EXPERT SYSTEMS FOR LIBRARY AND INFORMATION SERVICES *

S S Murthy

1. CONCEPT AND DEFINITION

An expert system (ES) acts as an expert or a consultant in some subject area and helps in decision-making in unstructured problem situations[1]. The expert systems comprise computer programs modelled after human experts. These systems rely heavily on knowledge, especially knowledge of their domain[2]. These systems store information relevant to the problem domain and seek to emulate problem-solving methods of human experts.

Many definitions of ES have been provided by experts in this field, but the one given by the British Computer Society Expert Systems Specialist Group (BCS ESSG) appears to be adequate and comprehensive:

"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"[3]

2. STRUCTURE AND OPERATION

A simplified form of an ES structure is given below :
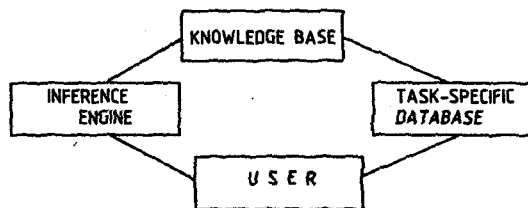
Computer programs normally organise knowledge at two levels— data and program. According to Yaghmai and Maxin[4] most expert systems organise knowledge at three levels—data, knowledge base and control.

The database of an ES system contains the data that is used for providing expert advice. This database keeps track of input data for the current problem, the status of the current

---

problem and the relevant history of what has been done so far. The knowledge base contains the procedural knowledge specific to the problem area or domain. The inference engine provides the control structure of the ES. It makes decisions about how to use the system's knowledge by organising and controlling the steps it takes to solve the current problem. This control mechanism is actually a collection of procedures or strategies on how to solve problems.

Since most expert systems store the knowledge on the problem area as production rules, the knowledge base is referred to as the rule base, and the control mechanism is called the rule interpreter. Rules in the knowledge base are actuated by input from the database. This rule application modifies the database and identifies the pertinent rules. The rule interpreter decides which rules to apply via inference procedure or conflict-resolution strategies that match the rules in the knowledge base. All rules are evaluated during every cycle as the system tackles current problems and data, and keeps track of the status of the development of solutions.

3. DEVELOPMENT

3.1 Seven stages

The development of ES can be visualised in seven stages. These are: (i) system design, (ii) system development, (iii) formal evaluation of performance, (iv) formal evaluation of acceptance, (v) extended use in a prototype environment, (vi) development of maintenance plans and (vii) system release. Yaghmai and Maxin claim that no current system has completed all these stages so far.

3.2 Trends

Yaghmai and Maxin visualise two kinds of trends, the short term trends and the long term trends. According to them, the short term, ES with thousands of rules will appear. In addition to an increase in rule-based systems, there will be an increase in the number of non-ruled-based systems. By the late 1990s, there

will be knowledge acquisition systems which, after being given the content of a basic problem area, will rapidly guide a human in the formation of the needed knowledge base. Around the year 2000, systems will semi-autonomously develop knowledge bases from text. The result of these developments may very well herald a maturing information society where expert systems put expert knowledge at everyone's disposal.

4. ES FOR LIBRARY AND INFORMATION SERVICES

4.1 Computerisation — A Pre-requisite

Computers have been used in libraries for over two decades in advanced countries like UK and USA. Many libraries and information centres made impressive achievements in automated operations like circulation, cataloguing, serials control, acquisition, etc., and also information storage, retrieval and dissemination functions. In India also, a few large libraries computerised some of these operations over the past decade. However, the use of ES programs (heuristic programs) in this field is relatively new even in advanced countries where it is yet to record notable achievements, while in India, it has yet to enter even the planning stage. But the work in this direction has progressively been increasing in advanced countries over the last eight to ten years. Automatic classification, cataloguing, translation, and information retrieval are some of the specific areas where these efforts are going on and the potential for application of ES in this field is now well understood which imparts good impetus to on-going efforts.

4.2 Applications of ES

A review of recent literature on the development of ES for library and information services not only reveals the possible areas of such application, but also acquaints one with the studies conducted on these applications in certain organisations. The applications, both tried and planned, and also some of the potential areas could be grouped as follows:—

a) Subject classification
b) Cataloguing

c) Reference service

d) Indexing

e) Specialised thesauri

f) Information retrieval and

g) Publishing.

4.2.1 Subject Classification

Classification of documents was considered as a problem of pattern recognition in the ES studies. Since the classification is generally based on the subject content of the documents which can be represented either by codes (class numbers) or by keywords, the system matches the subject content of a new addition to the library collection with the existing codes or keywords for achieving automatic classification. Earlier studies in automatic classification used factor analysis or other clustering algorithms to create subject matching or subject associations based on keywords in the text[5]. Although these studies were helpful in demonstrating the possibilities of automatic classification, they were not adequate to handle large collections. Further work in this direction is in progress.

4.2.2 Cataloguing

There are different sets of rules for cataloguing. A scrutiny of any of these rules would convince any one that there are many complexities in the cataloguing function, particularly in entering the authorship of a document or work. An expert system can store the rules according to a particular code and apply after taking appropriate decisions with regard to the authorship, type of document, linkage between multiple volumes of document and so on. An ES model for cataloguing[6] of maps was recently developed at the University College of Los Angeles (UCLA) on similar lines with MYCIN, an expert system on diagnosis of infectious diseases.

Another application of ES in this area could be the development of a suitable system for applying relevant cataloguing rules to a document while the text of the document is still under electronic composition.

4.2.3 Reference Service

A prototype ES for reference service was designed and tested at UCLA in 1980 using a TRS 80 microcomputer. The system was called REFLES— Reference Librarian Enhancement System. The objective of the system was to enhance the performance of a reference librarian by providing him access to ephemeral data, non-current data and data derived from knowledge of local resources. The study showed that with improved hardware and software facilities, the system could be augmented to encompass other reference material, search strategy development and user participation making it a true expert system.

Vickery and Brooks[7] describe an ES developed at the University of London, titled PLEXUS covering the subject of gardening. This system was expected to provide reference service based on the description of the information requirement given by the user, and by supplementing the user's statement with additional search terms or concepts from the knowledge base of the system and by making the search strategy that could be applied to a database. The development of its prototype took about 20 months and further work is in progress.

4.2.4 Indexing

Indexing of documents is highly subjective relative to the types of users, their specialisations, the depth of indexing required to be done for meeting individual requirements of users, the reading habits of the users and so on. Slant in indexing to provide access to the information required by different user groups from the same document is also an important factor. So, an ES, which would store knowledge to satisfy different view points of the users and retrieves information that matches their specific requirements, can be highly useful.

4.2.5 Specialised Thesauri

The quality of a good thesaurus is to be measured by its ability in recognising and incorporating new subject fields, ignoring obsolete

ones automatically. No single thesaurus would normally meet such requirements unless it is specially designed for the system and is constantly updated. On the other hand, the ES which can take care of the changing requirements of the users of the IR system and can constantly update the thesaurus based on its expert knowledge. Such an ES would be highly useful for IR system in general. No such ES appears to have yet been developed.

4.2.6 Information Retrieval

For retrieving information from a computerised IR system, the practice has been that the user either himself provides a search expression or describes the information requirement from which the search expression and a search strategy are worked out by an information specialist. The search is then conducted and the feedback is discussed by the information specialist with the user and the search expression modified as necessary and searches are again conducted till the user gets the most relevant information possible. Studies using ES for information retrieval included automatic formulation of search strategy and search expression based on the inputs provided by user through a menu provided by the system. The user gives his own version of his information requirement using his own terminology. The heuristic procedures used by the system, using its database and the knowledge base select the appropriate vocabulary, and formulate the search strategy and conduct the search. Further modifications in the search, based on the feedback provided by the user, are also carried out automatically. The system also attaches weights to the search terms based on its structured dialogue with the user, and in the repeated searches the weights are also modified suitably.

An ES entitled CONSEARC[8] was recently designed to enable doctors to specify queries to retrieve cancer— therapy related documents stored in the MEDLINE database of the US National Library of Medicine. The system embodies the expert knowledge of a librarian functioning as an intermediary between the physician and the database.

4.2.7 Publishing

Some traditional reference publications like Beilstein's Handbuch der Organischen Chemie are not only difficult to compile but are complicated to use. Some publications like, directories and handbooks get outdated before they are distributed. Development of a suitable ES for providing up-to-date and ready data or information to answer specific queries incorporating expert advice to a particular problem for which the query is made, would be highly valuable.

As the information explosion continues, reviews and digests covering the literature output in specific fields, have become very valuable to the working scientists as well as the busy technocrats or executives. So, an ES which automatically prepares up-to-date digests with different subject slants to meet the requirements of different users and linked with an electronic publishing system to provide on-the-spot copies of the digests or reviews, would be greatly welcomed by these busy users.

5. FUNDING of ESs

In USA alone a substantial amount ($ 10 million) is being spent by US government on some of the ES systems and $ 2 million under government contracts to non-profit organisations and universities. Some of the ES systems funded by US government are: Diagnosis of infectious diseases (MYCIN), Oedema patients (PIP), Diagnosis and treatment for glaucoma (CASENET), and knowledge acquisition and problem formulation (ROGET). ES research sites in USA include universities, non-profit organisations, government depts and industrial units.

6. CONCLUSION

As stated earlier, the use of ES for library and information services is not yet very common, and significant achievements in this area are still to take place, although some studies have

been made. These studies have demonstrated the high potential for the use of ES in the libraries and for information storage, retrieval and dissemination services including electronic publishing. In India, computerisation of library and information services has recently received a momentum with various national agencies actively carrying out work in this area and it will not be long before use of ES is also made for operating these information services.

REFERENCES

1. Shoval, P. *Inform Process Manage* **22(6)**, 1985; 475—87.

2. Sowizral, HA. *Annual Rev Inform Sci Technol* **20**, 1985; 178—91.

3. Pollitt, AS. *Aslib Proc* **36(5)**, 1984; 229—34.

4. Yaghmai, NS. and Maxin, JA. *J Amer Soc Inform Sci* **35(5)**, 1984; 297—305.

5. Borko, H. *Education Inform* **3(2)**, 1985; 103—14.

6. Borko, H. *Inform Process Manage* **23(2)**, 1987; 81—7.

7. Vickery, A and Brooks, HM. *Inform Process Manage* **23(2)**, 1987; 99—117.

8. Pollitt, S. *Inform Process Manage* **23(2)**, 1987; 119—38.

My message to the managers and administrators is to care for their workers, to treat them like human beings and not like cogs in a machine. Only then will they be able to motivate them and carry them together in a team.

MK Rustomji