Supporting the active learning

of collaborative database browsing techniques

M. B. Twidale, D. M. Nichols, J. A. Mariani, T. Rodden and P. Sawyer
Computing Department, Lancaster University

We describe the implications of a study of database browsing behaviour for the development of a system to support more effective browsing. In particular we consider the importance of collaborative working, both in learning browsing skills and in co-operating on a shared information-retrieval task. From our study, we believe that an interface to support collaboration should promote the awareness of the activities of others, better visualization of the information data structures being browsed, and effective communication of the browsing process.

Introduction

The proliferation of databases has led to a growing demand for skills in navigating and manipulating them. These skills are important not just in order to undertake study and research in Higher Education, but as an important transferable skill likely to be expected by industry and commerce of all graduates. We are developing a system that can more effectively support the browsing of library databases. In particular we are interested in developing an interface to support the process of collaborative browsing.

We have made a study (Twidale *et al.*, 1994) of the requirements of such an interface, chiefly to inform our design intuitions. The study has involved a substantial literature survey, interviews with subject librarians, and small-scale in-depth studies of authentic browsing activities undertaken by library users. The latter involved observing remote collaboration using existing simple tools including Unix *Talk* in order to reveal the problems that a more sophisticated system should address.

The use of library resources is stereotyped as a solitary activity, with hardly any mention in the library science and information retrieval literature on the social aspects of information systems. However, our study indicates notable collaboration, with users consulting both library staff and each other. Informal computer-based collaboration already exists through sharing or leaning over terminals and pointing at screens.

Traditionally, as computer scientists, we have designed databases to appear as single-user systems. We aim to build a system to support more active collaboration.

Browsing

Whereas some information scientists use the term *browsing* in a very restricted and precise manner, we use it to describe all search activities that involve a level of imprecision about the desired outcome of the search (Bates, 1989). A critical feature is that users refine their requirements over time as more information becomes available and set subgoals to satisfy them. The literature on browsing backs up our observations of its importance to library users as well as the problems that novices have. Researchers have stressed the importance of serendipity and its effective exploitation in the browsing process (Rice, 1988). But it also seems that expertise makes you luckier! This may be because an expert is better able to recognize and exploit the accidental find. There are two types of relevant expertise: generic knowledge of browsing techniques, and local knowledge of the subject area being explored. In addition, browsing involves two different kinds of skill: tactics, which are actions relating to the composition of a single query, and strategy, which involves actions in managing a sequence of queries. Examples of tactics are the use of Boolean operators, keywords, wildcards and indexes. Strategic skills include choosing the right database, restricting or widening a search, and managing your evolving set of goals and subgoals.

Browsing tactics involve knowledge that can be unique to a single database system, and their transfer between the many and constantly evolving systems that learners use is difficult. This is a frequent complaint both of users and of information professionals (Yee, 1991). By contrast, the strategies, being more abstract, transfer better and are the kinds of generic transferable skills that students should acquire. Unfortunately, this very abstraction can also make their acquisition all the more complex and their consideration may be masked by the more visible issues of tactics.

Learning about the system

Novices may have misconceptions about the browsing process. These can include an inappropriate overview or metaphor of the search process. An example we observed was considering an Online Public Access Catalogue (OPAC) system as analogous to an intelligent colleague. Associated with such a misconception will be a set of beliefs, such that results are presented in some order of importance, or that a 'common sense' filter will have been applied. For example, a novice with the 'intelligent colleague' misconception would find it bizarre that a search on the word *browse* will return references to the behaviour of reindeer, and consequently distrust any use of the search system as being so stupid as to be utterly flawed.

Information systems are complex, and it is entirely appropriate that novices learn a minimal subset of the available facilities to enable them to do some limited searching. Unfortunately, this can lead to users who are confident in their use of the system but who have settled into sub-optimal work patterns. Examples of these are: getting several

hundred hits and reading them all, failing to look at all the results when there are a reasonable number, and failing to pursue a lead or revise a strategy. Users may settle into these patterns either because they are unaware of the more sophisticated methods or because they choose not to use or practise them, due for example to cognitive overload (Rudd & Rudd, 1986). Therefore, in addition to providing them with an introduction to that subset, we must consider issues of how they can continue their learning of more advanced features. The process of continuing learning involves experimentation, blind trial and error, asking experts for help, informal learning from colleagues, and advice being volunteered by experts as a part of a consultation about a domain issue.

Potential losses from computerization

Computerized library databases offer features not available in traditional library information structures (such as keyword searching and the ability to work remotely), but they may also lose some useful features. These include the spatial nature of the way information is presented in a library: the ability to walk round the bookstacks and to exploit our powerful spatial memory abilities (Beheshti, 1992).

The organization of bookstacks can also facilitate browsing: a searched-for book on a shelf will have related books near to it that can be equally or even more relevant to the user's vague and constantly evolving information needs. An electronic view of information is remarkably impoverished compared with the vast amount of peripheral information of real books. These include the ease of seeing whether a book is brand new, well thumbed, borrowed a lot, or ancient but never consulted (Mitev, 1989). People also have well-developed spatial awareness and memory skills that can be exploited in navigating a physical library and in supporting the retrieval of information by its position (Chang & Rice, 1993). However, not everything is easily accessible in a physical library. For example, at Lancaster the Librarianship section is not on open access, and one cannot physically browse the books on browsing! Furthermore, one naturally cannot browse the shelves for books that others have borrowed. Even the old-fashioned card-index file can convey peripheral information from nearby cards, the colour and degree of ageing of the card and even how dog-eared it is. Such physical information structures enable the browser to get a sense of the size of the related information.

Collaboration issues

It is not just the inanimate contents of a traditional library that convey useful information. People can also be useful to a browser. One can observe and learn from the browsing techniques of others, discuss issues with co-learners or with subject experts, and also be aware of the activities of others that may be of interest and relevance to one's own work. For example, upon seeing a colleague in an unexpected part of the library, one might choose to ask what she has found there. Similarly, upon seeing someone in 'your' area, you may decide to introduce yourself as someone also interested in that field. A computerized library that is accessed remotely will lack these advantages unless we take steps to re-introduce them into the system.

Research in Computer Supported Co-operative Work employs a useful classification of collaboration (Rodden, 1991). Collaboration may be remote or co-located, as well as being synchronous or asynchronous. In conventional libraries, we can consider most co-operation to be co-located and synchronous, but the computerization process makes the other permutations possible, while offering new opportunities for the first.

We envisage a number of scenarios in which the system being developed might be used in an educational context. In all cases, the interaction may be synchronous (participants working at the same time) or asynchronous (participants leaving messages for each other):

- *Expert consultancy.* A learner is browsing the database and decides that she is not making the progress she would like. In a physical library, she might go and talk to the subject librarian. For the database use, she may wish to communicate by telephone or email (or even ultimately by video link). The expert calls up a representation of the student's browsing history and composes suitable advice. This advice can be specific, general and remedial. It can help to solve the current task, explain a generic browsing technique and also correct any apparent misconceptions. As part of the explanation, the expert passes on an annotated browsing procedure which the student can view and even use on her terminal.
- *A 'language laboratory' for database skills.* A number of individual learners are browsing a database as part of a practical class on database-browsing skills. They may all be in the same teaching room, or working remotely on their own terminals elsewhere. The expert can observe the browsing activities of the students in turn, and offer advice as necessary.
- *Collaborative browsing for learners.* Small groups of learners who have a similar browsing task make use of the awareness mechanisms to monitor and discuss the progress and activity of each other. This is known to be a useful learning activity both for the questioner (who asks questions such as 'Why are you doing that?'), and for the respondent who has to reflect on her action in order to generate a suitable explanation.
- *Serendipitous meetings.* If the groups in the case immediately above are made larger, database awareness can be a useful tool for supporting serendipitous meetings of the kind that naturally occur in physical libraries where strangers meet by browsing the same bookstack and finding common interests for collaboration, or when finding that a book they want is already being used by someone else, who therefore might be worth talking to. Now, one can become aware when others are, or have been, browsing the same parts of the database.

We are working on an interface to support collaborative browsing. Our study will inform the initial version of our interface, and one of our findings we believe to be particularly important. Collaborative working implies a need to share information: both the end product (the hits) and the process (the search strategy/tactics). Similarly, there is a need to share this information with the librarians, for whom inspection of the search process can

reveal not only gaps in the users' browsing techniques but also an indication of their degree of searching sophistication. In addition, an externalized representation of the search process reduces cognitive load and facilitates reflection, a vital component of learning. An initial analogy would be the Unix history list, but the actual representation will need to be considerably more sophisticated and flexible.

Conclusion

It seems that collaboration issues have not been addressed in OPAC design. We believe that an interface that takes them into account in providing visualization features will greatly facilitate both the learning and exercise of browsing skills.

Acknowledgements
Financial support for the research reported in this paper was provided by a grant from the JISC New Technologies Initiative.

References

Bates, M.J. (1989), 'The design of browsing and berrypicking techniques for the online search interface', *Online Review*, **13**, 5, 407–24.

Beheshti, J. (1992), 'Browsing through public access catalogs', *Information Technology and Libraries*, **11**, 3, 220–8.

Chang, S.J. & Rice, R.E. (1993), 'Browsing a multidimensional framework', *Annual Review of Information Science and Technology*, **28**, 231–76.

Mitev, N.N. (1989), 'Ease of interaction and retrieval in online catalogues: contributions of human-computer interaction research' in Hildreth, C.R. (ed), *The Online Catalogue: Developments and Directions*, London, The Library Association.

Rice, J. (1988), 'Serendipity and holism: the beauty of opacs', *Library Journal*, **113**, 3, 138–41.

Rodden, T. (1991), 'A survey of CSCW systems', *Interacting with Computers*, **3**, 3, 319–54.

Rudd, J. & Rudd, M.J. (1986), 'Coping with information load: strategies and implications for librarians', *College & Research Libraries*, **47**, 4, 315–22.

Twidale, M.B., Nichols, D.M., Mariani, J.A., Rodden, T. & Sawyer, P. (1994), Issues in Collaborative Database Browsing, Internal Departmental Report, Computing Department, Lancaster University.

Yee, M.M. (1991), 'System design and cataloguing meet the user: user interfaces to online public access catalogs', *Journal of the American Society for Information Science*, **42**, 2, 78–98.