CHAPTER 1

INTRODUCTION

1.1 BACKGROUND

Research into expert systems can be traced back to Alan Turing, Allen Newell, Herbert A Simon, John McCarthy and Edward A Feigenbaum who explored the general problem of designing “intelligent” machines (Giarratano and Riley, 1993). Artificial intelligence (AI) research is defined as ‘the part of computer science that is concerned with designing intelligent computer systems, that is, systems that exhibit the characteristics we associate with intelligence in human behaviour - understanding language, learning, reasoning, solving problems, and so on’ (Barr and Feigenbaum, 1981). AI started in the 1950s when it was recognised that computers could be used as logic machines that could process symbols, expressed as numbers, letters of the alphabet, or words in a language (Borko, 1985). It could be used in well-defined and well-constrained problems such as symbolic algebra, chess playing or game playing in general, puzzle solving, and simple theorem proving.

Some researchers started asking how people solved real-world problems. They found that people used problem-specific knowledge to help simplify their reasoning. Experts use domain knowledge (the knowledge surrounding the problem) to help them understand the issues, to suggest possible approaches, to judge the reliability of facts, or to evaluate the effectiveness of the solutions they generate. Thus began the development of expert systems.
An expert system is a computerised advisory programme that attempts to imitate or substitute the reasoning processes and knowledge of experts in solving specific types of problems (Turban, 1993). More precisely, it solves real-world problems requiring an expert's interpretation, employs heuristic (rules of thumb) knowledge and/or qualitative models of the problem domain (Pfeifer and Luthi, 1987), reaches the same conclusions that a human expert would have if faced with a similar situation, and employs a programming methodology based on a separation between knowledge and its application.

Human experts are scarce and expensive. The objective of an expert system then is the dissemination of expertise: transferring expertise from some expert(s) to a computer and then on to other human non-experts. The problem areas addressed by expert systems include interpretations, prediction, diagnosis, design, planning, monitoring, debugging, repair, instruction and control (McGraw-Hill, 1994). Building an expert system requires extracting relevant information, encoding that information in a form amenable to mechanical manipulation, deciding on an inference mechanism, and verifying that the system performs correctly.

In an expert system, three possible modes can be envisaged for interaction between the user and the system, viz. the user as a client, trying to solve a problem; the user as a tutor, providing knowledge for the system; and the user as a pupil, trying to learn from the accumulated information (Clarke and Cronin, 1983). Expert systems have and are being developed for many varied applications, among which are
medicine, fault diagnosis, science, electronics, engineering, computer science, geology and library science (Giarratano and Riley, 1993). This will be discussed in greater depth in Chapter 2.

In the field of library and information science (LIS), expert systems have been applied to reference work, online retrieval, indexing, cataloguing, classification, library management applications and abstracting (Poulter, Morris and Dow, 1994). The primary reason for developing expert systems for various areas in LIS is to bring the improvements that technology can provide to bear on existing tasks. Cataloguing has long been considered a bottleneck in the process of attempting to get newly acquired material to the patron in a timely fashion. Cataloguing therefore seems to be fertile ground for the development of expert systems application for reasons below:

(i) It is a rule-based activity.

(ii) The AACR2R (Anglo-American Cataloguing Rules, 2nd Edition, 1988 Revision) constitute the totality of guidance required in order to provide bibliographic description for a catalogue entry. The presence of a formalised code makes it feasible to develop a knowledge base.

(iii) Considerable numbers of professional librarians devote an appreciable amount of their working time to catalogue-related activities. The development of an expert system would reduce staffing costs or at least contain the increase in those costs.

Cataloguing involves three basic activities: describing the item and choosing access points for names and titles, assigning classification numbers, and assigning subject
headings. Most of the research to date has focused on the first activity, descriptive cataloguing, because it uses a rule-based cataloguing code, AACR2R (Hawks, 1994). However, the task is far from easy as cataloguers have not only to decide which rules to use, but also to decide the order in which they should be applied and how to interpret them. This is especially so in the area of cataloguing conference proceedings.

The creation of bibliographic records for conference publications is difficult for cataloguers. Those identified only by a generic term for a meeting and the name of the corporate body holding the meeting present special problems. For this study, conference proceedings are defined as collections of papers presented at a meeting organised or sponsored by a corporate body and published as a monograph. Titles of these works often consist of generic words such as “proceedings” or are non-existent, e.g. when the title page identifies the work as “the eleventh meeting of the such-and-such corporate body”. It is usual for particular meetings to be held on a periodic basis; they may or may not be numbered; and the title frequently fails to remain the same from one meeting to the next (McGarry and Yee, 1990).

Problems arise too in deciding whether conference proceedings should be viewed as monographs or serials when in published form. In addition, there is the problem of erratic publication patterns, variety of form of publications, and cataloguing procedures (McGlasson, 1983). This study aims to develop a detailed and specialised system where the knowledge base contains not only information about the relevant
rules in AACR2R, but also the knowledge of experts, in this case expert cataloguers of conference proceedings.

1.2 CONTEXT OF STUDY

The purpose of this study is to develop an experimental cataloguing expert adviser system that uses the domain of descriptive cataloguing, an expensive and complex portion of the cataloguing process, for published conference proceedings. The adviser's primary goal is to educate novice cataloguers in creating bibliographic records for published conference proceedings as well as to improve conventional instruction in the cataloguing of conference proceedings. In addition, the system also aims to serve as a consultant to the professional cataloguer of conference proceedings.

The system is essentially an interactive tutor that asks specific questions to elicit information, advises, and explains its decisions. It incorporates portions of AACR2R, heuristics obtained from expert cataloguers of conference proceedings, and sample bibliographic records. Specifically, the system aims to familiarise the novice cataloguer with the specialised vocabulary of descriptive cataloguing of conference proceedings; to simulate an actual cataloguing environment by providing the user (the novice and professional cataloguer) with a graphic representation of the item being catalogued; and to lead the novice through the creation of bibliographic records.
Within the context of this study the distinction between a novice and an expert needs to be explained. Experts are expected to know more and they use all their knowledge more effectively. Experts integrate facts with experience and reorganise it over time into more efficient memory units, a process which has been termed as chunking (Kolodner, 1984; Barfield, 1986). Experts 'just know' often without being able to explain why or how they know. The rules they originally learnt have sunk out of consciousness in a process of procedural embedding or compilation, leading experts to base their reasoning on cases or analogies rather than rules (Slatter, 1985; Kolodner, Simpson and Sycara-Cryanski, 1985). Consequently, experts often have difficulty in verbally expressing the reasons for their decisions and actions (Neale and Morris, 1989). Nevertheless, an awareness of these and other sources of distortion prior to conducting this study may alleviate the problems in knowledge acquisition. This study involves collecting a sample which exemplifies the various types of published conference proceedings; identifying experts' mental modelling when assigning bibliographic records for the proceedings; developing an expert adviser system and finally testing the system's performance.

1.3 NEED AND IMPORTANCE OF THE STUDY

There is a need and importance for this study to be carried out for the following reasons:

(i) There is a need to economise on the costs of cataloguing. An expert adviser system that could automate the professional aspects of cataloguing, in addition to the clerical procedures, would have a correspondingly greater
potential for achieving economies especially since professional librarians are paid more than clerical staff.

(ii) Novice cataloguers that are churned out through library science courses are inadequately prepared for the various types of materials that need to be catalogued which include conference proceedings. An expert system, used as a computer-aided-instruction package or adviser system, will be able to aid the novice cataloguer to better understand the various rules and thus be able to undertake cataloguing published conference proceedings more efficiently.

1.4 AIMS AND OBJECTIVES OF THE STUDY

The aims of this study are:

(i) To take account the private / personal, and public knowledge (McGraw-Hill, 1993) necessary and needed in giving bibliographic descriptions to conference proceedings which makes up the knowledge base of the expert adviser system. Public knowledge, also known as published knowledge is taught in cataloguing courses, and its tenets are generally accepted and practised by the Anglo-American cataloguing community. This body of knowledge includes published cataloguing principles, rules, rule interpretations, facts, theories and definitions contained in textbooks and references. Private knowledge or “knowledge by acquaintance” is a direct and specific sort of knowledge which expert cataloguers acquire by means of their long-term dedication to and scholarly practise of cataloguing.
(ii) To test the relative success rate of novice cataloguers in cataloguing published conference proceedings.

Among the objectives of this study are:

(i) to help novice cataloguers in solving a set of tasks in the domain of descriptive cataloguing of conference proceedings,

(ii) to identify interface characteristics related to the design and implementation of the system, and

(iii) to evaluate the performance of the expert adviser system in giving advice in the descriptive cataloguing of conference proceedings to the novice cataloguer.

1.5 RESEARCH PROBLEMS

Conference proceedings are indeed difficult to catalogue as shown through studies conducted and reported in the literature (McGarry and Yee, 1990; Zainab, 1991, 1996). Novice cataloguers are faced with the momentous task of determining if a work is a proceeding or not, and then to catalogue the proceeding. These inexperienced cataloguers have to be trained to construct main headings for published conference proceedings based on AACR2 rules. Confusion arises in situations such as when the title page carries both a complete conference statement as well as the corporate body organising the event; or in the case of constructing the main heading by wrongly including the entire date information in parentheses after the conference name. By mapping out the thought processes of expert cataloguers,
and from there implementing it to the expert system, it is hoped that the prototype system will aid these novice cataloguers in overcoming these cataloguing problems.

Research has shown that the use of an expert system as a computer-aided instruction tool in cataloguing training is feasible (Zainab, 1996) and would alleviate problems faced by novice cataloguers. An expert system could:

(i) make experts' knowledge accessible to professional cataloguers, novice cataloguers, as well as occasional users, and

(ii) relieve professional cataloguers from concerns for routine details allowing them to concentrate on critical tasks.

In addition, the following problems have been addressed:

(i) how many types/categories of conference proceedings there are and the characteristics that differentiate them,

(ii) how the expert cataloguers handle the different types of proceedings,

(iii) how the expert knowledge or expert practice can be best encoded and represented in a computerised format,

(iv) what are the design recommendations for the expert adviser system,

(v) what are the dependent variables for evaluation of the system's performance.
1.6 RESEARCH QUESTIONS AND HYPOTHESES

Based on the problem of categorising conference proceedings, the following research questions were formulated:

(a) What constitutes a main entry heading for a conference proceeding?

(b) Having identified the main entry heading, how does the bibliographic description for the proceedings take shape?

(c) What should be included in the subject headings for conference proceedings?

(d) How should the author mark in the call number be noted for conference proceedings?

To develop design recommendations for the expert adviser system, the following research questions were asked:

(e) How can expert cataloguers' knowledge be articulated sufficiently to incorporate them into a suitable interface design?

(f) What are the set or sets of boundaries among cataloguing access points for which various assistance modes can be made?

The system's performances on the dependent variables, which are determined later, were evaluated in a laboratory setting. The main question was:

Will the treatment group do cataloguing tasks "more accurately" than the control group?

The term "more accurately" means statistically significant higher mean scores of each of the answers included in the tasks.
Based on past research in this area of implementing expert systems in cataloguing, the following null hypothesis was formulated:

There will be no significant difference in mean scores obtained between the experimental group (treatment group) and control group in cataloguing all types of identified published conference proceedings.

More specific hypotheses are given in Chapter 5.

1.7 ASSUMPTIONS

This study assumes the following:

(i) The sample is undergraduate library science students who have been exposed to a semester of introductory cataloguing and cataloguing conference proceeding classes. It is expected that this sample would be competent in cataloguing conference proceedings.

(ii) The sample is computer literate, therefore there was no induction course on how to use the system as it will be Windows based and user-friendly.

1.8 METHODOLOGY

This is an experimental research in developing a prototype expert adviser system for use in the cataloguing of published conference proceedings. The study was carried out in three parts as explained earlier under the section regarding the context of the study. A detailed description of the methodology is given in Chapter 3.
1.9 LIMITATIONS

The limitations of this study are:

(i) The study is limited to undergraduate library science students.

(ii) The software to be used to build the expert adviser system is Asymetrix Multimedia Toolbook II where the interface design is user friendly and Windows based making it easy to use.
REFERENCES


