

library management. The emphasis on the systematic and effective deployment of resources to the benefit of users and library staff offers a valid alternative to traditional management approaches. This is particularly true in a time of rapid and fundamental change in the environment in which university libraries are operating.

However, this study also found that although marketing strategies were widely used in both countries, the concept of marketing is still frequently misunderstood. Consequently, marketing strategies are often used neither systematically nor coherently. Constraining factors in Germany seem to be rigid budgetary structures, and a system of staff representation with extensive veto rights. In contrast, the most important restricting factors in Britain are a lack of resources and rapidly changing higher education policies. The latter has made forward planning and even reacting to these changes difficult. In both countries, marketing management in university libraries is limited by a lack of specifically trained staff, and this lack of training poses particular difficulties in the area of financial management and of performance measurement.

This research also indicated that the awareness and knowledge of marketing has increased in both countries in recent years. However, some misconceptions about marketing prevail, and prevent some library managers from reaping the benefits which their services could gain from a fully implemented marketing approach. In a time of decreasing or stagnating funds, a marketing approach to management can be a tool for using available resources effectively in the best interest of users and library staff.

Intelligent information retrieval from the World Wide Web using fuzzy user modelling

GARY MOONEY

Division of Learning Development/School of Computing Sciences
De Montfort University, Leicester.

Introduction

The ever-changing nature of information sources, coupled with the increased demand on dwindling academic resources, led librarians and other information professionals to recognise the need for information retrieval (IR) systems that can incorporate the expertise of the information professional and gather knowledge about the user's experiences and preferences (Werckert & Cooper, 1989). The explosion of available information resources brought about by the development of the Internet and the World-Wide Web (WWW) has strengthened this need. The goal is to produce an 'intelligent' IR system which would work with the user to satisfy their information needs, so the application of Artificial Intelligence (AI) techniques seems a likely approach to the problem (Morris, 1990).

Attempts to incorporate intelligence into IR are numerous, have met with limited success, and involve many facets of AI (Mooney, 1996a). Many of these attempts rely on knowledge of the structure and content of the database (information store or knowledge 'space'). They are systems for performing dynamic, ie changing, queries on static databases and require extensive knowledge about the databases upon which they are to search. Even if developed only to aid IR from fixed databases, their success has been limited (Shoval *et al.*, 1990; Zegher *et al.*, 1993).

However, the Internet is not static. It is a dynamic, distributed information store, meaning that it is *constantly* changing, and its contents reside on *innumerable* sites around the globe. This last point is as important as the first, for it ensures that the true content and extent of the Internet is effectively unknowable (Mauldin, 1996). Various tools known as search engines have been developed to combat the phenomenal growth of the Internet and allow the user to search and retrieve information effectively (Mooney, 1995). They possess no 'intelligent' capabilities.

This article investigates the effects of applying fuzzy logic and user modelling techniques to the process of information retrieval from the WWW, a major part of the Internet. This is a novel AI approach to the process of IR. To perform the investigation, a prototype system, the Fuzzy Query Modelling Assistant (FMQA), has been developed. The focus of the investigation was whether the results achieved by using the FMQA would improve upon those returned by using an existing search tool, specifically LycosTM (Mauldin, 1996), alone. To answer this question a user study of the FMQA is being performed and its early results are reported.

Fuzzy User Modelling and Intelligent IR

A major problem with IR lies in the vagueness and lack of precision of the prospective searcher's information need. This vagueness and lack of precision leads to the aforementioned problems and these are exacerbated by the nature of the WWW. The problem of information overload is one example. A search with the tool LycosTM using the search string 'information retrieval' produced 61,000+ hits (Mauldin, 1996). However, information about the user's experiences and knowledge of the search subject and of the WWW in general can be used to modify the query intelligently and produce better IR results.

User modelling research has shown that adaptive user stereotypes are often used to represent different sorts of user and their characteristics (Rich, 1979). Fuzzy logic, with its inherent ability to capture and represent partial knowledge, is a valid AI technique to use in IR - a process involving the representation of information needs as queries, with all the attendant vagueness and semantic ambiguities (Zadeh, 1993). Here, the user stereotypes are represented as fuzzy sets to ensure flexibility and adaptivity. This concept is at the heart of the prototype system, the FMQA.

Development of the FMQA

The FMQA is not designed to act as a *new* 'intelligent' search engine. Within the field of distributed AI and computing in general there has been much research and development into the notion of searching and intelligence through the development of agents. (Wooldridge & Jennings, 1995). The FMQA seeks to alleviate IR problems through 'intelligently' assisting the user in a search. In this sense it is similar to the concept of *interface agents* defined in Maes (1994) but the FMQA applies fuzzy logic and user modelling to the query formulation of searches. The aim is to refine a query before it is submitted to an existing search engine. This refinement is based on knowledge about the user's beliefs and experiences (in the Internet and the subject domain of AI) captured through an on-line interactive session.

The user interface is written in clear and simple HTML to make it compatible with WWW browsers and existing search tools. The two fuzzy knowledge bases (FKB) which contain the sets and rules were based on software developed by Cox (1994).

The captured user knowledge is used to adapt default user models in order to represent an individual user. This representation is then combined with the user's query to produce the refined query. The

knowledge is captured from two on-line interactive questionnaires. Each question is represented by a fuzzy set. The FKB combines the answers to these questions to produce two sets which represent the individual user model. The defuzzified values from these sets are used to refine the query.

The user is then given the option to submit either the original or the refined query to Lycos™. Presently, the user must choose the original query from a list of topics representing different areas of AI but this is just a facet of the prototype. The system could easily be expanded to include other fields of interest and eventually to allow the user to enter the query words themselves. Additionally, the final fuzzy sets are lost when a user finishes accessing the FMQA and new sets are produced with every new session. However, the system could easily retain these sets and use the information they contain in future sessions.

User study

A user study has been performed in which a number of DMU undergraduate students from the first year intake of two courses accessed the FMQA and provided feedback on its use and performance. The aim of the study was to answer the following question:

Would the use of the FMQA improve upon the IR results obtained from the WWW by using the search tool alone in a certain branch of the subject domain AI, when used by novice and intermediate users?

The students were chosen after a questionnaire (Mooney, 1996b) showed that these courses would provide sufficient subjects who were at a novice or intermediate level with regard to WWW and subject domain (a branch of AI) experience. The system itself decides whether a user is of novice or intermediate type from their answers to the questionnaires.

The study took place over a number of weeks with each user accessing the FMQA at a similar time each day. This ensured that each user's reaction to the system was not unduly influenced by differences in the Internet network traffic speeds. Each user was asked to submit the original topic they chose as a query and the refined query produced by the system. This is equivalent to using Lycos alone to search for the AI topic and then using the same FMQA-modified AI topic and allows the results to be used to answer the central question of the study. As part of each session, the user was presented with an e-mail form and asked to list the best and worst results for each query, and to rank these as well by giving them a score between 1 and 10. They were also asked to comment upon the results, in terms of usefulness and relevance, and upon the system, in terms of ease of use and design.

Results

Early results from the study indicate that the FMQA does indeed improve upon the IR results achieved by using Lycos™ alone. The following are sample comments of users on the results:

'If I had the choice I would use the modified query. It seemed to produce results more appealing to me (subject matter).'

'The search result with the modified query depending on the response gives you a more specific response which is more relevant than the unmodified response.'

'Yes, the modified query gives you a more detailed search, giving you a better response on the topic.'

In terms of scoring the results, the modified queries have generally received higher scores (out of 10) than the unmodified queries. For example, in a query for information on 'Knowledge-based Systems', the following scores were given:

best two results - unmodified	5/10, 5/10
best two results - modified	9/10, 8/10
worst two results - unmodified	3/10, 1/10
worst two results - modified	6/10, 6/10

During a query looking for 'Fuzzy Logic', which afterwards the user remarked that the modified results were more relevant, the best result for the modified query was a website dedicated to fuzzy sets and systems (Brown, 1996). The dedicated website contains many WWW links to general sites of interest to fuzzy logic reserachers and would be a good starting point for a novice to the area, which was the category this user was placed in by the FMQA.

Conclusions

This article has examined the application of fuzzy logic and user modelling to the process of IR from the WWW, the concept being to assist intelligently the user in searching for information and reduce the problems commonly associated with IR in general, eg irrelevance and redundancy. A prototype system, the FMQA, was developed, which realises the concept by employing knowledge about the user to modify queries before they are submitted to an existing WWW search tool. This knowledge is represented in fuzzy sets which act as adaptive user stereotypes. Early results from a live user study of the FMQA show that, in the opinion of the users, the results achieved from using the system do improve upon those obtained from using the search tool alone.

References

- Brown, S. 'Fuzzy Sets and Systems' at <http://www.abo.fi/~rfuller/fuzs.html> , 1996
- Cox, E. The fuzzy systems Handbook - A practitioner's guide to building, using and maintaining fuzzy systems New York: Academic Press, 1994
- Maes, P. Agents that reduce work and information overload *Communications of the ACM* 37 (7), 1994, pp. 31-40.
- Mauldin, M. 'Lycos: The Catalog of the Internet' at <http://www.lycos.com/> , 1996.
- Mooney, G J. Information retrieval from electronic sources CMS Working Paper No. 62, School of Computing and Mathematical Sciences, DMU, May, 1995.
- Mooney, G J. Intelligent information retrieval CMS Working Paper, School of Computing and Mathematical Sciences, DMU. (in progress), 1996a.
- Mooney, G J. 'E-mail questionnaire to DMU Undergraduate Students, 1995/96'; contact the author, gjmooney@dmu.ac.uk, for details, 1996b.

Morris, A. 'The Use of Expert Systems in the Library and Information Service Environment' in Computers in Libraries International 90 J Eyre (ed.), London: Meckler, 1990, pp 104-107.

Rich, E. User modelling via stereotypes *Cognitive Science* 3, pp. 329-354.

Shoval, P. *et al.* ERSE: An expert retrieval system for electronic databases *Expert Systems for Information Management* 3 (2), 1990, pp. 85-114.

Weckert, J., Cooper, C. Libraries and expert systems *Riverina Library Review* 6 (4), 1989, pp. 323-328.

Wooldridge, M., Jennings, N R. Intelligent agents: theory and practice The Knowledge Engineering Review 10 (2), 1995, pp. 115-152.

Zadeh, L A. 'The role of fuzzy logic and soft computing in the conception and design of intelligent systems' in Proceedings of 8th Austrian AI Conference (FLAI '93) E P Klement, W Slany (eds) Lecture Notes in Artificial Intelligence 695, Springer, Berlin, 1993, pp.1-2.

Zegher, de I. *et al.* 'IRHIS: Intelligent, adaptive information Retrieval system as Hospital Information System front end' in Advances in medical informatics: results of the AIM exploration action J Noothoven van Goor, J P Christensen (eds), 1992, pp. 188-198.