DESIDOC Bulletin of Information Technology, Vol. 16, No. 4, July 1996, pp. 3-8

Expert Systems and their Application in Library and Information Science

Yogendra P Dubey*

Abstract

The article explains the concept of expert systems and how they function. The components of an expert system: the knowledge base, the inference engine, user interface and the knowledge acquistion module are explained. The various activities in the LIS field where such systems can be put to use are also discussed.

1. INTRODUCTION

Expert Systems have evolved from a long tradition of Artificial Intelligence (AI) research. Almost all types of organisations have recognised the enormous potential of expert systems which have forced the frontier of knowledge forward at a rapid pace. Expert systems are no longer confined to the research laboratory. Although, still in an evolutionary stage, these have added a new dimension to information processing. Having achieved remarkable success in both industry and commerce, expert systems are now denuding the frontiers of library and information systems. There is a significant surge in activities on assessing and debating the likely impact of expert systems on library and information profession.

Artificial Intelligence technology introduces a new paradigm for dealing with knowledge and reasoning process in human experts. This new paradigm emerged from the research efforts of Newell and Simon [1], who together had designed 'Logic Theorist', a programme that could prove the validity of theorems in Russell and Whitehead's *Principia Mathematica*. There was however, no intelligence in Logic Theorist; it worked on brute force trying every possible connection of symbols until something worked, but in one case the system discovered a more elegant proof than the one set forth by Ruseell and Whitehead. Logic Theorist laid the foundation for Newell and Simon's next project, the General Problem Solver (GPS), which they began in 1957.

The idea behind the GPS project was to build a machine incorporating problem solving techniques that could be applied to a broad range of problems. In the process of developing the GPS model, Newell and Simon investigated the information processing behaviour of human subjects. They discovered that people use domain specific knowledge as a basis for solving significant problem; and that the knowledge incorporates a large amount of fragmentary, judgmental, and heuristic knowledge. People reason with this knowledge forward and backwards at the same time, keeping track of the current status of the problem while simultaneously thinking

* *Professor Yogendra P Dubey*
Professor (Computer Application)
Department of Library and Information Science
Banaras Hindu University, Varanasi-221005

backwards from the goal they are trying to reach. This technique used by Newell and Simon for learning the way people reason is used today by knowledge engineers to create expert systems to solve complex problems.

By 1965, work in Artificial Intelligence had progressed to the point at which Edward A Feigenboum and Joshua Ledenberg, both from Stanford University, could join forces to develop a system for determining the structure of an organic molecule based on the input obtained from mass-spectrograph. The problem was the one that taxed the abilities of an expert organic chemist: there were too many possible solutions to allow the use of pure logic, and chemistry experts solved the problem using a combination of intuition, rules of thumb, and just plain good guess-techniques that were known as heuristics[2].

By interviewing organic chemists on the methods they used to develop a picture of a molecule's structure, Feigenbaum and Lederberg developed a combination of rules and heuristics for interpreting spectrographic data. The result was a programme called 'Dendral'; its descendants are used today in laboratories all over the world. Dendral marked the beginning of a new branch of inquiry for researchers in Artificial Intelligence–that is expert systems. Dendral's immediate successor was 'Mycin', an expert system which has been even more influential. Mycin diagnoses bacterial infections of the blood, and prescribes suitable drug therapy.

2. FEATURES

Although, expert systems are the subject of much discussion these days, their definition is somewhat vague. If basic definitions from some introductory texts on expert systems are examined (such as those of Barr and Feigenbaum[3]; Forsyth[4]; Harmon and King[5]; Hart[6]; and Jackson[7]) then it is found that only two elements are consistently cited:

1. An expert system embodies some representation of knowledge about a given task domain.
2. The expert system emulates the capabilities of a human within the given task domain at a level of performance equivalent to an 'expert'.

Whitaker and Ostberg[8] have observed that these requirements presuppose that the task domain is of such a nature that consistent facts about it can be derived, i.e., there is knowledge about the domain and that there is differential distribution of said knowledge among some population (there are individuals possessed of greater relative expertise). Such presumptions address the character of the task domain into which an expert system is inserted.

In contrast, much of the literature emphasise on the expert system's structure. Most commonly the discussion on an expert system's structure concentrates on two discrete components–a knowledge base, and an inference engine carrying out logical operations over that base. Occasionally a discussion extends to the four essential components of a full-fledged expert system.

3. COMPONENTS OF EXPERT SYSTEM

A detailed discussion of the components of an expert system is presented here to clarify this further.

3.1 The Knowledge Base

The knowledge base in an expert system contains codified knowledge that is structured very differently from book knowledge. Typically, the knowledge base of an expert system contains 'facts' and the relation between them. There are two kinds of facts : public facts and heuristics. Public facts are defined as published rules (generally available) and agreed knowledge, e.g. cataloguing rules. Heuristics are the personal (and often unwritten) knowledge of the human expert (e.g., skill in cataloguing). In addition to facts, heuristics imply several features of human decision making based on experience, perhaps acquired on a trial-and -error basis, 'rules of thumb', and inability to take all possible factors into account due to cognitive limitations.

The major difference between 'conventional' database and expert system database methodology is that a knowledge base of an expert system is more creative. Facts in a database are normally passive; they are either present or not present. A knowledge base on the other hand, actively tries to fill in the missing information. Production rules are a favourite means of encapsulating 'rule-of-thumb' (heuristics) knowledge. These have a familiar IF (CONDITION)–THEN (CONSEQUENCE) format , for example:

IF frequency of word < = upper threshold AND frequency of word > = lower threshold THEN use word as index term.

There are other methods also to encode the facts and relationships that constitute knowledge. They include: semantic net; object-attribute-value triplet; frames and predicate calculus. It might be inferred from some texts that the methods are alternatives; one or the other is used exclusively. In fact, they can be used in association with each other. Different methods have different advantages and disadvantages[9].

3.2 The Inference Engine

Inference engine stands between the user and the knowledge base. It performs two major tasks: first, it examines existing facts and rules, and adds new facts when possible; second, it decides the order in which inferences are made. In doing so, the inference engine conducts the conclusion with the user.

The inference strategies used in expert systems are:

- ☐ Modusponens
- ☐ Reasoning about uncertainty
- ☐ Resolution

Engine part of the system means it is used to drive around amongst the various inferences it might make. This means it allows to pursue reasoning strategies or control strategies as they are sometimes called, to decide what operators to apply at each stage of the search. The most common control strategies used in expert systems are forward chaining, backward chaining and bi-directional. Broadly speaking, forward chaining involves reasoning from data to hypotheses, while backward chaining attempts to find data to prove, or disapprove, a hypothesis. Pure forward chaining leads to unfocused questioning in a dialogue mode system, whereas pure backward chaining tends to be rather relentless in its goal-directed questioning. Most successful expert systems use by-directional reasoning method[10].

Whether inferencing procedure works primarily backwards or forwards, it will have to deal with uncertain data. There are too many ways of dealing with uncertainty. There is Fuzzy logic, Bayesian logic, multi-valued logic and Certainty Factor, to name only four.

3.3 User Interface

User interface is the component which enables the user to communicate with the expert system. Most expert systems are interactive; they need users to input information about a particular situation before they can offer advice. Most of the existing user interfaces of expert systems are menu-driven, accepting single words or short phrases from the human user. A few have natural language capabilities. A good user interface is an expert system that will allow the user[11]:

- to ask question, such as why an advice has been given, how a conclusion has been reached or why certain information is needed
- to volunteer information before being asked
- to change a previous answer
- to ask for context-sensitive help on demand
- to examine the state of reasoning at any time
- to save a session in disk for later perusal, and
- to resume a session previously abandoned mid-way.

The expert system aims to assist or advise the non-expert users, but it will be consulted only if it helps the users to perform that task more easily. Thus, expert systems must meet all the requirements of good interface design.

3.4 The Knowledge Acquisition Module

Knowledge acquisition is the most challenging and important element in expert system development. It is the process of eliciting knowledge from experts and converting it into facts and rules that are compatible with the knowledge representation form of the chosen development tool.

Feigenbaum[12] calls those who build knowledge-based expert systems as 'Knowledge Engineers', and refers to their technology as 'Knowledge Engineering'. Knowledge engineers are concerned with identifying the specific knowledge that an expert uses in solving a problem. Initially, the knowledge engineer studies a human expert and determines what facts and rules-of-thumb the expert employs. Then the knowledge engineer determines the inference strategy that the expert uses in an actual problem solving situation. Finally, the knowledge engineer develops a system that uses similar knowledge and inference strategies to simulate the expert's behaviour.

4. SCOPE OF EXPERT SYSTEMS IN LIBRARY AND INFORMATION SYSTEMS

Expert systems are knowledge intensive computer programmes. They contain lot of knowledge about their speciality. They use rules-of-thumb or heuristics, to focus on the key aspects of particular problem domain and to manipulate symbolic descriptions in order to reason about the knowledge they are fed with. They often consider a number of competing hypotheses simultaneously, and they frequently make tentative recommendations or assign weights to alternatives. The best expert systems can solve difficult problems within a very narrow domain, as well as or better than human experts can.

All this suggests that expert systems are confined to well-circumscribed tasks. They are not able to reason broadly over a field of expertise. They cannot reason from axioms or general theories. They do not learn and thus, are limited to using the specific facts and heuristics that are 'thought' by a human expert. They lack common sense, they cannot reason by analogy, and that their performance deteriorates rapidly when problems extend beyond the narrow task domain they are designed to perform. A suitable domain should have clearly circumscribed borders and contain a limited number of entities and relations.

Now the question is whether the library and information science (LIS) field is suitable for application of knowledge engineering to build expert systems? Brooks[13] and Sparck Jones[14] claim that subdomains of LIS field are suitable for application of knowledge engineering as they possess clear-cut boundary with lots of public and heuristics knowledge which can be identified, elicited, codified and formalised in the knowledge base. Their claim is vindicated by a study reported in Poulter and others[15] which has identified 139 expert system projects in the LIS domain.

However, systems are usually grouped in literature by the sub domain (e.g., online intermediary, reference work, cataloguing, indexing, classification, etc.) that they address. Some particular LIS sub domains have received relatively little attention (e.g., abstracting) while others (e.g., acquisition of new stock) are hardly represented at all. A breakup of the avilability of expert systems in LIS field is given below:

reference work	32 percent
online retrieval	28 percent
indexing, cataloguing, or classification	12 percent
library management application, and	7 percent
abstracting	1 percent

Within each sub domain, systems tend to be restricted to even more limited domain to avoid problems associated with ill-defined borders and large number of entities and relationships. For example, 'Plexus' by Vickery and Brooks[16] of expert system devoted to a very narrow domain of gardening while 'Pointer' by Smith[17] is limited to US Government publications and 'Maper' by Ercegovac[18] which advises novice map cataloguers.

5. CONCLUSION

We are witnessing gradual increase in the development and research activities in the LIS expert systems. However, as already stated, some areas such as reference and online intermediaries are more appealing to researchers than others. In Western countries, knowledge engineers are gearing up for the production of commercial expert systems in LIS domain, but most work is still in experimental stage. It is however, expected that expert systems may find increasing application in special libraries (such as DESIDOC) which may collect and codify knowledge within their areas of specialisation.

REFERENCES

1. Newell, A. and Simon, H. Human problem solving. Prentice Hall, Englewood Cliffs, NJ. 1972.
2. Williamson, M. Artificial intelligence for microcomputers: The guide for business decision maker. Brady Communications Company, New York. 1986.
3. Barr, A. and Feigenbaum, E.A. (Ed.). The handbook of artificial intelligence. Addison-Wesley, Reading, MA. 1987.
4. Forsyth, R. (Ed.) Expert Systems: Principles and case Studies. Chapman and Hall, New York. 1984.
5. Hermon, P. and King, D. Expert systems: Artificial intelligence in business. John Wiley, New York. 1965.
6. Hart, A. Knowledge acquisition for expert systems. McGraw-Hill, New York. 1986.
7. Jackson, P. Introduction to expert systems. Addison-Wesley, Warkingham, England. 1986.
8. Whitaker, R and Ostberg, O. Channelling knowledge: Expert systems as communications media. *A I & Society*. 1988, (2), 197-208.
9. Dubey, Y.P. Knowledge representation and reasoning. *In* Encyclopaedia of Microcomputers, edited by Allen Kent and James G Williams. Marcel Dekker, New York, Vol.15.
10. Kemp, D.A. Computer-based knowledge retrieval. ASLIB, London. 1988. pp.103-10.
11. Ref. 4
12. Ref. 5
13. Brooks, H.M. Expert systems and intelligent information retrieval. *Information Processing Management*. 1987, **23**(4), 367-92.
14. Sparck Jones, K. Information retrieval. *In* Encyclopaedia of artificial intelligence. Edited by Shapiro, SC. John Wiley & Sons, New York. 1987, Vol. 1, 419-21.
15. Poulter, A. and others. LIS professionals as knowledge engineers. *In* Annual review of Information Science and Technology (ARIST). The American Society for Information Science. Medford, NJ, Learned Information, Inc. 1994, Vol.28.
16. Vickery, A. and Brooks, HM. PLEXUS: The expert system for referral. *Information Processing and Management*. 1987, **23**(2), 99-117.
17. Smith, K F. Pointer vs. Using government publications: Where is the advantage? *Reference Librarian*. 1989, **23**, 191-205.
18. Ercegovac, Z. Research on knowledge-based descriptive cataloguing of cartographic publications: An experimental advice-giving system-mapper. Ph.D. Dissertation. 1990, UCLA, Los Angeles, CA.