「情報管理研究」, vol. 24, no. 4, 1993. 12., pp. 43~57

Expert Systems as a Search Intermediary

문 성 빈*
(Moon, Sung Been)

抄 錄

본 논문은 인공지능(artificial intelligence) 및 전문가 시스템(expert system)의 기본 개념과 이에 적용되고 있는 특정한 기술인 퍼지 이론(fuzzy logic)을 논하고 있으며, 지난 몇 년 동안 탐색 중개인으로서의 전문가 시스템을 조사해 보았다. 이러한 전문가 시스템은 1) 특정한 데이터베이스에 관련된 질문서 작성을 도와주며, 2) 탐색용어나 데이터베이스 선정에 관한 결정을 보조하고, 3) 탐색중에 있는 이용자에게 助言을 해주고 있다. 또한 전문가 시스템을 개발하는 데 있어 어려움 및 制限點을 논의하고 있다.

키 워 드

인공지능, 전문가 시스템, 탐색 중개인, 온라인 탐색, 퍼지 이론

ABSTRACT

This paper discusses the basic concept of artificial intelligence(AI) and expert system and a particular technique(fuzzy logic) applied to expert systems. It examines expert system as search intermediaries during the past few years, particularly in terms of the following functions: 1) handling certain classes of questions on a particular database, 2) assisting in decision making for selecting databases or search terms, and 3) offering advice while keeping the end-user in the control of the searching process. The limitations and difficulties involved in developing such expert systems are also presented.

KEYWORDS

Artificial intelligence, Expert system, Search intermediary, Online searching, Fuzzy logic

* 연세대학교 문헌정보학과 강사.
Dept. of Library & Information Science, Yonsei University.

1. Introduction

Information retrieval is central to the work of librarians and information specialists. Providing the most appropriate information to the user is a critical role performed by libraries and information centers. During the past few years, information scientists have been investigating new ways in which they can store and retrieve information and knowledge, and thus improve the effectiveness of information systems. Recognizing the significance of the new research efforts and encouraging further development, a number of innovative projects in library information science education and practice are stimulating and supporting new information systems, especially expert systems.

Expert systems studies grew out of Artificial Intelligence(AI) research. Expert systems are intelligently designed computer applications in which human expertise is incorporated into a knowledge database and control programs. The system can be used to aid or advise users on solving problems and making decisions. There has been much research which examines the extent to which expert systems can replace human experts in terms of their retrieval effectiveness. This field of study is still relatively new, and its importance is increasing. A variety of applications are now being developed(including medical diagnosis, engineering, finance, geological explorations, cataloging, and online database searching), and research is ongoing in many other areas of industry and the military. Brooks(1987) says "The realization of the need to incorporate and use knowledge within IR systems has led researchers to look at AI systems that also aim to incorporate and use knowledge, that is, intelligent knowledge—based systems." Expert systems, one of such systems, are the new frontier development in library and information science.

The last few years have seen information scientists such as Gerald Dejong (1985), Linda C. Smith(1980, 1987), and Harold Borko(1985) identify a number of areas of information handling which may be appropriate for the application of AI and expert systems. Particularly, the 1987 special issues No. 2 and No. 4 of Information Processing & Management deal solely with expert systems and AI, respectively. The American Society for Information Science showed a great concern with AI and expert systems, adopting a timely theme, "Artificial Intelligence:Expert Systems and Other Applications" for the 17th ASIS Mid-Year Meeting in 1988. Recognizing its tremendous promise and its potential applications in information environments, the ASIS conference examined much of the AI research and development underway to enhance information systems and service through a variety of presentations, demonstrations and panel discussions. However, the area with the longest history and the largest number of research activities is that of expert search intermediaries. These activities have been much more emphasized as the online environment becomes more complicated, with its various systems and databases to which users want to have access. The purpose of the expert search intermediaries is to make online systems directly available to end-users without the need to rely on human intermediaries(Smith, 1987).

This paper examines the concepts and techniques involved in expert systems and their prospects as search intermediaries for bibliographic information retrieval system, particularly in terms of the following functions:1) handling certain classes of questions on a particular database, 2) assisting in decision making for selecting databases or search terms, and 3) offering advice while keeping the end-user in the control of the searching process(Smith, 1987).

The application of knowledge-based systems' concepts and techniques to developing an IR system interface attracted many researchers. As online search systems become dependent on specialized access mechanisms such as commands, index terms, and query forms, it is natural for information scientists to seek ways of mapping effectively the user's request onto a search query. This is important not only because assistance by human intermediaries is expensive, but also because it would be nice to make a search system directly accessible to the end-user(Belkin et al., 1987).

2. Expert System

(1) Definitions

There are several definitions of what an expert system is, but a formal definition approved by the British Computer Society's Committee of the specialist group on expert systems is as follows:An expert system is regarded as the embodiment within a computer of a knowledge-based component, from an expert skill, in such a form that the system can offer intelligent advice or take an intelligent decision about a processing function. A desirable additional characteristic, which many would consider fundamental, is the capability of the system, on demand, to justify its own line of reasoning in a manner directly intelligible to the enquirer(Brook, 1987).

(2) Expert Systems Components

Computer programs written for an expert system are not the same as conventional ones, because they are differently organized. In other words, Expert Systems replace the software tradition of "Data+Algorithm=Program" with a new architecture, that is, "Knowledge+Inference=System"(Forsyth, 1984). These systems are designed to simulate expert human performance and to present a humanlike interface to the user. Today the new information science frontier designs, develops, and tests expert systems for use in libraries and other information centers(Borko, 1987). The four essential components of a full-fledged expert system are as follows(Brooks, 1987;Vickery, 1987).

i) Knowledge Base

A knowledge base, which embodies the human expert's knowledge about a specific problem domain, is the vital component of an expert system. This is the knowledge the system needs to accomplish its tasks during processing. Expert system may use various types of knowledge already incorporated into its system.

ii) Inference Engine

The inference engine is also an important system component. It examines the existing facts and rules and decides on the order in which inferences are made. There are two inference strategies:forward chaining in which the system reasons from a set of data to an outcome or goal-state, and backward chaining which involves working back from a goal or a conclusion to see if the conditions that would make it true are satisfied.

iii) Explanation Mechanism

An expert system includes the explanation mechanism that allows the system to justify its own conclusion. The system's capability of giving some kind of explanation seems an essential component which helps users take some action in response to the system's advice. Users would therefore understand why and how the system's decisions are made.

iv) User Interface

A user interface is the system component responsible for the communication between the user and the system. Most expert systems ask users to specify their problems to be solved and to provide the system with a set of initial conditions or observed data. A successful expert system relies heavily on this component.

3. Fuzzy-Set Theory and Expert Systems

(1) Fuzzy-Set Theory

The fuzzy-set theory, which involves a non-probabilistic method, is derived from the efforts aimed at developing effective information retrieval systems. The theory of fuzzy sets has already affected the research on information retrieval, and research groups around the world are involved in this alternative unconventional approach.

The basic concepts of this theory were introduced by Zadeh(1965). He extended the concept of classical Boolean logic, which uses "1" and "0" to represent truth and falsity, respectively. In addition, he adopted the following method to indicate partial truth. Thus

$$P(clever(X)) = 0.25$$

may indicate that the proposition that 'X is clever' is in some sense one quarter true, thus three quarters false. To combine pieces of evidence by using a fraction between zero and one, fuzzy logic defines the equivalent of the AND, OR and NOT operators as follows: (Forsyth, 1984)

P1 AND P2=MIN(P1, P2)(*smaller*)

P1 OR P2=MAX(P1, P2)(*greater*)
NOT P1=1−P1(*inverse*)

Bookstein(1985) outlined the following strengths and weaknesses of fuzzy−set model when applied to information retrieval. He includes as weaknesses:1) the ranking ability which is not sensitive to each term in the query, 2) no agreement on the way of making the best use of feedback information, and 3) uncertainty of the retrieval process which neither fuzzy−set nor traditional Boolean retrieval precisely recognized. And he considers its strengths to be:1) the flexibility of recognizing degrees of relevance, 2) ranking ability of documents according to the degree of estimated importance to the user, and 3) the capability of handling the relative importance of each term in document representation.

(2) Fuzzy−Set Theory Applied to Expert System

Much of the information stored in the knowledge base of an expert system is ambiguous, incomplete, and not totally reliable. This section reveals an urgent need for a better understanding of how to deal with uncertainty in expert systems.

As Negoita states, "Approximate reasoning is concerned with both the means for representing natural language expressions into a machine understandable form, and a methodology for making inference from these representations." Approximate reasoning involves techniques for drawing conclusions from an inference mechanism under uncertainty, in which the underlying logic is approximate or probabilistic rather than exact or deterministic. In its narrower sense approximate reasoning is an outgrowth of fuzzy logic. Fuzzy logic may be thought of as not only a unification, but also a generalization, of both predicate logic and probability theory. In this context, fuzzy logic offers a more powerful language for representing uncertain knowledge and a basis for developing rules to combine pieces of uncertainty evidence. Fuzzy logic is used successfully for example in the decision−support system REVEAL(Forsyth, 1984).

The first generation of expert systems such as MYCIN, PROSPECTORS, and CASNET(Gevarter, 1983) were designed using conventional probabilistic tools.

Expert systems, although of limited capabilities, then emerged as one of the most important applications of artificial intelligence. However, the introduction of the fifth generation computer(Brooking, 1984) and the new methodology based on approximate reasoning, has brought in new innovations in designing expert systems. This design methodology uses the expert's knowledge base extensively. The employment of approximate reasoning to characterize and manage uncertainty for the design of expert systems has played an critical role. In other words, the employment of fuzzy logic as a framework in the management of uncertainty in expert systems makes it possible to consider a number of issues which are not dealt with effectively or correctly by conventional techniques. As Zadeh(1983) points out:"Fuzzy logic provides a natural framework for the management of uncertainty in expert systems because—in contrast to traditional logical systems—its primary goal is to provide a systematic foundation for representing and inferring from imprecise rather than precise knowledge. In effect, in fuzzy logic everything is allowed to be—but need not be—a matter of degree. The greater expressive power of fuzzy logic derives from the fact that it contains as special cases the traditional two—valued as well as multi—valued logics."

Kiszka and Gupta(1985) suggested two fundamental stages for designing procedures for expert fuzzy control systems:a) knowledge acquisition and b) computer implementation. Although expert fuzzy control systems showed great promise, he encountered several difficulties during the realization of these two stages, which are:1) the correctness of the linguistic description of human expert actions, 2) the mathematical formulation of the linguistic description, and 3) the realization of software and hardware.

Unlike traditional Boolean systems, there exists no apparent demarcation in fuzzy logic expert systems between those documents that are not at all about a term and those documents in which the term is central, making it difficult to distinguish the border line separating relevant from nonrelevant documents. In this sense the boundary that separates documents that are relevant from those that are not relevant can be said to be fuzzy.

Unlike the early work in probabilistic retrieval, fuzzy-set theorists admit the importance of the current Boolean approach and make efforts to keep the basic Boolean framework while increasing its flexibility. Radecki(1982) suggested that "a combination of the existing theory of probabilistic retrieval into a practical methodology based on Boolean searches would be very promising."

4. Using Expert Systems as Search-Intermediaries

In designing many expert intermediary systems, certain common rules have been used which had already been developed in other AI systems. Brooks et al. (1986) have studied information interactions between human users and human intermediaries to derive specifications for an intelligent interface. They included functional discourse analysis as one of their methods. These investigations resulted in the specification for a distributed expert system architecture. Vickery (1984) describes the main features of the interaction between human and machine, and also characterizes essential elements that should be incorporated in an interface.

A number of information scientists suggest that we are on the verge of a major new kind of computer application which serves as an assistant in the performance of a difficult task. Recognizing the perceived advantages of expert intermediary systems, a number of such systems designed for information retrieval have been proposed in the last few years, and many of these have actually been implemented. Expert systems such as DENDRAL, MYCIN, and PROSPECTOR have proved the possibility of designing knowledge-based systems that can perform over a very limited domain. The performance levels of the systems approximate those of a human expert. More recently, expert systems have developed for use in information retrieval and online database searching.

Recent research and development activity has made efforts to experiment with the possibility that expert intermediary systems, performing as assistants to the end-user, could replace human search intermediaries(Marcus and Reintjes, 1981). The ultimate goal of designing an expert search intermediary is to allow the expert system to accomplish the same tasks as a good human intermediary in the same situation. I will now give some examples of currently operating expert system search intermediaries.

(1) Handling Certain Classes of Questions on a Particular Database

This approach has been followed by Pollitt(1987) in designing CANSEARCH, which is a rule-based expert intermediary system written in Prolog to handle cancer therapy questions in searches of MEDLINE. Unlike other online retrieval intermediary systems, CANSEARCH has been designed to assist a novice in the search of a restricted set of documents. In fact, it is not easy for beginners or infrequent users to select appropriate MESH terms to express their information need in a query form. The thesaurus is not easy to use and infrequent users are not well informed of indexing policies. CANSEARCH users do not need to know how to use the thesaurus. Using its knowledge of how to use MESH, CANSEARCH guides the user through the relevant sections of the thesaurus. The users can select the desired terms by pointing to a term on the screen, using a "touch-sensitive screen interface." In common with successful expert systems in other areas, the subject domain of CANSEARCH is well bounded and well defined. A test of CANSEARCH proved that the end-users using the expert system intermediary were successful in retrieving relevant information, and sometimes even outperformed human intermediaries. CANSEARCH is considered a promising example of an expert system applied to online retrieval.

(2) Assisting in Decision Making for the Section of Search Terms

This approach focuses on the type of decisions used in formulating a search strategy, such as the selection of search terms. Fidel(1986) illustrated in her article the way to model searching behavior of human intermediaries, by demonstrating that formal rules for the selection of search keys can be extracted from human experts. Even though these rules are not yet at the stage of being incorporated into a knowledge base of an expert system intermediary, the work she presented indicates that with more research a complete set of rules can be established. It also demonstrates that these rules can be automated to enhance significantly the adaptability of intermediary expert systems.

Shoval(1985) developed a prototype system which is designed to assist users to select the right search terms for a database search. The system uses a knowledge base, into which a thesaurus is incorporated. It is represented as a semantic network enhanced by adding information about terms, meanings and associations. The system allows the user to enter terms representing his/her query. The system then searches through its semantic network and selects other relevant search terms, and shows them on the list for the user to choose any of terms he regards as useful. The system performs the same job as the information specialist does. The system accepts the user's request expressed in his words, and as a result suggests a set of the possible candidate terms representing the user's problem. The system has a knowledge base which includes the expertise of human experts. The system does not necessarily simulate any specific expert, but incorporates the knowledge and the information processing performed by the information specialist into an intelligent computer program, with which the system can successfully carry out the tasks of the information specialist. Shoval (1985) indicated that an expert system for this kinds of task is necessary for following reasons:one is "the expansion and popularization of online databases," and the other is "the decentralization of computer systems and their users."

(3) Offering Advice While Keeping the End-User in Control of the Search Process

Meadow(1979) pointed out why a computer as a search intermediary is necessary. He developed an experimental system at Drexel University, IIDA (Individualized Instruction for Data Access) which was designed to provide diagnostic analysis of the user performances and to intervene by pointing out errors, offering advice, or giving instruction(Meadow, et al. 1982). IIDA watches the user's commands and suggests useful strategies when the user begins to repeat commands excessively or shows that help is needed. IIDA concentrated on search strategy aids and monitored the search as it was carried out.

The CONIT(COnnected Network for Information Transfer)(Marcus, 1983) system, developed and tested by Marcus for many years at MIT, demonstrates the evolution of an expert system intermediary with increasingly expert capabilities. Marcus considers that keeping concept representation and formation of the Boolean logic separate is important. CONIT assists the user in choosing a database, logs on to the appropriate database automatically, and helps to formulate the search statement. After the initial strategy is formed and passed to the databank host, CONIT suggests modifications after the first user's request if the search is not appropriate. Based on his experience with CONIT, Marcus concludes that a very extensive knowledge base development is required in order to provide a truly comprehensive expert assistant for the search process.

These two systems are good examples of the incorporation of expert system features, and demonstrate the areas in which further development is necessary.

5. Limitations and Difficulties

As Hawkins(1987) mentioned, "Most of today's systems are exploratory; only a few of them have emerged from the research or experimental stage. Many leaders in the information field have long recognized the advantages to be gained by applying the principles of AI to information retrieval, but little is available commercially yet."

In spite of the progress in the development of expert systems, none of the currently existing ones are full-fledged enough to be comparable to a human expert. Since expert systems are limited to a narrow subject domain, they cannot argue about the field of expertise, particularly any topic outside its domain: "They do not create their own axioms or theories, and they do not learn"(Borko, 1987).

In the last few years there has been considerable interest within the information society in the development of an expert system which is capable of retrieving "intelligently." Unfortunately, most of the past developments have focused on rather simplistic notions or definitions of what an expert system is, and they have consistently underestimated the difficulties involved in achieving expert system performance in any problem-solving domain(Brooks, 1987).

Successful applications tend to involve problem domains that are narrow, specialized, or homogeneous. Unlike the above domain, there are several limitations and difficulties for expert systems to be implemented in the library and information science environment. These are as follows(Vickery, 1987):1) the very wide subject domain of most library and information systems, 2) the language-dependence of library and information systems, and 3) the lack of knowledge about the tasks performed by expert librarians and intermediaries. Particularly, Brooks(1987) considers IR to be a less than ideal problem domain for an expert system application, since it is a domain that is neither well defined nor homogeneous. She points out that "in some retrieval environments and some aspects of the retrieval process, there may be no obvious human experts." On the other hand, Kehoe(1985) asserts that the first possible application of expert systems in a library and information system may be the fields of online database searching or cataloguing, in which human expert intermediary is required for everyday processing.

6. Conclusion

The development of online information services and the proliferation of databases has resulted in a new professional role for librarians or information scientists. This is the role of an intermediary who knows how to search information systems in order to meet the information needs of end-users. The intermediary is an integral part of a system which allows interaction between the user, the computer and the knowledge base.

It is said that end-users will easily find answers for their own requests when search processes become more simplified or user-friendlier. As the online environment grows to include more complex and diverse databases, the development of "intelligent" online assistants becomes a crucial research area. The prevailing approach to providing easier and friendlier user-system communication is to develop an "Intelligent Interface for Information Retrieval(IIIR) systems(Croft and Thompson, 1988). Researchers in many different fields are designing expert system as search intermediaries. A number of authors have argued that the successful use of expert system techniques in many information areas may not be possible until such techniques have been further advanced. However, advanced experimental intermediary techniques are now capable of providing search assistance whose effectiveness at least approximates that of human intermediaries in some contexts. Finally, information scientists envision future prospects for much more advanced systems which would automatically select databases and emulate human experts, and thereby make IR more effective for all classes of users. It is apparent that once more advanced techniques are fully developed for expert systems to be successfully applied to library and information systems, they will enable users to make more effective use of the automated systems and online databases that have been designed and implemented during the past decade. They will help the librarian be more productive and efficient in carrying out the many tasks involved in managing an information service center. Recent work on intelligent interfaces for IR systems and on information-seeking behavior has been useful to clarify some of the issues surrounding expert systems, making us aware of what intelligent retrieval might be like and of the system functions and components necessary to support an intelligent IR system(Brooks, 1987).

Although substantial progress has been made and much work has already been done, there is still a long way to go before full-fledged retrieval assistance based on expert system appears. But continuing research and development in this area shows great promise of deepening our understanding of the retrieval process from a basic scientific viewpoint, as well as improving search techniques themselves. The potential rewards and benefits of continuing research are, in my opinion, substantial(Smith, 1987).

〈Notes〉

- Brooks, H. M., "Expert Systems and Intelligent Information Retrieval", *Information Processing & Management* 23(4)(July 1987):367～382.
- Dejong, Gerald., "Artificial Intelligence Implications for Information Retrieval", *ACM SIGIR*(1985):10～17.
- Smith, Linda C., "Implications of Artificial Intelligence for End-User Use of Online Systems", *Online Review* 4(4) (October 1980):383～91.

- Smith, Linda C., "Artificial Intelligence and Information Retrieval", *Annual Review of Information Science and Technology* 22(1987):41～77.
- Borko, Harold., "Artificial Intelligence and Expert Systems Research and Their Possible Impact on Information Science Education", *Education for Information* 3(2) (May 1985): 103～14.
- Belkin, N. J., et al., "Distributed Expert－Based Information Systems:An Interdisciplinary Approach", *Information Processing & Management* 23(5) (September 1987):395～409.
- Forsyth, Richard., "The Architecture of Expert Systems", *In:Expert Systems, Principles and Case Studies*, Chapman and Hcomputing Ltd., 1984.
- Borko, H., "Getting Started in Library Expert Systems Research", *Information Processing and Management* 23(2) (March 1987):81～87.
- Vickery, A., Brooks, H., "Expert Systems and Their Applications in LIS", *Online Review* 11(3) (June 1987):149～165.
- Zadeh, L., "Fuzzy sets", *Information and Control* 8 (August 1965):338～353.
- Forsyth, Richard, "Fuzzy Reasoning Systems", *In:Expert Systems:Principles and Case Studies*, Chapman and Hall Computing Ltd., 1984.
- Bookstein, Abraham, "Probability and Fuzzy－set Applications to Information Retrieval", *Annual Review of Information Science and Technology* 20(1985):118～151.
- Negoita, C. V., "Fuzzy Sets in Knowledge Engineering", *In:Kacprzyk, J. and Yager, R. R. (Eds), Management Decision Support Systems Using Fuzzy Sets and Possibility Theory*, Verlag TUV Rheinland, 1985.
- Gevarter, W. B., "Expert Systems:Limited but Powerful", *IEEE Spectrum*, 20(8)(August 1983):39～45.
- Brooking, A. G., "The Fifth Generation Game", *In:Forsyth, R. Eds., Expert Systems:Principles and Case Studies*, Chapman and Hall Computing Ltd., 1984.
- Zadeh, L. A., "The Role of Fuzzy Logic in the Management of Uncertainty in Expert Systems", *Fuzzy Sets and System* 11(Novembe 1983):199～227.
- Kiszka, J. B. and Gupta, M. M., "Some Properties of Expert Control Systems", *In:Gupta, M. M.:Kandel, A., Eds, Approximate Reasoning in Expert Systems*, North－Holland, 1985.
- Radecki, T., "A Probabilistic Approach to Information Retrieval in Systems with Boolean Search Request Formulations", *JASIS*, 33(6)(November 1982):365～370.
- Brooks, H. M., Daniels, P. J., Belkin, N. J., "Research on Information Interaction and Intel-

ligent Information Provisio Mechanisms", *Journal of Information Science* 12(1−2)(January 1986):37∼44.

- Vickery, A., "An Intelligent Interface for Online Interaction", *Journal of Information Science* 9(1)(January 1984):7∼18.
- Marcus, R. S., Reintjes, J. F., "A Translating Computer Interface for End−user Operation of Heterogeneous Retrieval Systems;Part Ⅰ:Design:Part Ⅱ:Evaluations", *JASIS* 32(4)(July 1981):287∼317.
- Pollitt, Steven, "CANSEARCH:An Expert Systems Approach to Documental Retrieval", *Information Processing & Management* 3(2)(March 1987):119∼138.
- Fidel, Raya, "Towards Expert Systems for the Selection of Search Keys", *JASIS* 37(1)(January 1986):37∼44.
- Shoval, P., "Principles, Procedures and Rules in an Expert System for Information Retrieval", *Information Processing & Management* 21(6)(November 1985):475∼487.
- Meadow, Charles T., "The Computer as a Search Intermediary", *Online* 9(7)(September 1979):54∼59.
- Meadow, C. T., Hewett, T. T., et al., "A Computer Intermediary for Interactive Database Searching Part Ⅰ:Design, Part Ⅱ:Evaluation", *JASIS* 33(5)(September 1982):325∼332, 33(6)(November 1982):357∼364.
- Marcus, R. S., "An Experimental Comparison of the Effectiveness of Computers and Humans as Search Intermediaries", *JASIS* 34(6)(November 1983):381∼404.
- Hawkins, Donald T., "Applications of Artificial Intelligence(AI) and Expert Systems for Online Searching", *Online* 12(1)(December 1987):31∼43.
- Kehoe, Cynthia A., "Interfaces and Expert Systems for Online Retrieval", *Online Review* 9(6)(September 1985):489∼505.
- Croft, W. B., Thompson, R. H., "IR:A New Approach to the Design of Document Retrieval Systems", *JASIS* 38(6)(November 1988):389∼404.