance seems a point in a two-dimensional space. But these are not all the possible relevances, since two more dimensions have to be taken into account. First, the above mentioned entities can be decomposed into the following three components (third dimension) [5,6,7]:

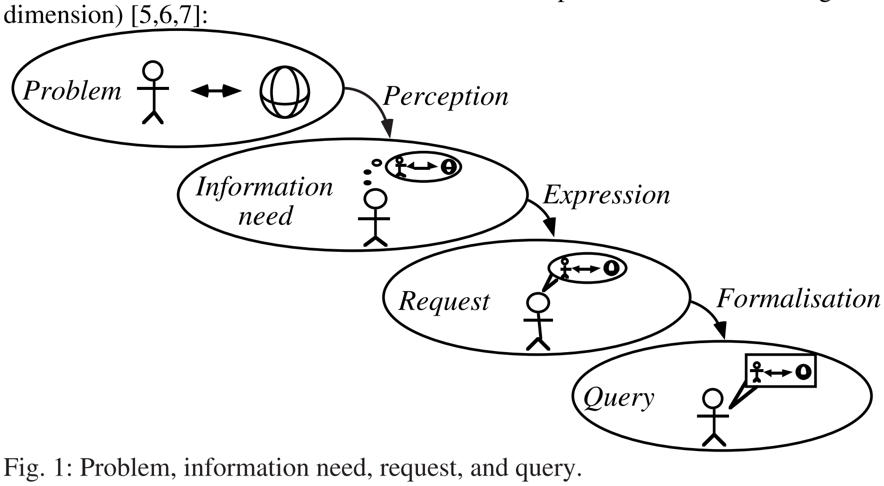


Fig. 1: Problem, information need, request, and query.

•*Topic*, that refers to the subject area to which the user is interested. For example, 'the concept of relevance in information science';

•*Task*, that refers to the activity that the user will execute with the retrieved documents. For example, 'to write a survey paper on ...';
•*Context*, that includes everything not pertaining to topic and task, but however affecting the way the search takes place and the evaluation of results. For example, documents already known by the user (and thus not worth being retrieved), time and/or money available for the search, and so on.

Therefore, a surrogate (a document, some information) is relevant to a query (request, information need, problem) with respect to one or more of these components.

2.4.Fourth dimension
The fourth dimension is the time: a surrogate (a document, some information) may be not relevant to a query (request, information need, problem) at a certain point of time, and be relevant later, or vice versa. This happens, for instance, if the user learns something that permits him to understand a document, or if the user problem changes, and so on. Thus the scenario represented in Fig. 1 has to be improved in order to take into account the highly dynamic interaction between user and IR system. In Fig. 2 the transformations of problem, information need, request, and query are illustrated. Four levels, represented by the four ellipsis, can be individuated, and refer to these four elements. At time t(P0) the user has a problem P0. The user perceives it, obtaining the initial information need (N0, at time t(N0)), he expresses it, obtaining the initial request (R0, at time t(R0)), and he formalises it, obtaining the initial query (Q0, at time t(Q0)). Then a revision takes place: the initial query may be modified (obtaining Q1, at time t(Q1)), the same may happen for the request and the information need, until the final information need (Np, at time t(Np)), request (Rm, at time t(Rm)), and query (Qn, at time t(Qn)) are obtained.

2.5.Relevance as a point in a four-dimensional space
Summarising, each relevance can be seen as a point in a four-dimensional space, the values of each dimension being:

1. Surrogate, document, information;
2. Query, request, information need, problem;
3. Topic, task, context, and each combination of them;
4. The various time instants from the arising of the user's problem until its solution.

The situation described so far is (partially) represented in Fig. 3: on the left hand side, there are the elements of the first dimension, and on the right hand side there are the elements of the second one. Each line linking two of these objects is a relevance (graphically emphasised by a circle on the line). It is more difficult to graphically represent the third dimension, since its elements are only partially (and not totally) ordered. In figure, the three grey levels represent the three components (induced in each of the relevances by the elements of the first two dimensions), emphasising that a relevance may concern one or more of them. For simplifying the figure, the time dimension is not represented. Finally, the grey arrows represent a partial order among the relevances. This order denotes how much a relevance is near to the relevance of the information received to the problem for all three components, the one to which the user is interested, and how difficult is to measure it. The number of steps (arrows) needed to reach the topmost small circle is an indication of the distance of each kind of relevance from the relevance of the information received to the problem.

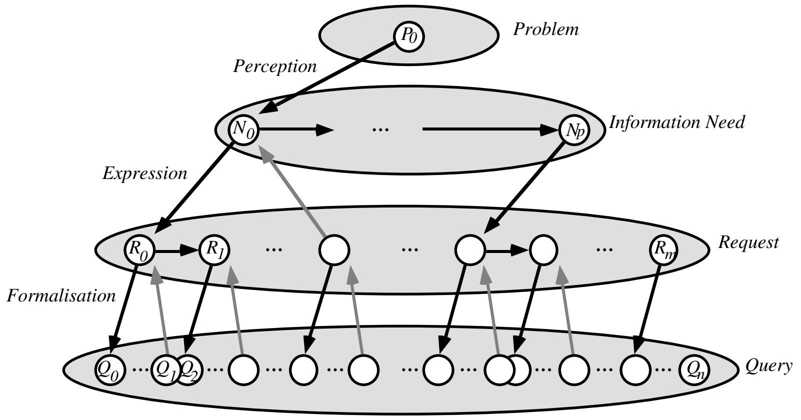


Fig. 2: The dynamic interaction user–IR system.

3.Relevance judgement
A relevance *judgement* is an assignment of a value of a relevance by a judge at a certain point of time. Similarly to what done above, it is possible to say that there are many kinds of relevance judgement, that can be classified along five dimensions:

1.The kind of relevance judged (see the previous section);

2.The kind of judge (for instance, it is possible to distinguish between user and non-user);
3.What the judge can use (surrogate, document, or information) for expressing his relevance judgement. It is the same dimension used for relevance, but it is needed since, for instance, the judge can assess the relevance of a document on the basis of a surrogate;
4.What the judge can use (query, request, information need, or problem) for expressing his relevance judgement (needed for the same reason of the previous point 3);
5.The time at which the judgement is expressed (at a certain time point, one may obviously judge the relevance in another time point).

4.Conclusions

The analysis presented in this paper:
•can easily be formalised using set theory, see [8,9];
•shows how it is short-sighted to speak merely of 'system relevance' (the relevance as seen by an IR system) as opposed to 'user relevance' (the relevance in which the user is interested), and how 'topicality' (a relevance for what concerns the topic component) is conceptually different from 'system relevance';
•is useful for avoiding ambiguities on which relevance (judgement) we are talking about;
•has to be considered in the implementation of IR systems working closer to the user; and
•emphasises that it is not so strange if different studies on relevance obtain different results, as one must pay attention to both which kind of relevance is measured and which kind of relevance judgement is adopted.

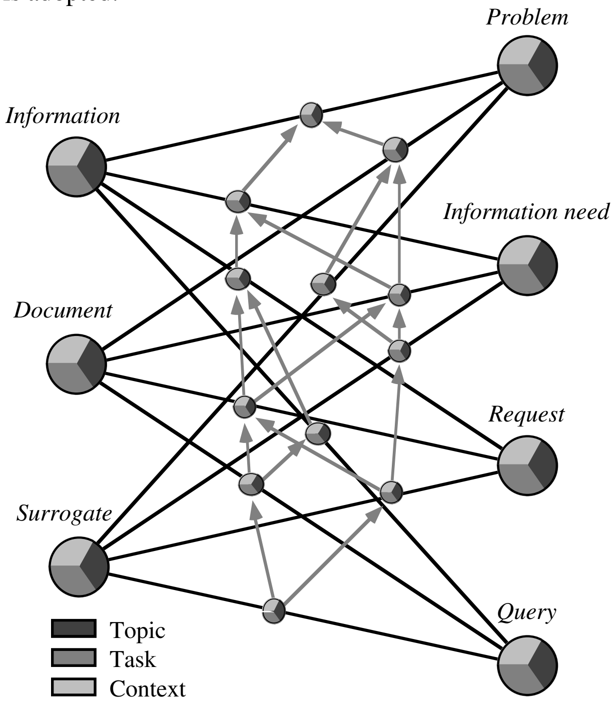


Fig. 3: The various kinds of relevance.

References

[1] S. Mizzaro, to appear. "Relevance: The Whole History". Accepted for the publication in *Journal of the American Society for Information Science*.
[2] S. Mizzaro, 1996. "A cognitive analysis of information retrieval." Accepted for the publication in *Proceedings of the CoLIS2 conference*, Copenhagen, October 13-16 1996.
[3] S. Mizzaro, 1996. "On the foundations of information retrieval." Accepted for the publication in *Proceedings of the Annual Conference AICA96*, Rome, September 24-27 1996.
[4] P. Ingwersen, 1992. *Information Retrieval Interaction*, Taylor Graham, London.
[5] G. Brajnik, S. Mizzaro, and C. Tasso, 1996. "Evaluating User Interfaces to Information Retrieval Systems: A Case Study on User Support". In *Proceedings of the SIGIR96*, Zurich, August 1996, pp. 128-136 .
[6] G. Brajnik, S. Mizzaro, and C. Tasso, 1996. "La valutazione di interfacce intelligenti per il reperimento di informazioni". In *AI*IA Notizie* (supplement to Nr. 3, Year IX, September 1996), pp. 36-37. (In Italian).
[7] G. Brajnik, S. Mizzaro, and C. Tasso, 1995. "Interfacce intelligenti a banche di dati bibliografici". In *Sistemi evoluti per basi di dati*, D. Saccà (editor), pp. 95-128, Franco Angeli, Milano. (In Italian).
[8] S. Mizzaro, 1995. "Le differenti relevance in information retrieval: una classificazione." In *Proceedings of the Annual Conference AICA95*, pages 361-368. (In Italian).

[9] S. Mizzaro, 1996. "Relevance: the whole (hi)story". Technical Report UDMI/12/96/RR, Dipartimento di Matematica e Informatica, University of Udine.

A Task-Based Approach To Usability Problems In Information Retrieval

Ann Doubleday, Michele Ryan, Mark Springett
HCID Centre, School of Informatics, City University, Northampton Square, London EC1V 0HB, UK.

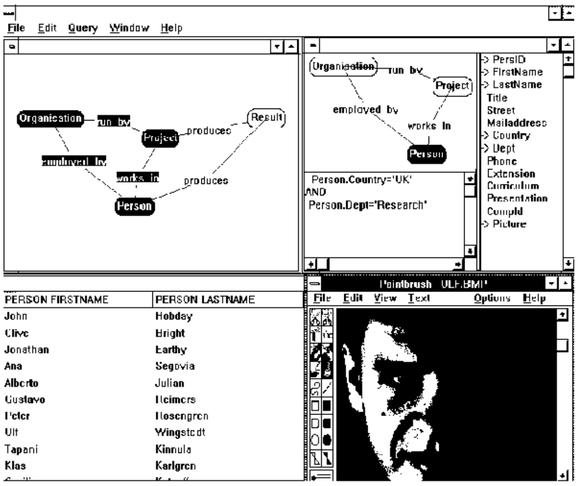
1. Introduction

The Esprit project INTUITIVE (INTeractive User Interface Tools In a Visual Environment) developed a support system designed to address many of the known usability problems in information retrieval (Belkin, 1982 and 1995; Fischer, 1991; Morris, 1994; and others). In this paper we report our studies on the usability evaluation of the INTUITIVE toolset and a directly comparable study on Microsoft Access, a widely used commercial product. We show how our analytical approach, using varied evaluation techniques, relating error-types to the retrieval subtask and seeking out underlying causes can identify generic usability problems and provide guidance for future designers.

The paper is divided into sections. First, we describe the applications under test and the steps involved in the information retrieval tasks. Next, we describe the methods used for collecting and analysing the data. There is then a section on the usability problems for each task step followed by a section on more general usability problems in information retrieval, and how the problems arise. Finally we show how looking for the cause, and not just noting the effect, of user difficulties can lead to design improvements.

2. The Applications

The INTUITIVE project was developed as a toolset to support information retrieval by graphical means. The four steps involved: (i) navigating the dataspace and choosing the entities to interrogate, (ii) forming the query by selecting and constraining attributes from the chosen entities, (iii) summarising the results of the query to enable the user to choose a subset for full viewing, and (iv) presenting the chosen items in full, were implemented as four tools in separate windows. The 'navigator' provides a map of the entities as nodes linked visually where appropriate by the relations between them, over which the mouse can traverse, selecting entities individually or collectively by clicking on entity or relation. Logical operators are dealt with implicitly by selecting linked (AND) or separate (OR) entities. The selected entities, with the relations between them, if any, are copied to the 'selector' window, where a mouse click on an entity displays a list of its attributes, enabling the user to select an attribute for retrieval (left hand side of list) or constraint (right hand side, and arithmetical operators). A textual rendering of the constrained attribute clause is displayed below the graphic. The 'results browser' shows a summary of retrieved results as thumbnails, if pictorial, or line descriptions if text. The 'presenter' presents the retrieved data in the most appropriate application, *.doc files as Word documents, *.bmp as Paintbrush, etc. It also arranges the windows on the screen for maximum visibility and spatially links the query with its results.



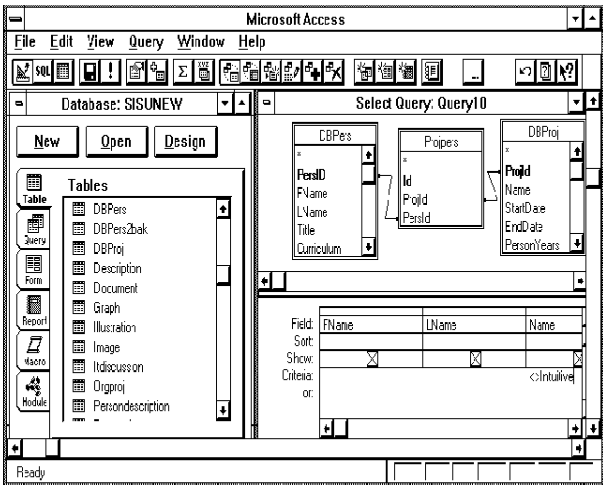


Figure 1. Intuitive Navigator, Selector, Browser and Presenter
Figure 2. Access Table list and Query Template

Microsoft Access is a commercial product, which although not requiring a knowledge of query (SQL say) syntax, still leaves the structure of the database explicit. The user selects entities from the list of tables in the database, and constrains them by typing in operators and values explicitly. Logical ANDs between clauses are made by joining tables; otherwise OR is assumed. Queries are submitted by a single '!' button or by the Run option on the Query menu. Results are in the form of a standard relational table and there is no screen planning.

3. Experimental Methods

Two methods of evaluation were employed: Nielsen’s heuristics (Nielsen, 1993) with HCI expert evaluators on the INTUITIVE interface only, and a second study on both products in which end users were videotaped and timed performing tasks on the toolset under control conditions (Preece, 1994).

3.1 Nielsen's Heuristics

Heuristic evaluation is intended to find the usability problems in a user interface design, in order to address them during the iterative design cycle. J Nielsen (1993) developed a set of heuristics to assess an interface according to recognised usability principles, such as users' language, consistency, and minimising memory load. He has shown (1995) that errors are found faster if evaluators are experts in the domain or in user interaction, maximally if in both. In principle, evaluators decide on their own how to proceed with evaluating the interface, but it is recommended that they go through the interface at least twice. The first pass is to is to get a feel for the flow of the interaction and the general scope of the system, the second is to focus on specific interface aspects.

In our case, we used five members of staff from the HCI design Centre at City University, most of whom had very little knowledge of the application. The recommended procedures were followed: after a familiarisation session including demonstration and explanation, they did a two pass supported walkthrough, employing our suggested usage scenarios at their own pace. The whole session took a little over an hour, but no time limits were applied.

3.2 User evaluation of INTUITIVE and Access

We were attempting to test the kind of problems users have in retrieving data, whether they are due to poor understanding of the task, faulty technical interface design or other factors. To do this we devised data access tasks of increasing technical complexity. It has been widely demonstrated that information retrieval is an iterative, exploratory process, so our experiments also tested how easy it was for users to modify and refine queries. The identical database was used for both interfaces, and as far as possible identical tasks were chosen.

Forty subjects, all undergraduate or postgraduate students, took part in the study. The subjects were divided four groups, equal numbers of SQL novice and experienced users on Access and the same on INTUITIVE. After completing a questionnaire on their level of computer, Microsoft, database and SQL experience, and being shown a short demonstration, each user was given a set of information retrieval tasks using a personnel database. The first three tasks required users to find straightforward information from the database, and the remaining tasks involved increasingly complex queries including zero results. Users worked at their own pace for up to 30 minutes. They were videotaped and asked to provide continuous verbal commentary on their thoughts and actions, following the protocol analysis procedure of Ericsson and Simon (1984). Each user then completed a satisfaction and usability assessment questionnaire, rating such items as flexibility, appeal, ease of avoiding and correcting errors, on 1-7 Likert scale, and a sample of subjects had a debriefing session to discuss the usability problems encountered. To elicit their post test memory of the task, users were asked, without looking at the system, to describe the steps needed to send a query to the database.

The quantitative data captured from the video tapes, observations and memory test included time taken to complete the tasks, number of tasks completed, and those correctly answered, error frequency, number of times user accessed help, and the task steps missed. Analysis of video data is always tedious and time-consuming. We found this method helpful. For each task step, the sequence (or set of alternative sequences) of expected user actions is tabulated. From the video a note is made in the appropriate box of any difficulties, repetitions, verbalisation or other points of interest. This can be done in a single pass and largely without pausing the tape, significantly reducing the tedium, and producing results which, together with timings per task and accuracy of retrieved results, produces an objective, accurate measure of usability problems across the subject population.

4. Findings

Quantitative results are to be published elsewhere (Sutcliffe et al., 1996). For this workshop we are interested less in whether one interface is better than another at some significance level, than in the *kind* of problems people have and how they originate.

From the questionnaires, users showed some preference for INTUITIVE over Access, especially for efficiency, flexibility and avoiding errors. The two products were *statistically* similar in overall performance and usability errors, (the only notable difference in task speed and correctness was that the experienced SQL-users on INTUITIVE had fewer errors and faster task completion times than the other groups) but this gives little insight into poor design features of either. The type of problems experienced can be considered under two heads, those within the design of INTUITIVE and Access applications and those more generally applicable to any information retrieval or indeed any windowing system. More interesting is variation within task steps over both applications. For this purpose, the task was decomposed into a series of steps, from creating the query, to evaluating the results and including query submission, modification and refinement. We will deal with the specific problems and questionnaire comments by task step.

4.1 Data visualisation and entity selection

As users are, in general, unaware of the contents of a database, INTUITIVE provides a map of the entities and relations, whereas Access shows a list of relational tables, though not joins between them. From the questionnaires, users liked the navigator design, and the video showed that they quickly learned from a training demonstration how to select entities, even though the Nielsen experts had found it unclear. Out of 25 noted INTUITIVE errors, only one applied to the INTUITIVE navigator, whereas a number of Access errors had to do with manipulating tables in order to select the right entity(s).

4.2 Query formulation - selecting attributes

To select from INTUITIVE's displayed list of attributes should be theoretically and practically helpful, but the manipulation of the list was difficult in a number of ways. Having to click on the left hand side to select and the right hand side to constrain attributes conforms to no standard convention, and many users showed by apparently aimless mouse wandering, repeated menu scanning and long pauses that they did not know how to select and deselect. The problem was exacerbated by the need to be very precise with the cursor and by poor feedback.

4.3 Query formulation - constraining attributes and joining the clauses

In INTUITIVE an operator (eg '=') is selected by point and click, the value then has to be typed in. In Access both the operator and the values have to be typed. Neither provides feedback for typing errors, and users did have problems with mistyping. Most Access users, even those with database knowledge, did not distinguish the inclusive from exclusive OR. In Access it is not obvious how a join is made or that it is needed, and no feedback is provided on invalid joins. In INTUITIVE this was done for them and they got the right answers, probably with no greater understanding.

4.4 Query Submission

In both systems submission from a menu was either an or the only option. This adds to memory load, was identified by experts as a probable source of confusion, and in fact did cause delay and error. As searching the database frequently causes visible delay and predicted confusion, it should provide some indication that the system is busy (and preferably how long it might take).

4.5 Retrieving and displaying results

In INTUITIVE, as an indication that the search is finished, a modal dialogue box appears, showing only the time the search has taken, not an indication of success. In both systems subjects repeatedly scanned the results, apparently unsure that they had successfully completed the query. Subjects had difficulty matching retrieved results to the appropriate query, and attempted to interact with the results window, probably due to the different window format between query and results tools. Users became disoriented when more than a very few windows were open at one time, even though INTUITIVE provides some presentation support.

4.6 Modifying queries

Many Access users were unable to work out how to modify a saved query, which involved using the ‘design’ button, even though this had been demonstrated. The terminology is not meaningful to the.users. Other problems included difficulty in joining tables, in order to extract related information

4.7 Windows manipulation problems

Where the application interface is clear and supportive, imperfect system features cause surprisingly little trouble, but as soon as the user model starts to falter, manipulative difficulty is enough to cause the user model to break down, as evidenced by the mouse wandering aimlessly round the screen or extended pauses requiring experimenter intervention.

In both systems, and for any task, the most confusing Windows feature in severity and frequency was the modal dialog box. Even experienced Windows users were baffled by the need to close the box before continuing. The next most serious problem was imprecise mouse pointing: clicking on the wrong (but close) hot spot produced unpredictable and uncomprehended effects, from which users had trouble recovering. A close third is the use of menus for essential steps in a task. Menu options are hidden at the point they are needed, and require the user, who is concentrating on the task not on its execution, to remember both a menu header and the option. This last problem is exacerbated by quirky, sometimes ambiguous, terminology and nomenclature and by apparently inconsistent greying out of options. Visibility was sometimes a problem, on account of lack of technical awareness of the working of scrollbars, change of window focus and moving and resizing windows. The randomness with which windows appear on the screen, in particular in INTUITIVE the Selector obscuring the Navigator, and results obscuring the query that produced them, causes considerable confusion.

5. Conclusion

The most severe and frequent cognitive problems in information retrieval tasks are caused when constraining attributes using logical and comparison operators. Maximum support should be provided in the form of conceptual maps of entity-relations, and also reducing manipulative and memory problems by visually linking attributes directly to the entity, providing appropriate operators and if possible values. The whole query should be visible, and as far as possible validated, before being submitted. Better feedback would have prevented many of the observed problems. It is a crucial principle of HCI, oft-repeated, but in our experience often ignored in commercial products, that every user action should trigger either success feedback, or an indication of what went wrong and how to correct it. This is at least as important in information retrieval - probably the most complex task in an everyday office environment.

Users encountered many manipulation problems, such as missably tiny hot spots, which tend to confuse results (and users). It should be possible to eradicate these at a very early stage in a short session with a single subject.

A number of problems identified by the expert evaluators were not experienced by end user subjects presumably because the training demonstration had adequately prepared the latter. This confirms that clear instructions are essential and suggests that problems might arise for the occasional user who might need a repeated training session, or more initial support. We would suggest that a small expert evaluation be performed in advance of the main testing, and the findings incorporated in the handbook or user guide. This guide should then be available during the user tests.

References

Belkin N J, Oddy R N and Brooks H, M. ASK for Information Retrieval: Part I, Background Theory. Journal of Documentation, vol. 38(2), 1982.

Belkin N J, Cool, Colleen, Stein, Adelheit & Thiel, Ulrich. Cases, Scripts, and Information-Seeking Strategies: On the Design of Interactive Retrieval Systems. In Expert Systems with Applications, Vol 9 (3), pp. 379-395, 1995.

Ericsson K. A. and Simon H. A., (1984), Protocol Analysis, MIT Press.

Fischer G, Henninger S. and Redmiles D, Intertwining Query Construction and Relevance Evaluation, ACM, 1991.

Morris R.C.T. Towards a User-Centered Information Service, Journal of the American Society for Information Science, vol. 45(1), pp. 20-30, 1994.

Nielsen, J. Usability Engineering. Academic Press. New York, 1993.

Nielsen, J. Getting Usability Used. Human-Computer Interaction, INTERACT. Chapman & Hall. 1995.

Preece, J. Human-Computer Interaction. Addison-Wesley Publishing Company, 1994.

Sutcliffe A G, Ryan M, Springett M and A Doubleday. Model Mismatch Analysis: Towards a Deeper Explanation of Users' Usability Problems. (City University, 1996).

Coordinated Support for Browsing, Searching and Monitoring : A User Interface for Networked Information Retrieval

Jonathan Furner*, David J. Harper** and David G. Hendry**
School of Information and Media , School of Computer and Mathematical Sciences**,*
The Robert Gordon University, Aberdeen, UK AB1 1HG
{j.furner,d.harper,d.hendry}@rgu.ac.uk

Abstract
User interfaces for accessing on-line collections of items tend to encourage information-seekers to either 'browse' or 'search'. This distinction is fundamental, and, when elaborated, it guides researchers in decision-making on topics as diverse as communication protocol standards, user-interface design, and empirical studies of information seeking. The goal of the research described here is to investigate how a user interface, or what we shall call a workspace, can provide coordinated support for both of these styles of access. Such a workspace needs to be cohesive, allowing people to shift between browsing, searching and associated monitoring tasks. A major design problem is to find ways for people to encode information on the workspace about what they plan to do, are doing, and have done. In short, the workspace must help people 'monitor' their progress. This point of view is elaborated through the discussion of a design study on how to harmonise 'browsing' and 'searching' on a direct-manipulation display.

1. Browsing and searching: Why couple them together?
The goal of the information-seeker may be expressed archetypically as:

> the retrieval, from a collection of items, of those items that collectively satisfy a particular need for information

Many activities can potentially be involved in satisfying, refining, or just better understanding an information need. A fundamental one, however, is what we shall refer to as *the articulation of an information need*. A distinction is commonly drawn between two styles or methods of articulation: 'browsing' and 'searching'. Browsing involves the successive retrieval of individual items, one item after another, on the basis of the existence of some relationship between each item. Searching involves the identification of sets of items on the basis of the existence of some relationship between a formal expression of information need (known as a query) and each item in the retrieval set. Although incomplete, this description of the difference between 'browsing' and 'searching' will serve for the discussion below.

Computerised information retrieval (IR) systems support information-seeking activity by automating the operations by which the relationships amongst queries and items are identified and represented to the user, and by which items are selected, retrieved and displayed. The details vary enormously from system to system, as does the extent to which support is provided for browsing or for searching. Hypertext systems, for instance, which enable relationships between items to be represented by links, on whose activation items are instantly retrieved and displayed, are typically far more supportive of browsing than traditional IR systems.

We have two research goals. First, to better understand the user-interface problems of supporting both browsing and searching. Second, to collect empirical data on how people shift from browsing to searching during the articulation of information needs. There are various possibilities for the design of hybrid systems that provide effective support for both types of activity. We are looking for a workspace that i) is simple; ii) has elements that differentially support both browsing and searching; and iii) maintains 'cohesiveness' such that an information-seeker may shift between browsing and searching activity easily. The design solution we propose is to create a workspace that gives prominent support for monitoring tasks.

2. Giving prominence to monitoring tasks
The articulation of information need, either through browsing or through searching, is just one of several different classes of activity in which the information-seeker might be engaged at any stage in the information-seeking process. This process is complex, dynamic and interactive. It often involves a number of different tasks competing for the information-seeker's attention. Making decisions about the ordering of tasks to complete or when to shift from one task to another can be difficult. Thus, the seeker needs to 'monitor', 'track' or 'record' their activity. Examples of monitoring tasks include: recording the order in which other tasks are carried out; monitoring the ways in which the objects (such as queries, targets, retrieval sets and retrieved items) involved in the search process are manipulated; and tracking the changes that occur in the characteristics of those objects and the relationships between them.

It is important that monitoring tasks are carried out effectively: the information-seeking process may otherwise suffer from error or unnecessary repetition on the part of the user. One problem faced by the information-seeker is that it is common for monitoring tasks to be poorly supported by existing user interfaces (Hendry & Harper, 1995); yet it would also seem that, in the networked environment, the success of information-seeking might depend even more crucially on the provision of effective support for such tasks. Any problems relating to the inability of the information-seeker to keep track of the characteristics of previous searches are likely to be exacerbated in a situation where the opportunity is given to issue queries simultaneously to multiple, heterogeneous targets.

Hendry & Harper (1995) argued that the execution of monitoring tasks can be well supported when searchers have direct, ongoing control of how the history of interaction is recorded. A direct-manipulation workspace allows searchers to compare queries or retrieval sets visually, side-by-side; to annotate particular searches; and to customise, according to personal preference, the shape, size, colour or location of the representations of particular objects. Other work also emphasises direct searcher control over the recording of the history of interaction in information seeking and organisation (Spink & Goodrum, in press; Shipman, Marshall & Moran, 1995). This design position is in opposition to systems in which the history of interaction is recorded automatically, such as Labrador (Rao, Russel, Mackinlay, 1993).

SketchTrieve is a direct-manipulation workspace that gives searchers control over ways in which their interaction is recorded (Hendry & Harper, in press). SketchTrieve enables information-seekers to formulate search queries, identify and connect to remote targets (which may or may not be Z39.50-compliant), execute searches and view retrieved items. Most significantly in the current context, it allows users to exploit space and layout in the planning and organisation of information-seeking activity, in just the ways described above. SketchTrieve is implemented using FireWorks, a general architecture for IR application development (Hendry & Harper, 1996).

3. Design study

The principal elements that searchers manipulate in SketchTrieve are called *services*. Services are used to represent a variety of objects involved in the information- seeking process, such as queries, remote servers, sets of references to retrieved documents, the contents of individual documents, etc. For example, *request* services expect some input (for example, a query) from the user, whilst *retrieval* services represent searchable collections stored on remote servers. The user may control the operation of each individual service by setting a variety of parameters in a customisation dialog. The user may copy, move or delete any service, and may save the entire contents of the workspace (i.e., the canvas) for later use.

Figure 1 shows a mock-up of a searcher's session using the SketchTrieve application.

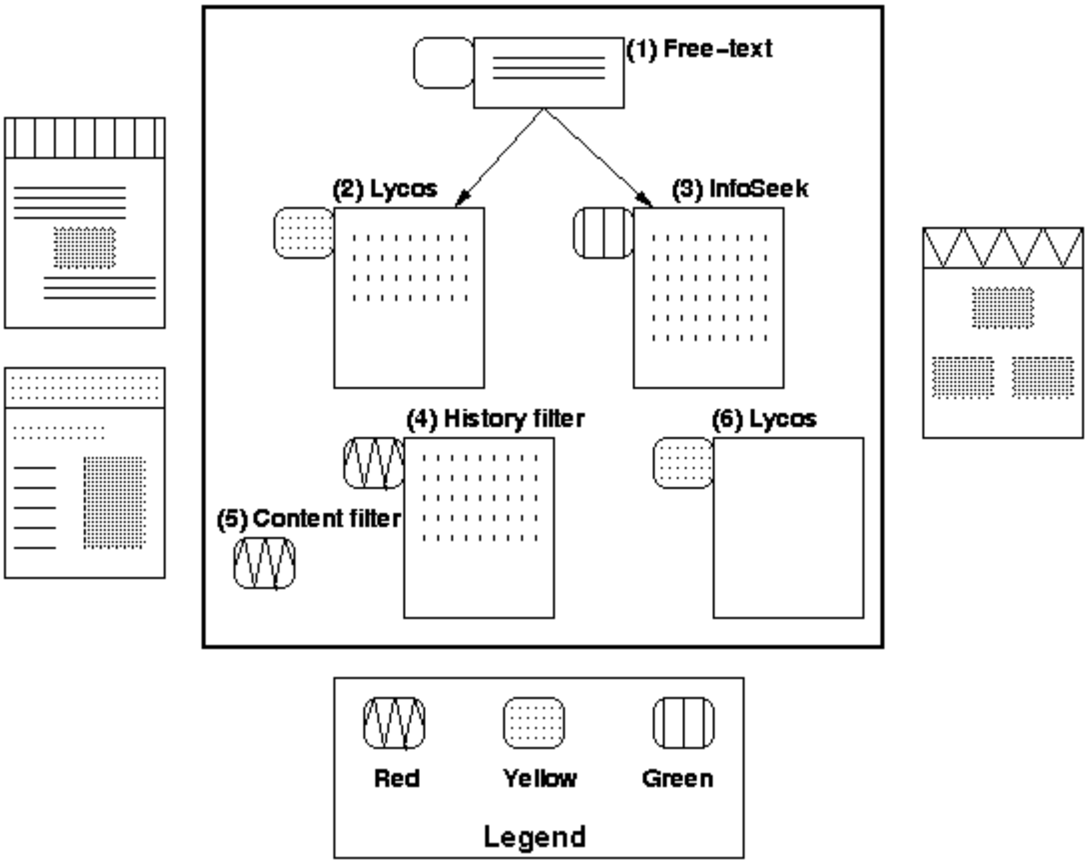


Figure 1. The SketchTrieve application. Shown in this mock-up is SketchTrieve with six different services on its canvas, and three WWW browsers.

In general, any request service may be 'wired' to any number of retrieval services with which the request service is compatible; the user's input to the request service is then directed to the remote servers represented by those retrieval services. In the mock-up, the information- seeker is working with six services. A single query has been input to the 'Free-text' service, which has been wired to two retrieval services representing Internet search engines ('Lycos' and 'InfoSeek').

In general, the output of a remote server (a list of references to those documents retrieved in response to the query) may be displayed in a *content view* associated with the appropriate retrieval service. In the mock-up, hypertext links to the Web pages retrieved by each search engine in response to the query are displayed in their respective content views.

The user may select any reference from such a list and the content of the document referenced will then be displayed in one of a number of *browsers* that have been associated with SketchTrieve by the user. A reference may also dragged from a content view and dropped either (i) at some point on the display, in order, for instance, to serve as a reminder for future plans, or (ii) on another service, such as a folder of saved documents.

In the mock-up, each of three browsers ('red', 'green', 'yellow') is associated with one or more services on the SketchTrieve canvas. Each association between browser and service is made through the service's customisation dialog, and is signalled through visual cues—in this case, a common colour. The functionality of these browsers is lightweight in comparison with that of a commercial browser such as Netscape: the user may view the content of documents, and use the links contained in viewed documents to navigate from one document to another, but little else.

Just as the user's searches that are carried out in the SketchTrieve workspace can cause documents to be displayed in the browsers, the user's browsing activity can be reflected in SketchTrieve. This is achieved through the use of a *filter*. This mechanism allows the user to capture in SketchTrieve references to whatever documents are viewed in the browsers. In the mock-up, for example, the 'red' browser is associated with two filters: the 'history' filter, which simply keeps a record of the last N documents visited; and the 'content' filter, which saves only those documents that match some expression—perhaps a regular expression or other type of profile.

4. Conclusion

In recent years, much attention has been paid to the problem of networked IR: i.e., optimising the retrieval of desired items from multiple, heterogeneous collections that are stored on server computers distributed across a wide-area network, such as the global Internet. There are several standard protocols that support the communications requirements of computers connected to the Internet. These protocols tend to be suitable for implementing user-interfaces that are either for browsing or for searching.

At present, we are designing a prototype that couples a HyperText Transfer Protocol (HTTP) client program, for browsing, with a Z39.50 client program, for searching (Eddie, Hendry & Harper, 1996). We claim that such a coupling is eminently desirable for at least two reasons: (i) HTTP is weakly supportive of searching, and Z39.50 is weakly supportive of browsing, yet a combination of the two would not suffer from either of these failings; (ii) use of Z39.50 is set to increase at an accelerated rate.

Our ultimate goal for the development of this hybrid browser/searcher is to allow information-seekers to coordinate more easily their networked information-seeking activity. We have already shown the potential value of the 'workspace' approach (Hendry & Harper, in press), but we have yet to explore the use of SketchTrieve in operational use. We intend to conduct an extensive study of how this prototype is used by real-life information-seekers. At the very least, we expect our work to enhance our understanding of information-seekers' behaviour and requirements in the networked environment, and to indicate improvements that might be incorporated in existing user interfaces. We believe that SketchTrieve shows how resources for information retrieval can be embedded directly in a users desktop.

Acknowledgements

John Eddie, a summer student, designed and implemented much of the code used in this prototype. This work is being carried out by the Information Retrieval Research Group, and is generously supported by the Principal's Research Fund at The Robert Gordon University and UBILAB, Union Bank of Switzerland. Additional information on SketchTrieve and FireWorks is available at: http://www.scms.rgu.ac.uk/research/ir/FireWorks/

References

Eddie, J., Hendry, D.G., & Harper, D.J. (1996*). The design of WebCoord: A system for browsing and searching the World Wide Web* (Tech. Rep. No 96/2). Aberdeen: The Robert Gordon University, School of Computer and Mathematical Sciences.

Hendry, D.G. & Harper, D.J. (1995). Coordinating information-seeking on interactive displays. In M. Collier & K. Arnold (Eds.), *Proceedings of the 2nd International Conference on Electronic Library and Visual Information Research*, pp. 127-136 London. Aslib.

Hendry, D.G. & Harper, D.J. (1996). An architecture for implementing extensible information-seeking environments. In H.-P. Frei, D. Harman, P. Schauble and R. Wilkinson (Eds.), *Nineteenth Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 94-100. ACM Press: New York.

Hendry, D.G. & Harper, D.J. (in press). An informal information-seeking environment. *Journal of the American Society for Information Science* (Special issue on Human-Computer Interaction).

Rao, R., Russel, D. M. & Mackinlay, J. D. (1993). System components for embedded information retrieval from multiple disparate information sources. *In Proceedings of the ACM Symposium on User Interface Software and Technology*, pp. 23-33.

Shipman, F. M., Marshall, C. C., & Moran, T. P. (1995). Finding and using implicit structure in human organized spatial layouts of information. *In ACM CHI'95 Conference on Human Factors in Computer Systems*, pp. 346-353. New York: ACM Press.

Spink, A. & Goodrum, A. (in press). A study of search intermediary working notes: Implications for IR system design. *Information Processing & Management.*

Speech for non-speaking people - how information retrieval and communication modelling can help

Dr Marianne Hickey,
MicroCentre, Applied Computer Studies Division, University of Dundee, Dundee, DD1 4HN, UK.
email: mhickey@mic.dundee.ac.uk

Abstract
This paper describes the application of models of spoken communication and speech production to voice output communication for non-speaking people. The models have been implemented, using text-based information retrieval, including a combination of rules, language statistics, and shallow methods of natural language processing. The resulting system, a PC based communication aid called WordKeys, allows the user to select pre-stored messages by entering one or more related words, where links between related words and messages are automatically generated. WordKeys has been evaluated in interactional and transactional speech encounters. The information retrieval approach taken has three key advantages: the physical effort required to speak a message is relatively low; the user does not need to remember exactly what messages have been pre-stored, or exactly how to access them; the user does not need to supply additional knowledge or descriptors when adding new messages - only the message text is needed.

1. Non-speakers and requirements for communication aids
Speech is our primary means of interacting with others; people make judgements about each other, form social bonds and learn about their environment via speech. About 5 per 1000 inhabitants - over 250,000 people in the UK alone - are physically or cognitively unable to communicate through speech or hand signs [1]. Many people with severe speech impairment are also physically disabled; often they can only be understood, and perhaps to a limited degree, by a few people such as family or care givers. Non-speakers are therefore limited in their opportunities for learning, personal development, a social life and work as a result of their impairments.

The current generation of communication aids do not meet the requirements of non-speakers, which include: easy and full of personalisation of the language which can be spoken by the communication aid; low physical and cognitive loads to allow fluent communication; and access to a wide range of utterances.

2. Applying models of speech production and spoken communication
The work described in this paper is based on Cheepen's description of spoken communication, [2], and Levelt's model of speech production [3, 4]. Most natural speech encounters consist of collaborative informal stories, which are built up in short turns taken by all participants. A communication aid, then, should allow a user to access messages suitable for short turn-taking, including messages that are topic initiators and continuations as well as responses. If we look at the natural speech production process, it seems that a suitable level to compensate for a severe speech impairment is before syntactic encoding [5]. At this level, key mental lexicon items are available, which a natural speaker would use to produce an utterance.

The author proposes that one way of being consistent with natural spoken communication and speech production, is to design a communication aid which helps the user to transform key words into a related message that is suitable for short turn-taking [5]. It is further proposed that the most effective way of implementing this thesis is with a text-based retrieval approach, where a combination of simple rules, statistical methods of information retrieval and shallow natural language processing is applied to a database of pre-stored messages [5].

3. WordKeys - word level compensation using text retrieval
A voice output communication system, based on an IBM compatible PC, has been developed to implement and evaluate these proposals. The system, WordKeys, predicts relevant pre-stored messages based on key words entered by the user. Links between messages and key words, which can be root morphemes, synonyms and other words related to message content words, are derived automatically. Retrieved messages are ranked according to some simple rules and knowledge of an individual's past use of the system. Automatic word to message links are made for most non-technical, non-specialised English words.

4. Evaluation
User evaluations showed that WordKeys can be used to achieve equal status in conversations and to achieve goals in transactional speech encounters. Subjects learned to use the system quickly and did

not need to know in advance the exact content of pre-stored messages or how to access individual messages [5].

5. Conclusion

The use of text-based retrieval techniques and language analysis brings three main benefits to voice output communication for non-speaking people:
(1) the message access method is flexible and intuitive;
(2) the user can speak an utterance with relatively little physical effort;
(3) the user or carer can personalise the system easily and fully.

References

[1] Lindstrom, J. I. (1989). Teleinformatics and disability. In R Oldenbourg (ed.) “Computers for handicapped persons.” Wien-Munchen.

[2] Cheepen, C. (1988). “The Predictability of Informal Conversation.” Pinter publishers.

[3] Levelt, W. J. M. (1989). “Speaking (From Intention to Articulation).” MIT Press.

[4] Levelt, W. J. M. (1994). “What can a theory of normal speaking contribute to AAC?” Proceedings of the Sixth Biennial Conference of the International Society for Augmentative and Alternative Communication. October 9-13. Maastricht, The Netherlands.

[5] Hickey, M. (1995). “Communication enhancement in an aid for severely disabled people”, PhD Thesis, Coventry University.

View-Based Searching Systems: A new paradigm for information retrieval based on faceted classification and indexing using mutually constraining knowledge-based views

A Steven Pollitt, Martin P Smith and Patrick A J Braekevelt
Centre for Database Access Research, School of Computing and Mathematics, University of Huddersfield, Queensgate, Huddersfield, West Yorkshire. UK. HD1 3DH

Abstract
View-based searching techniques employed in prototype HIBROWSE interfaces are exercising the principle recommendation made by the Classification Research Group that faceted classification should form the basis for all information retrieval systems. These interfaces present views onto databases using navigable hierarchies of subject descriptors from thesauri and grouped attribute values. Views inform the user of the number of documents resulting from inclusive searching of descriptors or attribute values and are mutually constraining. Each additional view provides an explicit constraint mechanism to modify the numbers of documents on all the views through automatic Boolean searching. Extensive search reformulation is achieved by the relatively simple interaction of refinement or expansion within a chosen facet or attribute value, or through the addition or removal of a view. The consequent tightening or relaxing of the constraints on retrieval is propagated through all the views. The result is an interface for information retrieval which exhibits increased searching and analysis power for the user without a commensurate increase in complexity or effort.

1. Introduction
The Classification Research Group, writing in 1955 [1], suggested that:
"The user may have a detailed understanding of the pattern of knowledge in the subject he explores, or he may have only an uncertain and confused understanding of it. An information retrieval system should be designed, first, to help even the ignorant user to pass from vague formulation of a subject in his mind to its precise formulation in the system; and then, having reached this precise formulation, to direct the searcher forward to literature references relating to it."

Developers of the HIBROWSE interfaces have been exploring ways of achieving this through employing view-based searching techniques making extensive use of thesauri, patterns of knowledge, in the interface. This work evolved from research that sought to implement an expert systems approach to document retrieval (CANSEARCH)[2,4], via a menu-based system which constructed multiple search statements (MenUSE) [5,6,7], to HIBROWSE, an interface which introduces a filtering paradigm for implicit Boolean searching using mutually constraining knowledge-based views organised into facets [7].

A research project, principally funded by the British Library Research and Innovation Centre, is examining how usability methods can be applied in the design and development of a HIBROWSE interface for the 6 million record EMBASE biomedical bibliographic database. An early evaluation can be found within these proceedings [10]. This brief paper describes the main features of the HIBROWSE interface and makes reference to the deliberations of the Classification Research Group [1].

2. The Impotance Of Thesauri and Classification
Problems of query formulation, indexing consistency and matching of terminology are well known and at the heart of information retrieval. The simplistic matching of keywords and an inability to refer to and relate concepts in querying are currently a major concern in the provision of search engines for searching the World Wide Web. The absence of organisation and structure facilitating reference to concepts and relationships was a concern of the Classification Research Group when examining the alternative searching mechanisms available in 1955 [1]. Much of the discussion focused on the difficulties experienced by users in identifying the exact subject of their search and the devices needed to overcome them. These difficulties include:
".. The searcher may use the correct concepts to express the subject but may not use the standard words used in the index ... may choose the correct terms, but may use only some of those which collectively describe the subject ... may choose concepts which do not correctly represent the subject sought .."

The solution to these problems was seen to consist of classified schedules of terms:
"Only one type of mechanism - the classified index - is in any position to deal with this problem. If to the serial list of numbers are attached verbal 'feature' headings, expressing the structure of the classification in words, the pattern of knowledge is made visible to the searcher. He may then often be led towards the correct heading."

The mechanism adopted in the major bibliographic databases (MEDLINE, EMBASE, INSPEC, WPI and many others) is the thesaurus. Using the thesaurus to its maximum benefit in the user interface has been the concern in a series of prototype developments outlined above. The most recent HIBROWSE interface embodies the main devices identified by the Classification Research Group but does so with the benefit of an online searching power and query manipulation capability unimagined by the members of that group.

3. The Interface Design

HIBROWSE interfaces present views onto databases using navigable hierarchies of subject descriptors from thesauri and grouped attribute values. Views are accessed by word entry and look-up in a wordlist or through browsing from the top level of a thesaurus facet and inform the user of the number of documents resulting from inclusive searching of descriptors or attribute values and are mutually constraining.

Figure 1. shows a subject view from the disease facet of the EMTREE thesaurus informing the user of the numbers of records concerning *degenerative disease* in a 300,000 record subset of EMBASE.

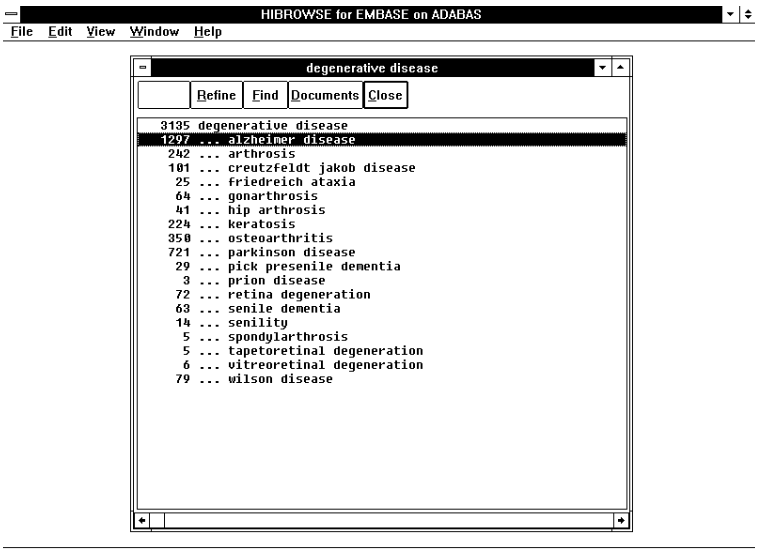


Figure 1. The view of *degenerative disease*

The user may choose to either expand the view, navigating up the hierarchical structure of the thesaurus to view more documents, or refine the view, moving down the hierarchy or selecting only part of a view (e.g. *Alzheimer disease* in Fig. 1).

Figure 2. shows the effect of adding a second view from the EMTREE facet which includes *therapy*. This additional view provides an explicit constraint mechanism to modify the numbers of documents on all the views through automatic Boolean searching. The numbers in the *degenerative disease* view are reduced from 3,135 to 1,074 excluding references which are not indexed by any of the terms in the *therapy* hierarchy of EMTREE. The therapy view alone would contain 91,633 references.

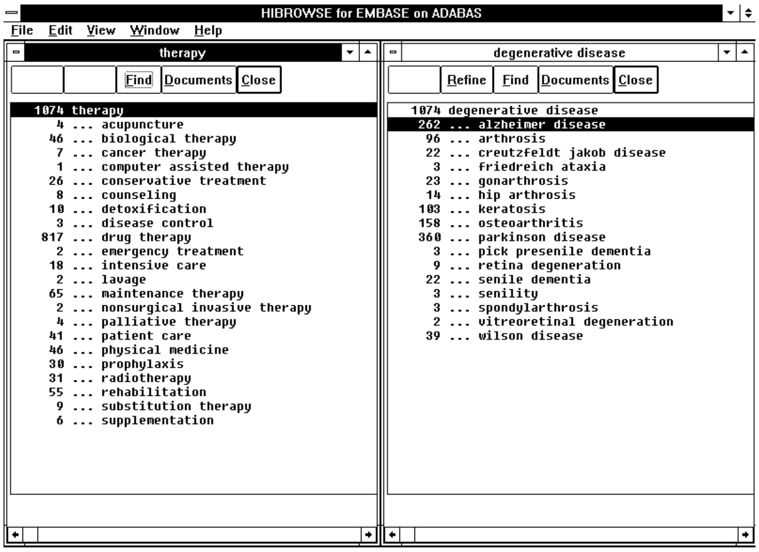


Figure 2. Mutually constraining views of *degenerative disease* and *therapy*

A refinement to drug therapy has a known outcome, and will reduce both views to 817 documents. What is unknown is the coincidence of drug therapy with specific degenerative diseases and the specific drug therapies, both of which are determined from a single button click in the therapy view having selected *drug therapy.* This result of this refinement is shown in Figure 3.

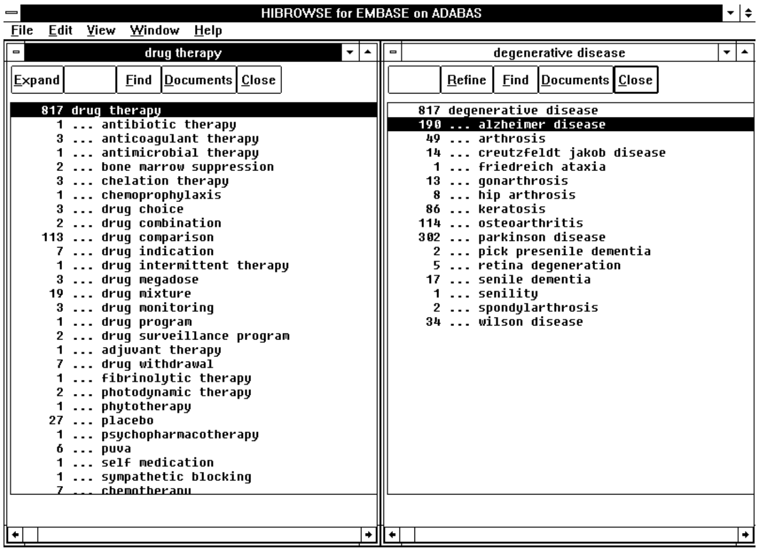


Figure 3 Refining the views to *drug therapy*

The result of further refinement to examine indexing coincidence concerning drug therapy of Alzheimer disease is shown in figure 4. This refinement significantly reduces the number of drug therapy descriptors. Alternatively, expansion of the concept *degenerative disease* would increase the number of documents being viewed.

The interface enables the user to examine coincidence of attributes that are not part of the subject indexing but which are recognised as valuable for discriminating within large document sets. These attributes include year of publication, journal, language of original article, institution and country of author. In addition to set refinement, viewing these attributes can also be seen as valuable for database analysis and bibliometrics. Figure 5 shows the country of Author for the 3,135 references on degenerative disease.

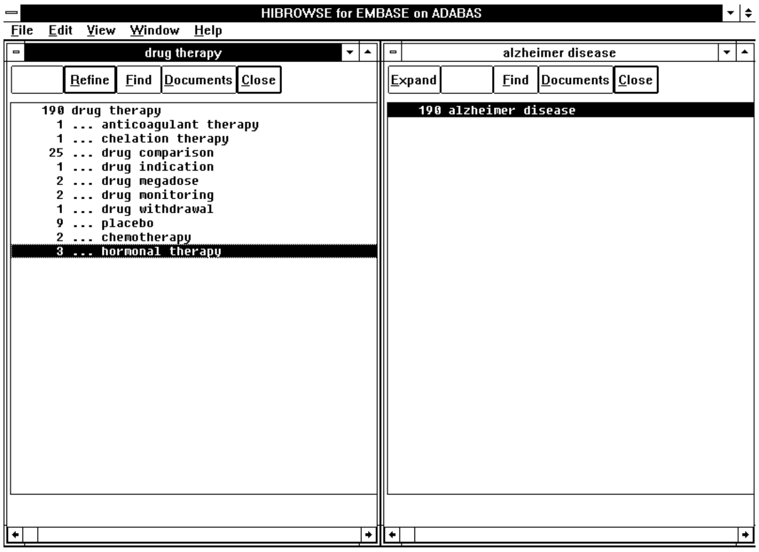


Figure 4 Further refinement to drug therapy and Alzheimer disease

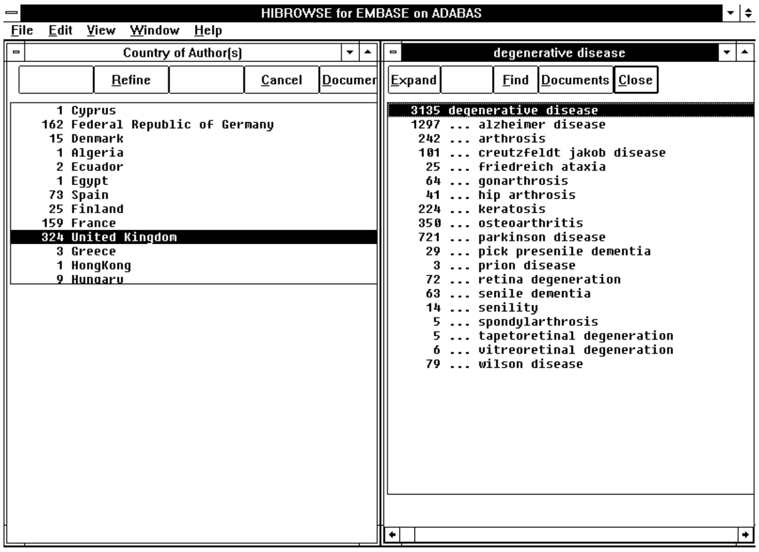


Figure 5. Views analysing degenerative disease references by country of author.

4. Multilingual Retrieval

The view based techniques described above lend themselves to multilingual retrieval. Given translations of the thesaurus there is no additional programming required to provide a system that can be used in the preferred language of the user. A system which demonstrates this is HIBROWSE for EPOQUE demonstrating access to the multilingual European Parliament Online Query System indexed using the EUROVOC thesaurus [following on from 6].

5. Application To Free Text And Corporate Databases

Although designed to exploit the use of controlled vocabulary indexing, view-based searching techniques are applicable for searching unindexed datasets using a thesaurus to generated free text search statements to exploit concept-based inclusive term searching. These techniques are equally applicable to corporate databases using attribute value summaries that can include hierarchical presentations or simple value range aggregations. The first HIBROWSE interface was in fact developed for an ORACLE held database of UK Hotel Groups.

6. Evaluation

Experiments are being designed to investigate usability issues seeking to apply techniques from the MUSiC methodology with a number of different types of user. The types of HCI techniques being applied are the User Action Notation, Cognitive Walkthroughs [9,11] and others from usability methods, mainly MUSiC [12].

Acknowledgements
The HIBROWSE for Bibliographic Databases project is funded by the British Library Research and Development Department (Grant RDD/G/282) and co-sponsored by Software AG of the UK, Sun Microsystems, Elsevier Science BV, STN-International (FIZ-Karlsruhe), Elonex and the University of Huddersfield in collaboration with the University of Leeds.

CONTACT DETAILS:
Dr A Steven Pollitt
Director, CeDAR - Centre for Database Access Research
Tel: +44 (0)1484 472147 Fax: +44 (0)1484 421106
email: a.s.pollitt@hud.ac.uk

REFERENCES

[1] CRG (1955) The Need for a Faceted Classification as the Basis of all Methods of Information Retrieval. A memorandum of the Classification Research Group, published as UNESCO document 320/5515, 26 May 1955.

[2] Pollitt A S (1981) An expert system as an online search intermediary Proceedings of the 5th International Online Information Meeting pp 25-32, Learned Information, London, December 1981

[3] Pollitt A S (1985) End user Boolean searching on viewdata using numeric keypads Proceedings of the 9th International Online Information Meeting, pp 373-380, Learned Information, London, December 1985

[4] Pollitt A S (1987) CANSEARCH: An expert systems approach to document retrieval Information Processing and Management, 23 (2), pp 119-138, 1987

[5] Pollitt A S (1988) A common query interface using MenUSE - A Menu-based User Search Engine Proceedings of the 12th International Online Information Meeting, Vol. 2, pp 445-457, Pub: Learned Information, London, December 1988

[6] Pollitt A S, Ellis G P, Smith M P, Gregory M R, Li C S & Zangenberg H (1993). A common query interface for multilingual document retrieval from databases of the European Community Institutions Proceedings of the 17th International Online Information Meeting . pp 47-61. Pub: Learned Information London, December 1993

[7] Pollitt A S, Ellis G P, Smith M P (1994) Peek-a-Boo revived - End-user searching of bibliographic databases using filtering views. Proceedings of the 18th International Online Information Meeting, pp 63-74, London, December 1994.

[8] Pollitt A S, Ellis G P, Smith M P (1994) HIBROWSE for Bibliographic Databases Journal of Information Science, 20 (6), pp 413-426, December 1994.

[9] Treglown M (1996) Results of applying the Cognitive Walkthrough Method to the HIBROWSE for EMBASE system. Research Report RR96/02 School of Computing and Mathematics, University of Huddersfield.

[10] Treglown M, Finlay J E (1996) Usability Inspection and Evaluation of Interactive Information Retrieval Systems. Proceedings of the Information Retrieval and Human Computer Interaction Workshop, Glasgow, September 1996.

[11] Treglown M, Finlay J E (1996) Driving Design with the Cognitive Walkthrough. Ancillary Proceedings of the 11th BCS Conference on Human Computer Interaction - HCI '96, London, August 1996.

[12] Macleod M (1994) Usability : Practical Methods for Testing and Improvement. Proceedings of the Software '94 Conference, Scandvika, Norway, February 1994.

Applying Ostensive Functionalism in the place of Descriptive Proceduralism: "The query is dead"

Iain Campbell
Computing Science, University of Glasgow, Glasgow, G12 8QQ.

iain@dcs.gla.ac.uk

In this short report I outline work that makes a case for the abandonment of queries in approaches to information access. I first highlight problems with queries. I then point out the complementary nature of query-based and browse-based approaches. I describe the "Ostensive Model" that forms the core of a new approach combining the two traditional approaches. It replaces the current reliance upon "Descriptive Proceduralism" with one of "Ostensive Functionalism". I believe the hybrid has potentially important media- and language-independence characteristics that make it ideal for a new generation of Information Retrieval systems.

1. Current approaches

Within interactive IR I see two main classes of approach. Both represent fundamentally different approaches to exploration of an information space by the user - that of query-based and that of browse-based. This division corresponds, in effect, to the reliance upon describing information needs.

1.1 Query-based

I regard as a query-based system any one that requires of the user the production and refinement of a description of their information need. Broadly speaking, this description encompasses the vast majority of current IR systems - both commercial and research.

The most prominent examples of query-based systems are the Boolean systems offered by commercial dial-up online information providers - eg *STAIRS* [Ibm72], and its derivative *BRS-Search*. Other Boolean systems available for most personal workstations and bundled with CD-ROM information bases - eg *Topic* [Leh94], *Personal Librarian*. The most modern of the query-based systems are the weighted systems that are generally restricted to the research world: *ConQuest* [Nel94], *NRT* [San91], *WAIS* [Ste91], *Inquiry* [Cal92], *Smart* [Sal71], *Okapi* [Rob93].

What all these systems have in common is that their operation centres around a query. A query is a user-generated description of an information need, expressed in terms of some system supported language. These languages are direct derivatives of the internal structures used by the system to represent the stored information objects. For example, the presence or absence of individual words along with Boolean operators such as 'AND' and 'OR', constitute Boolean query languages.

In practice, a user generates an initial query, the query is evaluated by the system, a list of resulting documents is presented to him, he inspects these documents, refines his query, the new query is evaluated by the system, and the process continues iteratively until he reaches a point of satisfaction. This might be when he believes that he has found the information required, believes that there is no more information to be found, or believes that he is unable to find any more using the system.

Users must describe the documents they believe will satisfy their information needs by describing the pattern of presence or absence of features within them. This must be done without having seen the documents in advance. Expecting a user to be able to describe something he doesn't have does not strike me as a very promising approach. Placing the restriction that such descriptions must be made in terms of an artificial language, is to further hinder a user satisfying his needs. Query languages have the appearance of being designed primarily for their computational simplicity and correspondence to the internal representation schemes, and not for their suitability to convenient and effective expression of information needs.

Weighted retrieval systems (eg *Smart* and *NRT*), have slightly different internal representations and significantly different retrieval algorithms to that of the Boolean systems. Here, the 'hard' nature of Boolean operators on word occurrences is replaced by 'soft' weights attached to candidate words giving more informal, forgiving languages in which to express information needs. Nevertheless, the reliance upon description, and the directness of correspondence between the query languages and the internal representations remain.

The representations are also used by the systems to 'explain' particular retrieval decisions. The words that caused the retrieval of particular documents are often highlighted in some way, thus providing feedback to the user on the extent to which each of the query words is involved in the retrieval of documents. This explanation can be crucial to the effective exploitation of such systems. It is with such information that users make their decisions upon how to refine their query and thus continue their search.

Several iterations of query-refinement and query-execution are likely to be required before the user will be happy with, or confident in, the query as a representation of his information need. In the meantime, the very information need that he has been endeavouring to represent will have changed ‘beneath his feet’.

Query-based systems support the development of queries. This *allows* information needs to develop and the user to follow that development. Nevertheless, they do not *support* that development - the requirement for the generation of and refinement of descriptions serves only to hinder it.

1.2 Browse-based

In a browse-based system there is no query, and no explicit description of the user’s information need. The user furthers his search purely through selections or indications on information objects presented to him by the system. Such systems almost exclusively employ a graphical interface presenting documents as objects with links between them.

The classic browse-based IR system is I^3R [Tho89] - where documents and groups of documents are presented as icons with accessibility links between them. *Transient-Hypergraphs* [She91] exemplify the application of the browse-based approach to a traditional strongly-typed database. *TACHIR* [Ago94] is a system that allows the user to browse across a plane of documents and their accessibility relationships, across a similar plane of terms, and across the accessibility relationships between the various planes. *Galaxy of News* [Ren94], and *Bead* [Cha92], present documents in an approximately 2-D space that has been dimensionally reduced from an original higher dimensional feature space.

The accessibility links can be human generated, derived from structure that might be present in the object space, or derived from the content semantics of the objects concerned. They are commonly present between every object and every other object in the space. The strength of the links (ie the degree of accessibility) can be binary valued, multi valued, or lie on a continuous scale.

In order to move around within the information space, a user makes a selection from a list of links. These links are to the objects that are deemed to be the most accessible. This accessibility relationship may be anything from a simple mutual set membership, to a sophisticated measure of the semantic similarity of two objects. Usually it is a pseudo-semantic similarity measure based upon one of the classical retrieval models. Such measures typically produce fully-interconnected graphs. Nevertheless, it is not practical to present them all to the user - only a small number of the strongest links are presented at each object.

1.3 Comparison

Browsing overcomes the drawback of requiring queries. Instead, a user accesses the relevance of objects and nearby objects and traverses links between them, never having to describe the thought processes involved or the changes in information need resulting from those assessments and exposure to information. Associated with not requiring descriptions, browsing does not expose a user to internal representations, or surrogates of them, and it does not require a user to manipulate them - all such matters can be completely hidden.

The one area where Querying outshines Browsing is that of its dynamic ‘view’ of the information space. As a query changes, the structure of the space changes also. In that way it can respond (although somewhat indirectly) to individual information needs or changes to them that occur during a searching session. Browsing, in contrast, has an insensitive static structure, to which a user must adapt.

Putting the primary task of a searching user (ie the identification of, and exposure to, relevance) to one side, both approaches have drawbacks: Querying requires significant cognitive effort in steering the system towards the desired information through queries; Browsing requires effort to determine where relevant information might be located in the network of objects and links. Although this ancillary effort (ie that not directly related to the primary task) is different in the two approaches, it could be argued that the amount is approximately equivalent.

Query- and browse-based approaches could be said to be examples of *descriptive proceduralism*: Descriptive, in their reliance upon queries; Procedural, in their reliance upon users performing non-trivial unrelated tasks in order to effect their ultimate intentions. Both approaches require the translation of an information need into actions, plans of action, and strategy.

2. An ostensive hybrid

I regard the nature of the two traditional approaches as complementary . The work I am carrying out centres around a hybrid of the best features of the two. Essentially, it is a browse-based approach (ie graphical presentation of information objects and links between them), with the addition of dynamic views of the information space.

The dynamic views are achieved by restructuring the space at each step made by the user. This is done in a manner intended to be sympathetic with a user's changing information need. The view presented at an object is based upon the path taken to reach that object. Evidence is gathered from each of the objects in the path, and combined into a representation of the current information need. That representation is used to determine the objects that are made available to the user as candidate next steps.

The use of evidence from the path is motivated by the assumption that each of the path objects were chosen as a result of their relevance at the time of being chosen. The path is regarded as a history of relevant-indicated objects - a history that embodies the development of an information need. The path is considered to be an *ostensive definition* of the information need.

> *ostensive definition* (Philos.), the explanation of a word by presenting, pointing at, or otherwise indicating one or more objects to which it applies. [Oed93]

Ostension can be a useful technique for the capture or identification of spatio-temporal concepts such as relevance. Spatial evidence can be collected from the internal representations of the objects. Temporal evidence can be collected in the form of the relative time at which the various relevance indications were made. The *Ostensive Model* [Cam96] has a view of the relationship between relevance indications and the development of an information need. In essence, old relevance indications are less indicative of the *current* information need than more recent indications.

This results in an ostensive definition being influenced by all indications of relevance, but more so by recent indications. As a user browses, all existing evidence ages with each step. It is in that way that a system based upon this model can dynamically follow the developing information need of a user.

2.1 Multiple media, multiple languages

Information stored in a non-textual medium, by its very nature, requires different representation techniques to that of text. An example of initial attempts at providing access to such data is that of speech-only audio information being transcribed into text, and then that text being searched in the traditional manner. Music, photographs, video, etc cannot be given anything but the most trivial of textual 'transcriptions'. I believe that we will, by necessity, develop new representation techniques and similarity measures for non-textual media. These are unlikely to be accessible to humans. For example, one can imagine using neural networks, where the representation would consist of thousands of subtly different probabilities in complex networks. As a result, I believe the reliance of the query-based approach upon a user's exposure to, and manipulation of, internal representations will prevent it from being applicable to the new media.

The ostensive approach appears more promising in that it allows dynamic access to collections of information objects without any of the representation issues being visible to a user. Images will be accessible through the use of an opaque image to image similarity technique that can nominate the candidate next steps. Similarly, for audio, video etc. Further, the availability of, for example, audio to text or audio to image similarity techniques would allow heterogeneous collections of objects to be searched. The user would experience objects of a similar *theme* (ie regardless of their media type) being located close to each other. Further, as they browsed in a particular 'direction' the themes of the objects presented to them would become more coherent.

In a similar manner to that of media, multiple languages in, for example, a text or audio collection could be made accessible to users with greater capabilities of reading and/or listening than of speaking and/or writing in the languages concerned. As for the case of multimedia objects, the user would experience objects of a similar *theme* being located close to each other (regardless, in this case, of their language). Again this would require only that inter language similarity techniques are available to the system.

3. Current and further work

The background intuitions, a fuller discussion of the query/browse contrast, and the expected look and feel of a system based upon ostension can be found in [Cam95]. A more specific and detailed account of the cognitive model, the aging of evidence, and its incorporation into the probabilistic retrieval model to produce a complete implementable approach can be found in [Cam96]. I regard the approach as an example of *ostensive functionalism*: Ostensive, in its evidence gathering techniques (ie non-descriptive); Functional, in that it requires of the user only that he perform his primary task (ie that of the identification of relevant objects), and that he need not concern himself as to where such objects might be or how to get to them.

The work has now two distinct tracks. Firstly, a practical track where a prototype has recently been constructed. Initial informal experiences have been positive with a test collection of 800 photographic images with annotations (in French). Formal evaluation of the interface and of the approach in general is now being started. This will be carried out with groups of users charged with performing a variety of tasks using an IR system as a tool. The performances of groups using the ostensive system will be compared to that of groups using querying, and to that of groups using browsing systems. By setting a variety of tasks, it is hoped that both appropriate and inappropriate tasks will be identified.

The second track is theoretical. The ostensive approach appears to have philosophical probabilistic implications regarding the relationship between ampliative induction and ostension that go back to the work of 17th and 18th century philosophers such as Pascal, Bacon, and Hume. My particular use of ostension appears to overcome a traditional perception of incompatibility between variative and enumerative induction.

The arguments concerning the drawbacks of both query-based and browse-based approaches to IR, and the promising characteristics of the ostensive approach lead me to believe that it is the arrival of multimedia information that heralds the departure of the query.

4. Acknowledgements

This work has been supported by the ESPRIT Basic Research Action "FERMI - Formalisation and Experimentation on the Retrieval of Multimedia Information" [Fer94].

[Ago94] M. Agosti, M. Melucci, and F. Crestani. *TACHIR: a tool for the automatic construction of hypertexts for Information Retrieval* in proc. of RIAO 94, Rockefeller University, 1994, pp. 338-357

[Cam95] I. Campbell. *Supporting information needs by ostensive definition in an adaptive information space* in proc. of MIRO '95, Workshops in Computing. Springer Verlag, 1995.

[Cam96] I. Campbell, and C. J. van Rijsbergen. *The ostensive model of developing information needs* in proc. of CoLIS-2, Copenhagen, October 1996. (in press)

[Cal92] J. P. Callam, W. B. Croft, and S. M. Harding. The INQUIRY retrieval system *in proc. of 3rd Int. Conference on Database and Expert Systems Applications* 1992, pp 78-83

[Cha92] M. Chalmers, and P. Chitson. *Bead: explorations in information visualisation* in proc. of ACM-SIGIR, 1992, pp. 330-337.

[Fer94] The FERMI project. ESPRIT Basic Research Action 81034. Documentation and technical reports Web server. "http://www.dcs.gla.ac.uk/fermi".

[Ibm72] *Storage and Information Retrieval system (STAIRS) - general information manual* 2nd edition, IBM Germany, Stuttgart, 1972

[Leh94] J. W. Lehman, and C. A. Reid. *Knowledge-based searching with TOPIC* in proc. of The second Text REtrieval Conference (TREC-2), 1994, pp 209-221

[Nel94] P. E. Nelson . *The ConQuest system* in proc. of The second Text REtrieval Conference (TREC-2), 1994, pp 265-270

[Oed93] *Oxford English Dictionary* CD-ROM, 2nd Edition, Oxford University Press, 1993

[Ren94] E. Rennison. *Galaxy of news. An approach to visualising and understanding expansive news landscapes* in proc. of ACM-UIST '94, 1994,

[Rob93] Robertson, et al. *Okapi at TREC* in proc. of The first Text REtrieval Conference (TREC-1), 1993, pp 21-30

[Sal71] G. Salton (editor) *The SMART retrieval system - experiments in automatic document processing* Prentice-Hall, 1971

[San91] M. Sanderson & C. J. van Rijsbergen. *NRT: news retrieval tool* Electronic Publishing origination dissemination and design, Vol. 4, Num. 4, Pages 205-217, 1991

[She91] M.A. Shepherd, and C. Watters. *Virtual structures for hypertext* in proc. of ACM SIGIR / ISS International Conference of Multimedia Information Systems, pp. 201-219,1991

[Ste91] R. M. Stein. *Browsing through the terabytes* Byte, Pages 157-164

[Tho89] R. H. Thompson. *The design and implementation of an intelligent interface for Information Retrieval* Ph.D., University of Massachusetts, 1989.

Characteristics of users'needs and activities: A design space for interactive information retrieval systems

Jean-Pierre Chevallet and Laurence Nigay
CLIPS-IMAG BP 53 , 38041 Grenoble Cedex 9 , France
Phone: (+33) 76 63 57 23, (+33) 76 51 44 40
E-mail: {Jean-Pierre.Chevallet, Laurence.Nigay}@imag.fr

Abstract
One of the new design challenges in Information Retrieval (IR) and Human Computer Interaction (HCI) is to face a huge quantity of structured and accessible information which is available to a variety of users having different information needs. This article proposes a dimension space that should help in the design of current and future interactive information retrieval systems. To do so, we adopt a user centered perspective: our problem space is comprised of 3 sets of dimensions. The two first ones are used to characterize the user: her/his information needs and her/his knowledge about the system. The third set of dimensions deals with the notion of input expression: how the user specifies her/his information need. Based on our design space and Norman's Theory of Actions we then model users' activities while accomplishing a retrieval task. The discussion is illustrated with IRIS-M, an Interactive Information Retrieval System that we have developed.

1. Introduction
An Information Retrieval System (IRS) allows various users to search and retrieve information. One of the new design challenges for Information Retrieval (IR) and Human Computer Interaction (HCI) is to face huge quantities of structured and accessible information which is available to a variety of users having different information needs. IRS are inherently interactive systems with a constant growth of the number of end-users. Here we underline the key role that Human Computer Interaction (HCI) plays in an IRS. IRS must not only rely on effective retrieval mechanisms but must also provide efficient interaction with the end-users. Nevertheless little attention has been paid to the user interface in the design and evaluation of an IRS. For instance the evaluation of an IRS is based on recall/precision metrics which are computed according to a set of predefined tests. The user activities are not taken into account in the evaluation process. We address this need in our work: the global goal is to incorporate the user, characteristics and activities within the design process of an IRS. To do so we apply design methods from the domain of HCI. Design methods are based on a conceptual study of the user, her/his information needs and tasks. Results of this conceptual study are used to design the user interface. Finally the designed user interface is developed and evaluated.

This article focuses on the user's need and activities while accomplishing a retrieval information task. We propose a framework that should help in the design of current and future IRS. To do so, we adopt a user centered perspective: our problem space is comprised of 3 sets of dimensions described in section 2.

- The two first set of dimensions are used to characterize the user: her/his information needs and her/his knowledge about the system.
- The third set of dimensions deals with the notion of input expression: how the user specifies her/his information need.

After a presentation of our design space, we then go one step forward in the design by describing the users tasks. To achieve this goal, we apply the seven step model of Norman's Theory of Action [8]. Section 3 presents our interpretation of the Theory of Action applied to the information retrieval task coupled with the dimensions of our design space. Finally, we illustrate the discussion with IRIS-M, an Interactive Information Retrieval System that we have developed.

2. Characteristics of user's situation: a design space
As pointed out by N. Belkin, information retrieval is a tool that supports human problem management. Consequently it does not attempt to answer questions or solve problems but is intended to help users find information that might be useful for those purposes [3]. The goal of an IRS is therefore that the user obtains a response appropriate to her/his situation. We characterize a user's situation by her/his information needs, her/his knowledge about the domain and the system: information needs and knowledge are two sets of dimensions of our design space. The third set of dimensions or space deals with the notion of input expression: it describes the expression specified by the user in a presentation independent way. This last space is closely related to the two first ones as shown in Figure 1.

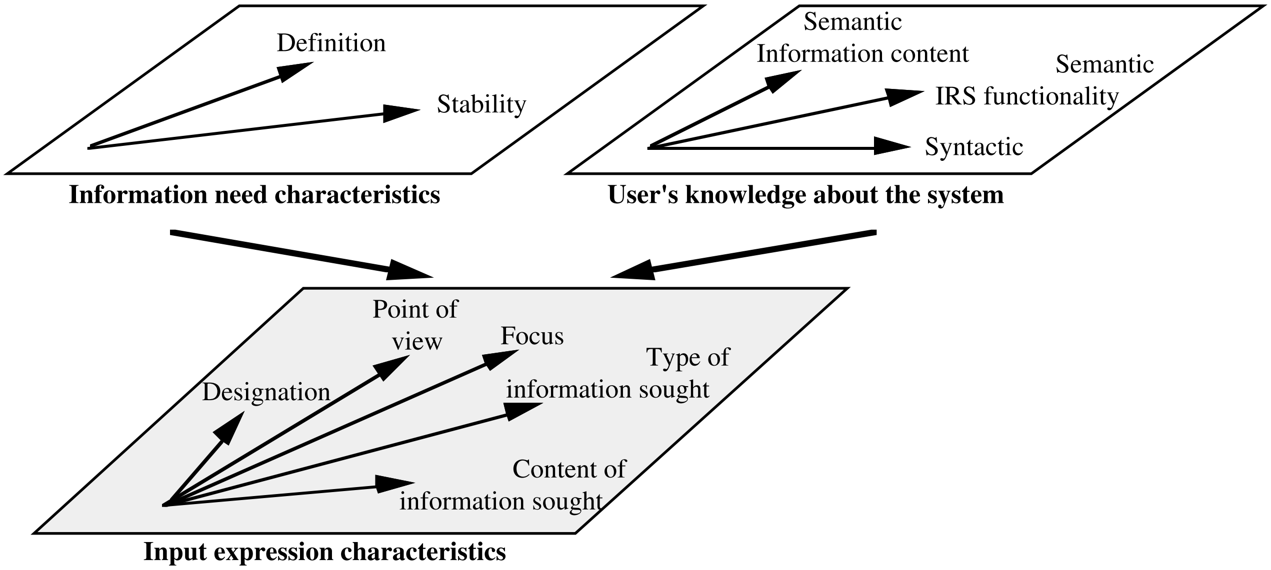


Figure 1: Our design space, information need, knowledge and input expression.

2.1. Information need characteristics

This first set of dimensions expresses the intrinsic properties of an information need. Such properties are closely related to the user's cognitive state. The user's cognitive space includes various cognitive elements such as the work-task, the interest domain, the cognitive state and the problem space that are identified in the Global Model of Polyrepresentation in IR [5]. The "information need characteristics" dimensions include two basic properties, namely definition and stability.

2.1.1. Definition of information need

This axis primarily covers the precision or vagueness of information needs. According to her/his knowledge and cognitive space, her/his information need can be well or ill defined. We adopt here a user's point of view. Nevertheless it is important to notice that a well-defined information need from the user's point of view can correspond to a very vague query from a system point of view and vice-versa.

2.1.2. Stability of information need

This axis covers the stability or variability of information need. P. Ingwersen [5] shows that information need varies because it depends on dynamic cognitive structures. We identify two cases of variability: topic and definition variability. The user can decide to change the topic of research while interacting with the IR system. Moreover the definition (precision) of information need can vary. An ostensive approach for IR navigation is a promising way for dealing with instability [1].

2.2. User's knowledge of the system

We here classically distinguish two types of knowledge: the syntactic and semantic ones. On the one hand, syntactic knowledge refers to the knowledge about how to use the system. Such knowledge has a direct impact on the user's actions while accomplishing a task. On the other hand, semantic knowledge covers the knowledge that the user has about the domain of the research ("*Information content*") and about how the system is internally functioning ("*IRS functionality*"). For example the user may identify documents that are not properly indexed from her/his point of view: Such knowledge will influence the user to choose special query words that may not be natural. The user adapts to the system: such a situation especially arises when too many results are provided to the user or when no satisfying results are retrieved.
Syntactic and semantic knowledge develop while interacting with the IR system. The user learns how to use the system while interacting with it. In addition she/he develops a mental representation about how the system works.

2.3. Input expression characteristics

This set of dimensions are used to characterize the input expressions specified by the user. Nevertheless such characteristics do not describe the form of the input expression such as the interaction modality used to specify the expression. The axes define more general properties that can be matched to several interaction techniques while designing the user interface.

2.3.1. Designation (connote/denote)
The "Designation" axis deals with the way information can be designated. We can compare intrinsic properties of an entity so that we connote this entity. For example, when we are looking for a book on computer science, we specify a property of the object sought: the topic of the book. It corresponds to a connotation of the document. On the other hand we can also denote a book by giving its exact title or its unique ISBN identification number. A denotation should always designate only one element (an article, an image, etc.). The distinction between connotation and designation can be difficult to establish when the connotation property is discriminatory. Moreover a unique identification (denotation) is also a property of an object (connotation).

A designation may be either a connotation or a denotation, as well as a combination of the two latter. If the system is able to distinguish these modes of designation, it can accordingly adapt its matching process in a more efficient way. Indeed the denotative part of a query can be processed by a simple non fuzzy matching algorithm. This matching process can be done using a classical database retrieval engine. The connotative part of a query is more complex to process: connotation can refers to various properties of the object sought. For example the following query "I am looking for a book on computer science that is often used by students" is related to the utilization of the book. This way of characterizing a book (connotation) is rarely embedded in an IR system. In addition there is very often a direct relationship between connotation and user's point of view: For example if the user specifies a query like "a picture with flashy colors", the connotation property refers to the user's feeling or preferences. The following axis "Point of view" captures those notions.

2.3.2. Point of view (subjective/objective)
This axis expresses the way information can be described. Two values are distinguished along the "Point of view" axis: subjective and objective.

Subjective characteristics depend on the user's experiences and more precisely the user's feeling like happiness or fear. A query like "movies that frighten me" is typically subjective. Subjective aspects of a query depend on a unique user and are time-dependent: it is thus not a reliable information.

We can contrast subjective points of view with objective ones. For example a query about "a movie with blood" is based on an observable characteristic of images. Nevertheless there is still subjectivity in such a query: "which kind of blood! Everyone knows that in horror movies the red blood is fake!

To sum up, the axis cannot be valued: a continuum from subjective characteristics to objective ones must be considered. Objective characteristics can be easily processed by the system since these just involve measurement. Unfortunately there is often a subjective part in queries. An IRS has to tackle this fact mainly through interaction. Interaction can reduce the subjectivity of a query in favor of more objective aspects. To achieve this goal, behavior databases can be available that allow the system to refine the query. For example the system can incorporate all the words related to "fear" (in a thesaurus). Another way to solve this problem is to let the user navigate and see the document. Thus, by reducing the key role of a query in the global process of retrieval, we can also reduce the subjectivity.

2.3.3. Focus (generic/specific):
This axis characterizes the precision of the expressed need. The need can be fuzzy like "I am looking for pop music" as opposed to "I am looking for Pink Floyd audio compact discs" which is more precise. The expertise of the user does not define the "focus" of query. Indeed an expert may be generic on purpose and a novice may only know few specific words. Again minimizing the key role of the query and increasing interaction and navigation in the information space may eliminate the problem of "focus".

The precision of the need can concern the information content itself as well as the information type or media. The query "a record of Pink Floyd" is more generic than "Pink Floyd audio compact discs" because the record media include magnetic tape, audio CD. Information type and content are described in the two following sections.

2.3.4. Type of information sought
Shannon has defined information as the correlation measure between two random objects [9]. A communication channel can be viewed as the temporal, virtual, or physical link that makes the exchange of information possible between communicating entities. Instead of considering the “linkage dimension” of communication, we stress the importance of the sources and recipients involved in a communication act. Thus, a communication channel covers a set of sensory (or effector) means through which particular types of information can be received (or transmitted) and processed. Several types can be identified along this axis: CD, picture, movie, text etc. Complementing this type, we consider the content of information in the following section.

2.3.5. Content of information sought
We identify two notions in order to describe the content of information sought: the meaning and the code of information. In IR the meaning of a document is represented by its index. But information

meaning depends on the receiver's knowledge and the way information is coded ("html page", "secam video") and presented. Information depth can vary according to the expertise of the user: for example on the one hand a digital sound may be understood by everybody if it is played. On the other hand a digital sound may be understood by experts if its frequency spectrum is displayed.

3. User tasks

After a presentation of our design space, we then go one step forward in the design by describing the users tasks. To achieve this goal, we apply the seven step model of Norman's Theory of Action [8]. Figure 2 presents our interpretation of the Theory of Action applied to the information retrieval task coupled with the dimensions of our design space.

3.1. Theory of Action

Starting from an information need, the different steps underline the mental and physical actions of a user interacting with the representations and using the I/O devices of the IRS. To emphasize the discrepancy between the user and the system, Norman denotes the semantic and articulatory distances as a gulf that must be bridged by the user and the system designer. Two gulfs, the execution and evaluation gulfs, are identified each of them being unidirectionnal:

• The execution gulf is comprised of the input semantic and articulatory distances (Figure 2). It proceeds from goals to physical actions. The starting point is the user's goal "get information". We call this step "user's motivation". Based on the goal, the first step is to build an information need. The latter can be very different according to the user and her/his knowledge ("User's knowledge" dimensions) and can evolve over time. The "Information need" dimensions identify characteristics for describing this step. The second step consists of organizing and transforming the query meaning resulting from the information need into a mental representation of a query: a query shape. Such activities are characterized by the "Input expressions" dimensions. Based on the query shape the user can then specify the actions to be performed requiring translation of the psychological variables in terms of the physical variables, defined and maintained by the system. After defining an appropriate action sequence, the user executes the actions using physical input devices.

• Symmetrically the evaluation gulf includes the output semantic and articulatory distances (Figure 2). Evaluation requires comparison of the current interpreted state of the system with the initial goal and intention. To make the evaluation possible, the current system state must first be perceived and then interpreted. The results to be perceived are very often presented as a simple ranked list of documents. Nevertheless more information can be presented and perceived. For example in TIAPRI [6], users can see a star field on the screen where each star represents a document. In a glance, she/he can perceive and interpret the density of retrieved documents and their distribution according to main keywords. The following section focuses on the evaluation gulf.

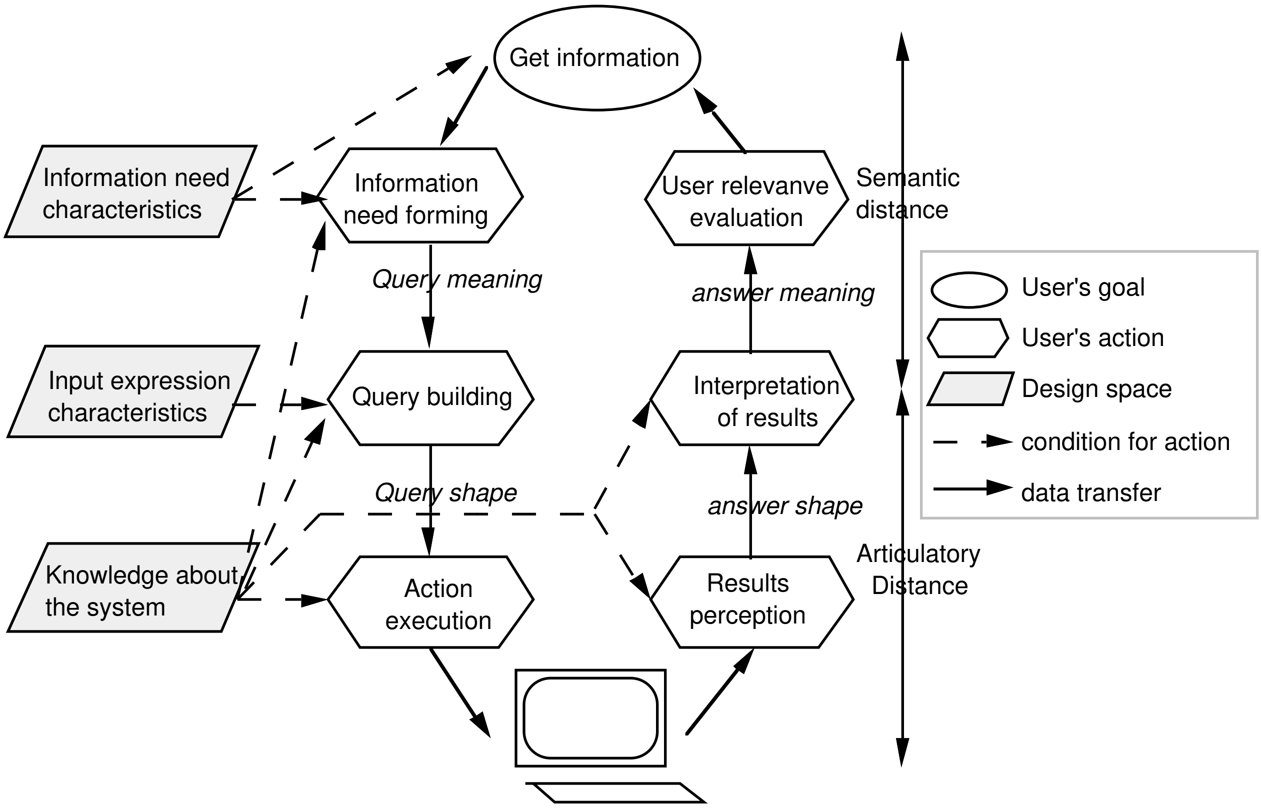


Figure 2: User's activities while accomplishing a retrieval task.

3.2. Perceiving the corpus

While experimenting with our TIAPRI interface, we have pinpointed the importance of perception and interpretation. As shown in Figure 2, the user's knowledge about the system influence the perception and the interpretation of the system output. This interpretation in turn has a direct impact on the evaluation of the system state. The evaluation step is in fact the most important one because it defines the user's satisfaction. That is why we concentrate our effort in improving the user's perception so that she/he can establish a correct mental representation of the system functioning. We underline here the key role of the user's knowledge about the system while interacting with the IR system.

The system can present the documents, its choice among them and its knowledge. The knowledge includes document indexes, thesauri, document links and matching functions. Several user interfaces [2, 4, 10] are efforts to present such knowledge. Our work focuses on the presentation of the system concepts that help the user to improve the definition of his need and the perception of the results. The results belong to the collection of documents (corpus) managed by the system. On the one hand the system has access to the whole set of documents. So the system selects the relevant documents based on the examination of the entire set. On the other hand the user's evaluation is only based on a sub-set of documents that are selected by the system. The relevance evaluation step is then severely distorted by this fact. In order to establish a more accurate relevance evaluation, we think that the user has to be able to perceive the whole corpus from which the selected documents are extracted. The remaining problem is to design an output interaction technique that enables the user to perceive the selected documents as well as the whole corpus. Our IRIS-M system presented in the following section is one such attempt.

4. The IRIS-M system

As shown in the appendix (Screen dumps 1 and 2) IRIS-M is an IR system that displays a representation of the corpus by clustering documents and presenting each document cluster as a 3D cone. Each cone belongs to a plane. The stack of planes represent levels of abstraction:

- The lowest plane is the lowest level of abstraction where documents are clustered.
- The clusters in plane n are clusters of the ones in plane (n-1).

The whole hierarchical structure presents the corpus. This structure is computed using a simple cluster algorithm. Using a zoom the user can focus on a cluster: the documents and the number of documents that are related to a given topic. When the user selects a cluster, all the embedded clusters in the bottom planes are lightened. Using a toolglass (magic lens), the user can see documents through the cones. Only few elements of documents are visible (title, main keywords, abstract). The navigation between planes and cones enables the user to perceive the corpus without initial queries. It is possible to keep track of the interaction because the path is lightened. The user can therefore always perceive the context of the search: the neighborhood and the path. Moreover the user can designate a document using a connotative value. Denotation corresponds to a query identifying one document. The different planes embody a hierarchical structure: the user can focus on a large spectrum of documents ranging from the generic to the specific. Finally toolglasses enable the user to perceive the documents according to their information types and contents.

The software design of IRIS-M is based on the PAC-Amodeus software architectural model [7] and IRIS-M is running on NextStep environment using RenderMan for 3D visualization.

5. Conclusion

In this article we have proposed dimension spaces that should help in the design of current and future interactive information retrieval systems. They draw upon psychology, HCI and IRS with their associated notions of information need, user's knowledge and input expression. By doing so, they identify salient characteristics for design studies of interactive IR. Based on our design space and Norman's Theory of Actions we then model users' activities while accomplishing a retrieval task and stress the key role of the perception and evaluation steps. Finally we have shown the relevance of perception of the whole corpus in a retrieval session.
In future work we plan to refine our design space and then establish a taxonomy of UI properties based on the identified characteristics. Such a taxonomy may be directly useful for the design as well as the evaluation.

Acknowledgements

We wish to thank Jianyun Nie and George Serghiou for reviewing the paper.

6. References

[1] Campbell, I., Supporting information need by ostensive definition in an adaptative information space, FERMI Technical Report, may 1995.

[2] Chalmers, M. and Chitson, P., Bead: Exploration in Information Visualisation, Proceedings of SIGIR'92, 1992, pp. 330-337.

[3] Belkin, N., Interfaces for Information Retrieval Systems, User Modelling in Information Retrieval Systems, Lecture Material, First European Summer School in IR, 1990, pp. 274-334.

[4] Hemmje, M., Klundel, A. and Willet, A., LyberWorld: A Visualisation User Interface Supporting Fulltext Retrieval, Proceedings of SIGIR'94, 1994, pp. 249-259.

[5] Ingwersen, P., Cognitive perspectives of information retrieval interaction: elements of a cognitive IR theory, Journal of Documentation, (52), 1, march 1996, pp. 3-50.

[6] Mulhem, P. and Nigay, l., Interactive Information Retrieval Systems: From User Centred Interface Design to Software Design, Proceedings of SIGIR 96, 1996, pp. 326-334.

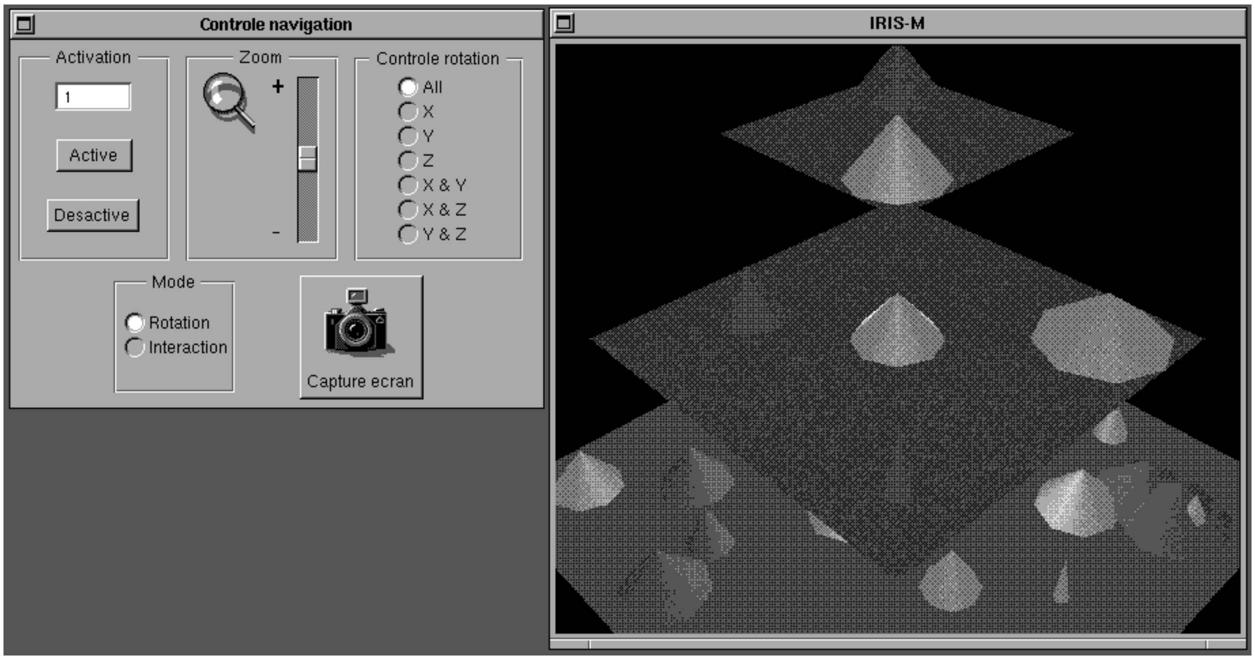
[7] Nigay, L. and Coutaz, J., A Generic Platform for Addressing the Multimodal Challenge, Proceedings of CHI'95, 1995, pp. 98-105.

[8] Norman, D., Cognitive Engineering. Chapter 3, User Centered System Design, New Perspectives on Human-Computer Interaction, Hillsdale, New Jersey: Lawrence Erlbaum Associates, pp. 31-61.

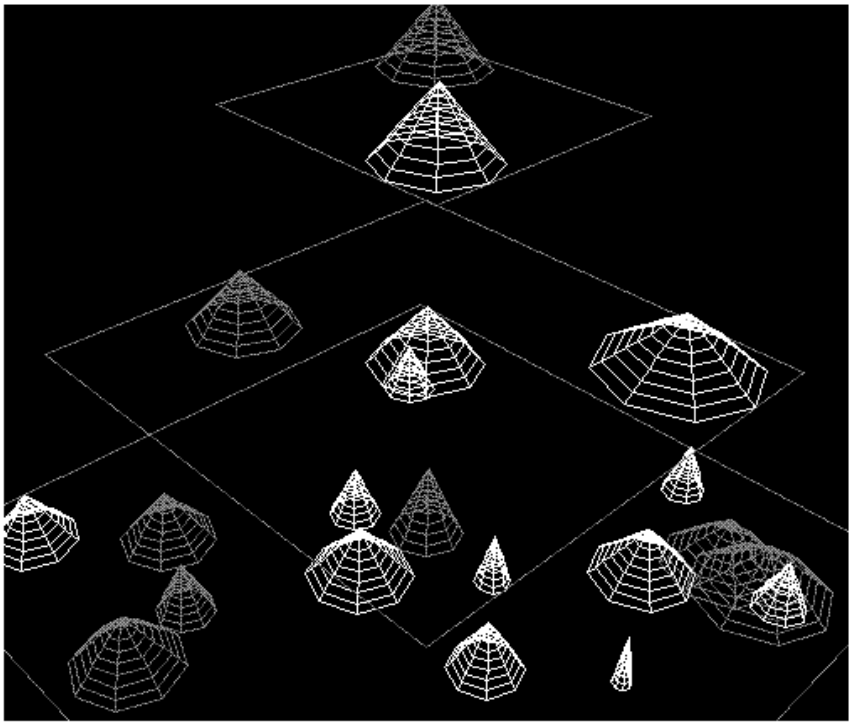
[9] Shannon, C., A mathematical theory of communication, Bell System Technical Journal (27), 1948.

[10] Zizi, M., Cartes Dynamiques Interactives : une métaphore spaciale pour l'exploration des espaces informationnels, Thèse de l'Université Paris-Sud, Orsay, France, 1995.

Appendix: IRIS-M



Screen dump 1



Screen dump 2