An expert selector of online public access catalogues.

BLACKADDER, A.

1992

An Expert Selector of Online Public Access Catalogues

by
Alistair Blackadder
May 1992

The Robert Gordon Institute of Technology
School of Librarianship and Information Studies
in collaboration with
Dundee Institute of Technology Library

A thesis submitted in fulfilment of the requirements for the award of a Master of Philosophy Degree of the Council for National Academic Awards

PRINT FADED IN ORIGINAL

Abstract

This project looks at the possibility of creating an expert system which would aid with access to OPACs available on the JANET network. This system would aid access by being able to select the OPAC most suited for a search in a particular subject area.

Looked at initially are the OPACs available and the JANET network. The problems with access and use of these OPACs are demonstrated. Expert systems and how they operate are looked at in relation to how an expert system might be used to solve these problems.

The uses already made of expert systems within library and information services are examined in order to establish the potential of such a system. Examples are given of systems in order to demonstrate this potential.

The differences between an expert selector and an inexpert selector are then examined. The system objectives required to make such a system "expert" are defined.

Tools for constructing an expert system are detailed. The different options are considered and the selection of the expert system shell Crystal is examined.

The knowledge required to construct an expert system is then examined. Problems in knowledge

acquisition are considered and the knowledge to be used in this project is evaluated, particularly the Scottish Conspectus project.

Details are given of how this knowledge is used to create an expert system using the Crystal expert system shell. How the system operates and how it would communicate with the OPACs available through JANET is examined.

The conclusion is reached that an expanded version of the demonstration system produced as a part of this project could be a valuable tool in aiding access to sources of information that are currently underused.

Acknowledgements

Douglas Anderson

(Senior lecturer, The Robert Gordon Institute of Technology. Director of studies.)

David Coleman

(Lecturer, The Robert Gordon Institute of Technology. Project supervisor.)

Robert Newton

(Lecturer, The Robert Gordon Institute of Technology)

Ivor Lloyd

(Chief Librarian, Dundee Institute of Technology Library)

Reg Pringle

(Chief Librarian, University of Aberdeen Library. Provision of access to Scottish Conspectus database.)

Contents

Abstract 1

Acknowledgements 3

Contents 4

1. Introduction 7

1.1 JANET OPACs 10

1.2 Expert systems 16

1.2.1 What is an expert system? 16

1.2.2 Components of an expert system 17

1.2.3 The knowledge base 18

1.2.4 The inference engine 28

2. Uses of expert systems within libraries 32

2.1 Cataloguing with an expert system 36

2.2 Assistance with classification 39

2.3 Reference and referral systems 45

2.4 Expert help with on-line searching 51

2.5 Examples of systems within LIS 59

2.5.1 Answerman 60

2.5.2 Cansearch 62

2.5.3 Expert 62

2.5.4 HEADS 63

2.5.5 Indexes 65

2.5.6 IGP 67

2.5.7 MARKUP 68

2.5.8 MITINET/MARC 69

2.5.9 PLEXUS 71

2.5.10 Pointer 73

2.5.11 Userlink/OASIS 74

3. An expert selector of online public access catalogues 77

3.1 Inexpert selectors 77

3.1.1 SALBIN 81

3.2 Expert selectors 84

3.2.1 System objectives 85

4. Expert system shells 89

4.1 What is a shell? 89

4.2 Shells vs. languages 93

4.3 Superexpert 97

4.4 Crystal 101

5. Knowledge acquisition 104

5.1 Problems with knowledge acquisition 104

5.2 Conspectus 119

5.3 University Funding Council report 134

5.4 University Grants Committee report 137

5.5 Questionnaires 140

5.5.1 Design and response 140

5.5.2 Dundee Institute of Technology survey 143

6. System construction 149

6.1 Explanation of Conspectus 149

6.2 Procedure for OPAC selection 152

6.3 Rules used in the expert system 157

6.4 Communications 161

7. Conclusions 165

7.1 Evaluation of the project 165

7.2 Future developments 168

Bibliography 174

Appendix A Questionnaire 194

Appendix B Dundee Institute of Technology survey 201

Appendix C Examples - Screen shots of the demonstration expert system 208

Appendix D How to run the demonstration expert system 214

1. Introduction

Over the last few years, within the area of library and information services (LIS), a number of expert systems have been created which can guide a user to relevant sources of information. These include such systems as Answerman (1) which was designed to guide users to a specific reference source in the field of agriculture and Indexes (2) which does the same thing on the subject of engineering. Such systems have, however, usually been limited to the sources available within a single library. This same type of technology which facilitates access to information sources within a library or information system can also be used to offer access to sources outside the library.

This project investigates the feasibility of constructing an expert system which would improve access to materials held in UK academic libraries and which are listed in those Online Public Access Catalogues (OPACs) that are accessible through the Joint Academic Network (JANET). The proposed expert system will select the OPAC most suited to a search for information on a particular subject and will activate communications software which will link the user with the OPAC.

Looked at later in this introduction are first Janet and the OPACs that are available on this network. This should demonstrate the possibilities that exist for an expert system to be used in this area by showing the difficulties that lie in the way of the OPACs accessible on JANET. Also examined are expert systems, this is to demonstrate what expert systems are and how they operate. This shows why expert systems differ from other systems and, by looking at the component parts of a typical expert system, how they differ.

The next section examines some of the uses that have already been made of expert systems in LIS. This shows that expert systems are already productively used within LIS and that the use of such systems is increasing. It demonstrates that expert systems are considered in some cases a better way of solving a particular problem than more conventional systems. This is illustrated by examples of expert systems that have been developed in this field some of which are used in-house and some of which are used and available commercially. These examples have been selected, as far as is possible to give a view of expert systems within a range of area e.g. cataloguing, online information retrieval, and so on.

The following section looks at a selector of OPACs. It looks at an inexpert selector and how it operates. It then goes on to look at what an expert selector would be able to do and what it is about an expert system that would allow it to be considered an expert, rather than an inexpert, selector.

The next section looks at ways of creating an expert system. It provides arguments for the selection of the expert system shell Crystal for the construction of this system.

Following this is an examination of knowledge acquisition. This shows problems in this area and methods that are used to overcome these problems. It details the sources of knowledge that were considered for this project, which have been selected and how the information in these sources is being used.

The penultimate section deals with the construction of a small demonstration system to provide an example of what an expert system selector of OPACs could do and what it might look like. It shows how knowledge can be turned into effective rules to govern the operation of such a system and how such a system can be linked into JANET via communications software.

The conclusion looks at the project as a whole. It considers if expert systems are a valid solution to the problems of access to JANET OPACs and what such a system might lead to in the future.

1.1 JANET OPACs

JANET is a network linking computers in organisations concerned with research that are funded by the Department of Education and Science (DES). It thus involves all British universities (except Buckingham) and many polytechnics and colleges. In many cases where the organisation concerned has an OPAC it is possible to gain access to this catalogue via JANET. At the present time it is possible to access the OPACs of 45 organisations in this way, as well as such other services as the Scottish Union List of Current Serials (at Edinburgh University), ARTel, Blaise-line, etc (3).

When this network began functioning in 1984 it was as a wide-area network serving to interconnect the local-area networks (LANs) at the sites served. Its prime task was to offer its users the following services (4) -

terminal access
file transfer
electronic mail
job transfer

It is important to realise that JANET is a wide-area network which operates by connecting the local-area networks LANs at the participating sites. There are thus no terminals and no host computers attached directly to the wide-area network as such. Access is therefore from a terminal or host computer on the LAN at one site to a terminal or host computer attached to the LAN at another site.

Janet does offer other services such as access to other network services such as British Telecom's Packet Switch Stream (PSS) and operates as a gateway to other networks (not all of which it is fully compatible with, so some services are not available on these networks), for example the European Academic and Research Network (EARN). It also provides access to some host-related services, such as access to various commercial databases.

Janet thus links together a large range of various services and makes them available to a wide range of potential users. It has been estimated (4,103) that Janet occupies an area of about 1000 km by 600 km, with about 130 sites connected, about 1,000 host computers, about 10,000 terminals, and has about 250,000 potential users. These include end users (as distinct from those users who are responsible for operating the campus networks to which end users connect, and who 'use' JANET in order to achieve interconnection of the campus networks) in various categories including undergraduates, research students, senior academic staff, technical staff, and administrators.

It manages to do this despite the fact that it has to link to a wide range of different types of host computer using an even larger range of operating systems which are connected on different LAN systems.

It is therefore true to say that the fact that the OPACs of many of the connected institutions are available to end users of the network cannot be seen as any deliberate act of resource sharing (5). They are only available in such a fashion because of the fact that they were each part of the LAN at the site concerned.

However the large number of OPACs available do not present in any way a coherent or consistent source of bibliographic information. This lack of consistency becomes obvious when the network has been used to access several different OPACs. These catalogues vary enormously with several different systems being in use. The most common of these is GEAC but others used include Libertas, Dynix, BLCMP, Dobis/Libis, CLSI, URICA and also some in-house produced systems. The capabilities of these OPACs available differ considerably. They vary from being relatively crude to being highly sophisticated devices. For example the OPAC at Aberystwyth, in the University of Wales, provides access to material only by the author's name or by the title of the work concerned whereas the OPAC at Oxford University is much more sophisticated and allows searching by 12 different means of access (including keyword, ISBN, subject heading, etc.) and the use of the Boolean AND operator This type of variation will affect how easily and with what degree of sophistication it is possible to search for information in the various OPACs.

However just as they vary in how it is possible to retrieve information they also vary as to just what information is available to be retrieved. Many OPACs do not cover the full range of the library's stock. The OPAC at the Brynmor-Jones library at the University of Hull covers 100% of the library's stock but in some cases either due to a deliberate decision, a lack of resources or simply because the OPAC has only recently been installed some OPACs cover only a portion of the library's available material. The OPAC at the University of Swansea, for example, only covers 60% of the library's stock (6).

Not only may only a certain proportion of the total number of items be held on the OPAC but, naturally enough, the content of the collections of the various academic institutions differs according to which subject areas are given priority within that particular institution. If a department within a university has a particular reputation for excellence in a certain subject area, then the collection of materials within the institution's library is liable to reflect this. For example The Robert Gordon Institute of Technology (RGIT) being situated in Aberdeen is highly committed to certain aspects of the offshore oil industry. Its department of Mechanical and Offshore Engineering has a high reputation for excellence in its field and its offshore survival course is internationally recognised. RGIT places a high priority on these areas and the library's stock reflects the support it gives to these departments and to the expertise within the institute. RGIT has recently acquired a URICA OPAC on which most of its stock is now listed. If this OPAC were ever to become available to external users via JANET it would be a suitable location in which to search for sources of information on the North Sea oil industry and its related topics. Other institutions will also have their own specialisations, and reputations for excellence in various fields of expertise, and the library of that institution will be likely to contain material which reflects and backs up this specialist knowledge.

However, to a user searching for materials on a particular topic, assuming little or no prior knowledge or experience with searching these OPACs, all these 45 possible sources would seem equally valid. Even to a user with experience while those OPACs that are particularly user friendly and easy and effective in use will become apparent with extensive practice to come to a judgment on the subject strengths of a particular OPACs' collection would be extremely problematical.

An expert system which contains data on the strengths of the collections of libraries, whose contents are bibliographically organised in OPACs, may be able to assist the user. Such a system would be capable of eliciting a user's subject interest and by relating this to data on the strengths of library collections, guide the user to the OPAC most likely to satisfy that information need. Such a system would be expert in as much as it would be concerned with assisting the decision making process of the user. The user would provide a need, the system would provide a recommendation as to a source of information capable of satisfying that need.

1.2 Expert Systems

If an expert system is to be considered as a means of solving some of the problems in accessing JANET OPACs that have been looked at in the previous section, it is important to understand what an expert system is and in what ways it differs from other systems. This will provide indications as to why an expert system may be better suited to providing solutions to some of these problems than other more 'conventional' systems.

1.2.1 What is an expert system

One thing that is agreed is that there is no agreed definition of what an expert system is (7). Some have suggested that definitions are usually tailored so the author can use the term to describe his own work in the hope of financial benefit (8). Despite the lack of a definition of what an expert system is, or what it does, they can generally be seen to be similar in their objectives. According to Ercegovac (9) the key goal of most expert systems is to model the human expertise of a well defined field and offer intelligent support (not replacement) to a practitioner that equals or surpasses even the best practitioner in the field.

Some other definitions which may be useful in order to illustrate what people mean, or may mean, when they talk about expert systems are -

An expert system is an artificial intelligence computer program which uses knowledge and inference to address problems of the sort which human experts would normally solve in a particular domain of expertise (10).

Expert systems are intelligent computer applications that use data, a knowledge base and a control mechanism to solve problems that are difficult enough to require significant human expertise for their solution (11).

... a system which performs the job of an expert or consultant in some area, and supports in making decisions in unstructured problem situations (12).

However probably the best way to understand what an expert system actually is, is to look at what functions the programs carry out and how they do them.

1.2.2 Components of an expert system

Two main components can be identified within an expert system. These are the -

a) Knowledge Base

b) Inference Engine

although some do identify a third component as being a user interface, usually a natural language interface (13,14).

The Knowledge Base is a collection of information or knowledge which can be drawn on or manipulated and the Inference Engine is the reasoning component. This determines what can be done with the information in the knowledge base.

1.2.3 <u>The Knowledge Base</u>

The Knowledge Base of an expert system contains the knowledge available to the human expert, or experts, whose skills are being reproduced within the system. But what exactly is this knowledge and how can it be contained within the memory of a computer in such a way as to make it usable?.

There are various kinds of knowledge that humans possess and in order to behave in a manner fitting the various definitions of an expert system that have been given, an expert system must also contain these kinds of knowledge. Barr and Feigenbaum (15) consider that the following are some of the types of knowledge required for an intelligent expert system -

- Objects: facts about objects that we have experience of: such as people have arms, houses have roofs and so on. There must be a way therefore to represent objects as a series of facts about them. Various categories of objects can also be dealt with; bungalows are houses, houses are buildings, etc.

- Events: actions and happenings in the world; for example, the Sun will rise tomorrow. As well as storing data on the nature of the event some consideration has to be placed on the causation of events and the relative time lines of a sequence of events.

- Performance: this is knowledge about how to do something. Skill related knowledge such as how to play the violin or how to fix a burst pipe.

- Meta-knowledge: this is knowledge about our own knowledge. Knowing the extent and origin of our information, about its reliability and the relative importance of specific items of data. For example if a robot guided by artificial intelligence were planning a journey, its knowledge that it can read the road signs along its route to find out where it is, is a form of meta-knowledge.

Also highly important in considering the ways in which knowledge can be represented is what the eventual use of the knowledge is going to be. That is not so much the purpose of the knowledge but rather the way in which it is used to achieve this purpose. This can be broken down into three basic stages in the usage of knowledge (15) -

- Acquiring more knowledge
- Retrieving facts relevant to the question
- Reasoning about these facts in order to solve the problem

Acquisition of new knowledge may not be strictly necessary in some applications. Once the knowledge of the human expert has been entered into the system it could just remain there as a definitive repository of that expert's knowledge. In many subject domains however such as computing, medicine, genetics and so on, the rate of development and change within the subject domain is such that any static knowledge base, while it could still remain useful, would be quickly outdated.

The potential knowledge available and necessary for accurately dealing with all possible demands on the system will tend to outstrip that contained in the system. There should therefore be some mechanism by which new knowledge can be added to the existing base. This new knowledge must also be integrated into the structure of that already existing in order to facilitate the proper interaction between the old and the new items of information.

The retrieving and the determining of what knowledge is relevant to a particular problem is highly important especially where the knowledge base is a large one. Two techniques that have been developed to deal with this area of difficulty are known as linking, and lumping. If it is known that one particular piece of knowledge is inevitably going to involve another in the solving of a particular type of problem then an explicit Link can be created between the two; but if several pieces of data are typically useful together then they can be Lumped, or grouped, together into a larger data structure.

Reasoning involves the system in working out something it has not been told, from various things that it has been told. In problems of any complexity the ability of a system to do this becomes of increasing importance. There are various ways in which the knowledge can be structured within the knowledge base in order to allow this reasoning process to take place. Some of these structures are better suited than others to particular types of knowledge but in many cases a combination of forms is better than a single representation type. The most common representation formalisms are (16) -

- Predicate Calculus
- Semantic networks
- Production Systems
- Frames
- Scripts

Of these probably the most common is the production system. These generally depend on rules which use an IF-THEN form to express their knowledge and make decisions. For example-

IF it is raining outside
THEN take an umbrella with you

This rule type has two parts. The IF part is the condition or situation and the THEN represents the action or outcome of the condition. If the conditions are satisfied the action can trigger off another section of the Knowledge Base into action until the final outcome is reached. This is because many rules may be linked together to form inference chains where the conclusion of one rule becomes the condition of the next.

Of the other methods, predicate calculus is used to represent factual information about specific objects or individuals. It can express true and false prepositions and establish a relationship between objects and propositions.

Frames describe an object or class of objects. For example, one may describe an apple and another the class of objects fruit. They do this by storing values/facts about attributes of the item(s) being described in slots within the frame.

Scripts are a structure of frame like devices that can be used to represent a series of events and the relationship between them. This can provide some understanding of time and of cause-and-effect relationships.

Semantic Networks (17) are in the form of a net-like structure made up of nodes and links. Nodes are the connecting points where one, or more, link leads to. The nodes represent facts, concepts, situations, and so on. The links, or arcs as they are sometimes known, demonstrate the various relationships between nodes. A common type of link is the IS-A link. For example -

```
Clyde-is a-Robin-is a-Bird-has a-Beak
                :
                :
                -owns a-Nest(a)-is a-Nest
```

One of the major features of the semantic network is that items lower in the net can inherit properties from items higher up. For example the facts that Clyde has a beak and that robins own nests can be inferred from the above network despite the fact that neither of these facts was actually explicitly stated in the network. Such networks can become highly complex and large, capable of reflecting many different types of relationship.

A semantic network, as above, is typical of the declarative scheme approach. Declarative schemes separate facts from procedures, this means that there is a pool of data, organised in some way, and a separate set of rules which operate upon it in order to solve a problem. The procedural view however, such as production (IF-THEN) systems, see knowledge as a series of rules which are activated according to the stimuli of some data; that is, the existence of a specific piece of data (triggered off initially by the users' input) causes rules to be activated which in turn causes additional rules to be activated until an objective is reached.

An expert system is designed to contain the accumulated knowledge of a human expert, or experts. That knowledge however must somehow be transferred from the human's brain to the computer's "brain". The process used for this transfer depends upon where and how this knowledge actually exists. A human expert may well have written a large number of books or papers and his knowledge is thus documented and presumably already structured in some way. This structure may of course not be of any particular use in the structuring of the knowledge base but at least it does make the extraction of the important aspects of the knowledge somewhat easier.

However the possibility also exists that this expert's knowledge has not been documented well, or even at all. In this case the knowledge can be said to be "personal" knowledge which is embodied only in the mind of the expert. But even in the case of an expert whose work is well documented much of the expert's skill will undoubtedly lie in the form of approximate "rules of thumb" which are seldom, or never, recorded. Some way must therefore be found of extracting it prior to its codification, in the form of rules or examples, and entry into the knowledge base.

The task of collecting this knowledge generally lies with a knowledge engineer who, once he has collected it, must organise it into rules. The collection period is usually characterised by a period of highly intensive discussion between the knowledge engineer and one, or more, subject experts. This often includes the analysis of a sample of test cases. This discussion need not be in any way formal or particularly structured. Indeed it may often be better if it is not, thus allowing the expert to express his knowledge in a natural way and allowing him to determine the structure of the knowledge as the interview proceeds.

Brooks and Vickery identify some particular types of problem that may arise during this stage (16) -

- Actually identifying and then locating the experts in the first place
- Once they have been found, getting them to co-operate in the project
- Problems of articulacy. The expert may be able to effectively perform tasks and solve problems in his subject area but he may be unable to say how or why he does something, his knowledge being "subconscious", he may only be able to partially describe what it is that he

does and may end up describing it in a superficial manner, in the process missing out underlying strategies.

- Where a problem is "fuzzy", or uncertain estimates of probability by an individual are liable to uncertainty or bias

- The experts may disagree amongst themselves and be unable to reach a firm consensus of opinion on the matter.

There are however ways around problems of this nature. One of the most commonly used is to construct a prototype system and then let the expert interact with it. For example this was done in developing an expert system for monitoring oil production platforms (18). The aim is to cut down the number of unnecessary shutdowns of platforms and emergency situations that arise through the cognitive overload of the human operators. That is, where there are too many warning lights and indicators for a human operator to make sense of, leading to confusion and mistakes. A simple system was constructed using only twelve variable factors, as opposed to the hundreds that would be required in a real situation, and tested using experienced human operators. This type of testing can show up mistakes in the prototype's rules and may encourage the expert to come up with new items of relevant knowledge already in a rule format.

Despite this, rule collection can be a long and tedious business and attempts have been made to automate the process (19). Many systems now use automatic induction. That is they can deduce their own set of diagnostic rules from a database of examples, or case studies. This method apparently removes the need for the expert to directly communicate his knowledge and rules to the machine. It does however require that a database of documented examples be available and this may not often be the case (16). A domain expert could however supply the example and some have suggested that this is the most natural way of building up the knowledge base. In some systems however building a knowledge base by example can be almost identical to building it by the pre-creation of rules.

1.2.4 **The Inference Engine**

The Inference Engine is the reasoning structure of an expert system. This is the mechanism that really drives an expert system. It examines the existing facts and rules in the knowledge base and makes connection between them in order to retrieve the implicit knowledge therein. It also decides on the order in which the various inferences are made.

Two types of inference system architecture are in common use, they are the -

a) Forward chaining
b) Backward chaining

of rules (20).

The forward chaining of inference works in a data driven, event driven, or bottom up direction. Starting from the available information as it comes in and trying to draw conclusions that are appropriate to the goals (15). The inference engine would therefore start by using the attributes of the questions asked of it. This therefore means that it starts from a manageable number of initial conditions, that is, the question attributes. It then tries to determine the solution from among a large number of possible solutions. In other words it starts with known facts and makes inferences until it reaches a goal.

Backward chaining is goal driven, expectation driven, or top-down thinking. This means that it starts with a number of potential solutions and is trying to work backwards to reach the initial conditions, that is the attributes of the question, among a large number of possible initial conditions. It works backwards by assuming that a particular solution is correct and then testing to see whether this assumption is consistent with the initial attributes gathered.

Some systems can use both backward and forward chaining but most use only one. The more commonly used is backward chaining which seems to be more acceptable to users than a "random" gathering of facts followed by forward deduction from them (21).

Since the inference engine can be kept entirely separate from the knowledge base it need not be necessarily restricted to the one subject area originally covered by the expert system. It essentially carries out the same functions no matter what the domain of the system is. An inference engine is sometimes referred to as a shell. In a way, this is exactly what it is, a framework or skeleton of a system which is capable of handling many tasks depending on the information in the particular knowledge base.

Expert system shells are now common pieces of commercial software and are available in a variety of guises from simple induction systems usable by anyone to highly complex and sophisticated pieces of software. This expansion to shells rather than the use of dedicated programming languages, such as PROLOG and LISP, has brought expert system technology within the grasp of a far wider spectrum of users than was the case even a few years ago, library and information services included.

Indeed it could be argued that LIS as dealers in information to be used in problem solving situations should be, and will be, some of the main users of expert systems. Indeed such systems may be essential if LIS are to go beyond providing information and also provide 'knowledge'. The distinction between 'information' and 'knowledge' is a difficult one to make but, in this context, by knowledge I mean material that is not simply raw data, it is a more sophisticated for of information provision than is usually carried out at present.

In the longer term this may lead to systems that will in themselves provide the knowledge required to the end user. In the shorter term it will tend to mean systems that aid the LIS provider in producing the knowledge required, by easing the technical processes that are involved. This will mean, as is considered in the following section, expert systems that aid in classification, cataloguing, and information retrieval processes.

2. Uses of Expert Systems within libraries

For too long, information specialists - especially documentalists - have obviously thought that the compilation of data in retrievable form was a significant contribution to problem solving. With great astonishment they observed the disappointment of users confronted with piles of such 'valuable information', produced as a result of an extremely skillful information retrieval process. Indeed many documentalists thought this was the answer to the user's question: as if handing a huge encyclopaedia to a farmer, trying for the first time to design an irrigation system in the desert, would be the solution to his problem. It would probably not even contribute to the solution. Thus says Canisius (22), commenting on some of the reasons why, according to him, the use of expert systems will become essential in the library and information profession. In his view, in the future all kinds of knowledge will be represented in knowledge based systems and information intermediaries will have to transfer, into the future systems, the skills acquired in their daily work over the years, with all sorts of databases and questions.

Forbes Gibb (23) identifies the main areas of interest in the use of expert systems within libraries as -

- Intermediaries for database interrogation
- Classification aids
- Cataloguing aids
- Reference work aids

He is concerned however that developing systems in these, large and complex problem domains will prove to be demanding in terms of time and resources. It is possible that useful expert systems could be developed in areas other than these or to perform smaller, more specific tasks in certain areas. Some of what are possible application areas within the general area of librarianship are -

- Rare skills archiving
- Records management
- Library guiding
- Careers advice
- Course selection
- Distribution lists
- Expertise directories
- Machine translation
- Systems selection

- Decision support
- Education and training
- Electronic reference works
- Community/citizens advice

While applications in many of these areas would no doubt prove useful additions to a library service they are in the main not concerned with what could be seen as the "core" domain of librarianship. It is within this "core", primarily as embodied in the previously listed four areas (page 22), that will probably prove to be most significant to librarianship and information retrieval as a whole.

The expert system for aiding access to JANET OPACs could be seen in the context of these "core" areas as being a reference work aid providing access to bibliographic information but also providing information as to the location of specific items of interest on a subject basis. In a more sophisticated form it could also be looked at as an intermediary for database interrogation. It will not operate as such at present but it would be possible for the expert system to operate as a front-end to the OPACs being accessed. This could be done by providing a common command language, for the searching process, making the actual OPAC being searched "invisible" to the end user.

Looking at how systems operate within LIS and at examples of how specific systems have solved problems can provide clues as to how an expert system for accessing OPACs would have to operate, what it could be expected to do, and also what it cannot at present be expected to do.

2.1 Cataloguing with an expert system

Cataloguing is a rule based activity. The accurate application of rules, as for example those given in AACR2, would seem to guarantee the production of a definitive catalogue record every time. Given that expert systems depend on a knowledge base consisting of rules, this immediately suggests itself as an area in which such systems could successfully be used.

Unfortunately, it is not that simple. AACR2 may be comprehensible to some humans, provided that they are experts in cataloguing, but at least some of the rules would be very difficult for a computer to act upon. Davis (24) points this out with the example of rule 22.2C2 which instructs the cataloguer to prefer the predominant name if an author uses different pseudonyms or the real name and one or more pseudonyms. Unless the system's knowledge base, as well as containing a rule or rules to carry out this task, also contained the information as to author's names, pseudonyms and which were the most known or most used, it could be unable to implement that rule and it would still be up to the cataloguer to make the decision. In creating the catalogue record it is not just the rule that is important, or the information contained in the work to be catalogued but also background knowledge, that is, information that is not explicit in either the rule to be applied or in the work being catalogued. The rule will specify how such information is to be used, but to be applied the information must first be available. Such problems are probably not insurmountable and if expert systems are to be applied in cataloguing there are two main ways in which they could be used (20) -

- An advisory system where the intellectual effort is split between the expert system and the human cataloguer. The system copes with deciding which rules should be applied and in which order, to any particular type of document, but the final choice of entry terms and so on is left to the human with advice from the system.
- An expert system with full catalogue capability hooked up to some form of electronic publishing system. This means that as text is generated on-line on the electronic publishing system it can be passed directly through to an expert system cataloguing process which would mean that there was no intellectual input from any human intermediary in the cataloguing process.

Most systems developed up to now have been of the first type, including systems such as MITINET/MARC, HEADS and MARKUP (which will be looked at later as examples of systems in use), advisers and trainers only. Other systems developed have been experimental systems looking at how cataloguing rules, particularly AACR2, could be used as a knowledge base for an expert system (25).

2.2 Assistance with classification

The introduction to the Universal Decimal Classification scheme (UDC) states that classification is something of an art and that proficiency in it depends to a large extent on a general background knowledge of the subject being classified. This would tend to imply that the process would be difficult to automate. However rules, and rule based operations, lie at the heart of expert systems and any operation which in itself is rule based, as classification certainly is, has the potential to make use of expert systems technology. Not only is "the application of existing classification schemes possible, but the creation of newer and improved schemes is also possible. Most commonly encountered hierarchical schemes fail to accurately reflect the conceptual network which represents the world of knowledge (26) and it may be possible with an expert system to construct new classification orders (27).

Existing classification schemes as illustrated by the printed classification schedules of UDC, BC, DC, and so on, theoretically contain, implicitly or explicitly, all the necessary rules for classification (28). All that would seem to be needed would be to transfer those rules into the knowledge base of an expert system in order to create an expert classifier. However this may not be as simple as it would seem, given that the same type of problem may arise as it does in cataloguing, where the rules contained in AACR2 may be difficult to frame in the way in which an expert system would require them, or even if framed correctly may not act 'intelligently' enough to justify the work involved in its creation.

Clarke and Cronin (20) give a simple example of some rules for an expert classification system. The system having gathered the title of the work, subject matter and key terms involved could sift and rank these in order of importance followed by a rationalisation exercise to select an optimum classification following rules such as -

- IF medicine THEN 61
- IF 61 AND physiology THEN 612
- IF 612 AND stomach THEN 612.32

The application of rules such as these of course depend heavily on citation order. This would reflect the ability of the system to sift and rank the various terms involved.

They envisage this as part of a system which has the ability to use natural language as a part of its information gathering process. A problem however remains in how to enter, and more importantly what to enter as data to allow the system to make its judgment as to what is the optimum classification for any given item.

Enser (26) in his experiments used two hundred and fifty books from which their title, table of contents (chapter and part headings) and the complete back-of-book index were entered into his system. The transcription of the complete index avoided some of the problems associated with phrase selection and thus removed some of the intellectual effort from the human experts, making it a simple clerical task, but did however mean that the resulting data structures were rather large. In an expert system this would of course necessitate a large knowledge base to translate the input terms into the terms required by the particular classification scheme required. Working on the data, Enser found that the automatic classification system produced results that were more effective than the manual classification had been in the collections from which the items had been drawn. Despite the fact that he was not working with an expert system, as such, his very success in this area using more conventional programming techniques implies uses for an expert systems in this area. An ES could be used as a natural language intermediary, an explanation mechanism, a knowledge base containing sufficient data to aid in translation of input data to the classification's required terms and as an automatic gatherer of electronically published material which can then be accurately and automatically classified.

Burton (28) suggests some of the areas of classification that he considers could be benefited by using an expert systems approach, these include -

- Inconsistent classification. This can be a problem where there is a large information service with a number of classifiers who will require some form of overall co-ordination if it is to be ensured that there is uniformity in the classifications produced. This can be particularly true if there are specific local classification procedures being used.

- Verification of classification. Even with elaborate procedures for checking numerical accuracy and so on, inconsistencies can arise, such as successive copies of the same work, or successive editions, being given different classifications. An expert system based classification would provide a consistent and uniform classification no matter which classifier dealt with the item.

- Efficiency of service. It should be possible to increase the accuracy, for some of the reasons above, and also the speed at which classification can take place when using an expert system.

- As an elucidation tool. The application of an expert system to these problems, even in some cases where the use of such a system does not prove viable, should (since it is a new approach) at least throw up ideas and provide a new view on the problem enabling further development in the area to take place by more conventional means.

- The removal of the more routine and clerical aspects of the procedure would free staff time for more demanding activities. It is ultimately possible that such systems will finally succeed in de-skilling the entire procedure by removing absolutely all need for a human intellectual input but, at the present time, such a system would seem unlikely to be developed quickly.

In attempting to apply an expert system to these areas Burton (28) has created an ES based on DC 19's Arts class. Two versions were created using different expert system shells (for more detail on expert system shells see section four). The systems were found to work reasonably adequately, considering their experimental nature, but one of the major problems encountered was that of vocabulary control. The system only used the vocabulary of the DC 19 schedules, clearly a limiting factor since all that has effectively been eliminated is the turning of the pages of the schedules. Further research was carried out on this problem looking (as was Enser) at the potential of using the work's title, contents page and index as sources from which index terms could be derived. According to later work by Sharif however (11) Expert system shells are not suitable for creating large scale ES in classification, they lack the ability to accept natural language input, the knowledge domain is large which causes problems with vocabulary control and the required knowledge base would be too large for an existing shell to handle.

Classification is still an area in which the application of ES has not been extensively investigated. There are however clearly some areas in which they can be applied and as research proceeds it may well be that expert systems will lead us to an entirely new view of classification.

2.3 **Reference and referral systems**

The systems looked at in sections 2.1 and 2.2 have so far been concerned with what can be termed as "primary" information. They possess knowledge on a subject area and answer questions, give advice or carry out an expert type task. An expert system designed to cope with reference work can contain knowledge of a more "secondary" type. That is, knowledge as to where other information can be found. This means it must possess the same type of expertise as a human reference librarian. In designing a system to carry out these functions Brooks (29) identifies what an intelligent interface for an information system would have to be like -

- Something which stands or mediates between the user and the information sources.
- An interface which is not human but instead some computer hardware and software.
- Is able to accomplish the same functions that a good human intermediary would in the same situation.

Performing the same task as human experts is intended to be one of the major features of an expert system, so on the last criteria alone, an expert system could certainly have a part to play in this role. Before we can be sure of this however it is clear that the tasks a human in this situation would perform must be identified. Parrot (21) considers that there are five main tasks that a reference librarian must perform -

- Draw information from the user to determine the salient characteristics of the question; i.e. the reference interview.

- Be able to remember the scope and coverage of a large number of reference tools.

- Match the attributes of the question with those of the reference tools and be able to arrive at a list, possibly ranked, of tools that may be of use.

- Explain how to use the reference tools and other library tools such as the catalogue. This can be in general terms for classes of tools or in more detail for specific important tools.

- Bridge any gaps between the vocabulary of the user and that of the reference tools. Acting almost as a human thesaurus capable of framing a question in standard or preferred vocabulary.

Parrot's way of determining which of the above tasks could usefully be carried out by an expert system is to consider what it is about certain of the tasks performed by the reference librarian that -

a) Qualify reference librarians as experts.
b) Take up so much time of the librarian that the time and expense in constructing an expert system to perform them can be justified.
c) Are sufficiently low level that an expert system can be constructed in a reasonable amount of time.

Many researchers in the area seem to think that a large number of tasks in this area fulfil sufficiently these criteria for there to be a great amount of activity in the development of automated systems for reader guidance expert systems appearing to have natural application to the reference process. Thompson (30) claims three main benefits for expert systems as a part of this process -

- Memory
- Objectivity
- Consistency

The benefits in memory lie in a consistent and deep knowledge of the database (contained in a knowledge base). Ignoring for the moment physical storage and processing constraints, the system has the potential of knowing everything there is to know about the structure of a knowledge base and of never forgetting.

The consistency benefit appears when we look at the user interface. In all dealings with people it is necessary to adjust to their idiosyncrasies and biases. We make the adjustment, then have to alter the interaction to accommodate swings in mood or because they are unavailable at the moment or have gone on to greener pastures. This type of problem is obviously evident in the reference interview. An expert reference system will always treat the user consistently. We never have to be concerned whether an automated system likes us or not. We do not have to be concerned as to how intelligent, or otherwise, it might think our question is. The user will come to know what to expect of it and will never be disappointed in its reaction.

Such a system could also have other advantages. The imposing qualities of a computer system's keyboard and screen may intimidate most users into more accurately formulating their query in the first place. According to Krulee and Vrenios (31) computers are notoriously unfeeling and unforgiving; one unconsciously prepares for the severest punishment at the slightest mistake.

There are still many problem areas in the use of expert systems for reference work; natural language understanding, vocabulary control, and difficulties in subject domains that are not well defined, narrow and discrete. However the (at least partial) success of system in this area to date, and the continuing research, indicates that successful and publicly available systems of this type are not too far away. Examples of systems in this area including Plexus, Answerman, Indexes, Pointer are given later in section 2.5.

Budgetary restraints and a high demand for services at information desks have made it difficult to provide effective reference services where they are needed. Information technology, in the form of expert systems supplemented by computer-aided instruction, offers a viable means of maintaining a good service without overworking staff. It can also help to compensate for any cases where such staff may have gaps or deficiencies in their professional knowledge or expertise.

The system considered in this project is in fact an example of a system of this type. It determines the area of the users subject interest and refers him to a source which is likely to provide bibliographic information on that subject. In its full form it would be providing access to a vast range of material that is not normally easily or quickly available. By determining the OPAC most able to provide information on a subject it would either be able to provide a reference service that was not being offered, or to enhance a service that was offered.

2.4 **Expert help with online searching**

Many, if not most, libraries and information services now have access to remote databases. These include, among others, those provided by such hosts as Dialog, BRS, Data-star, Infoline, Blaise, and so on. Available directories list several hundred different databases held by various hosts (32,33) and this is by no means an exhaustive listing with over 2,500 databases world-wide providing bibliographical, statistical, numerical data on all subjects. As the demands of modern society for information are tending to increase, at an ever greater pace, so libraries are also having to increase the scope of their work and in some situations access to these on-line resources is essential to the provision of an adequate information service.

These on-line systems however have developed separately over a number of years. There has been little or no standardisation between them. This has led us to a situation where each host has its own command language, logon procedures, database file structure and searching abilities. This means that to have full access to a range of databases held by different vendors a user must be familiar with a range of communications/access procedures and command languages, familiarity with one will not normally allow you search on another. For example (34) searching Chemical Abstracts via CAS on-line, Dialog and System Development Corporation's ORBIT requires three different set of procedure varying from logon protocols to alternative methods of term truncation and search logic.

Due to these factors many users of such systems are not in fact direct users. They work through some kind of intermediary, in most cases a librarian or some other expert human who is more familiar with the systems being used. This intermediary can not only carry out such searches generally faster, thus cutting the cost of on-line time and telecommunication charges (due to this familiarity) but have several other functions. One of the most important of these is a conversion process. This process involves changing the user's expressed information need into a firm search strategy which can be input into the database system. It may not only be a case of converting a user's natural language enquiry into a structured command language. As in any reference work the question itself may not have been put in such a way as to extract the information actually required. As Paice and Ramirez (35) have said most user's statements of an information need have about them a certain element of ill-definition and will therefore require topic elucidation in order to arrive at a satisfactory picture of the users true requirement.

If an expert system is to be used instead of a human intermediary it must thus be able to duplicate not only this, but all the major functions of the skilled information handler it is replacing. Williams (36), who designed a system called Userlink for this very purpose, considers the essential tasks which the human at present carries out and the knowledge he must possess to be -

- Choosing a database which is likely to contain the required information.
- Choosing retrieval terms which accurately describe the search topic.
- Knowledge of the retrieval language commands.
- Knowledge of the communications systems and protocols.
- The ability to interact with the host computer.
- The ability to react to error messages given by the host.
- An ability to modify the search strategy in the light of the results obtained from the

database being searched.

An expert system may not be required in order to carry out all of these functions as some of them are perfectly capable of being handled by non expert computerised intermediaries. Automatic communication to a chosen database including dialling and logon does not require 'expertise' as such, merely a database of numbers and codes which are automatically transmitted when a key is pressed. Such a system however could easily be integrated into an expert system and if the system is in any case going to be able to automatically choose a database to go in to for a particular search then it may as well be able to communicate and logon automatically as well.

Looking at which criteria the experts themselves consider to be most important in a database was done by Morris and Tseng as a part of the construction of a prototype expert system for database selection. Two tables showing the results for various criteria are (37) -

Table 1: Criteria considered important when selecting a host

Host Characteristic	Average rating on scale 1-5
Own familiarity with search language	4.6
Flexibility of search language	4.0
Variety of search fields	3.7
Speed of operation	3.4
Cost	3.1
Quality of documentation	2.8
Online cost accounting	1.8

Table 2: Criteria considered important when selecting a database

Database characteristic	Average rating on scale 1-5
Currency of material	4.8
Types of source covered	4.7
Quality of abstracts/text	4.5
Familiarity with database	4.2
Search fields available	3.7
Cost	3.5

It can thus be seen that in this area there is a range of factors which must be taken into account in just the selection process alone.

Most of the factors mentioned in the preceding tables could be taken into account by an expert system and be related to the JANET OPACs.

Looking at the factors in Table 1 first. Familiarity with search language, flexibility of search language and variety of search fields can generally be related to the type of OPAC in use at the particular site, e.g. GEAC, URICA, or other type. If desired the expert system could enquire if there were any particular type of OPAC that the user was happier using in relation to these factors. If so the system could give greater importance to this and select an OPAC of the desired type, although it might therefore have to give less prominence to other factors. It might in this case end up not recommending the best OPAC for a subject area overall, but recommend the best OPAC of a particular type. This type of operation would also cover factors in table 2, the familiarity with the database and the search fields available.

Other factors in table 2 which can be covered are the currency of material and the type of source covered. As looked at in section 5.2 of this project, Conspectus is the major source of information around which the demonstration expert system has been created. Conspectus provides ratings for libraries as to their collection strengths in subject areas. Two ratings are provided in Conspectus. The first covers the strength and comprehensiveness of the current collection, i.e. types of source covered, and the second provides a rating as to the current collecting intensity in a subject area, i.e. likely currency of material. For more detail an how Conspectus does this section 5.2 should be consulted.

It can thus be seen that even the demonstration system does take into account some of the important selection factors in making its choice and could be expanded to consider more of the factors identified by Morris and Tseng.

Several systems have thus far been able to automate parts of the procedures carried out by a human intermediary. Given the amount of interest in this area it surely will not be long until the entire procedure is automated, perhaps leading to a vastly greater amount of direct end user searching. There are already several systems in this area e.g. Cansearch, IGP, Userlink/OASIS which are considered in section 2.5 as examples of expert systems within the field of LIS.

2.5 Examples of systems within LIS

This section includes brief details on a range of expert systems that have been developed (either as commercial products or as research projects of various types) with the LIS field. These systems have been selected in order to provide a range of examples illustrating the types of things that can be done with expert system technology. They were selected as well known examples of systems within the "core" areas previously looked at. HEADS, MARKUP and MITINET/MARC are systems aiding with cataloguing. Answerman, Indexes, PLEXUS and Pointer are reference and referral systems. Cansearch, Expert, IGP and Userlink/OASIS are systems to aid with online searching. I have not provided an example of a functioning system in the area of classification. As pointed out in section 2.2 there are some particular difficulties associated with producing systems in this area with the result that examples of systems that are any more than either partial, or very experimental are rare. For examples of systems and theoretical systems in this area Sharif (11), Clarke and Cronin (20), Enser (26), and Burton (28) should be consulted.

The consideration of expert systems that have already been created in the LIS area is important since these systems can be used to establish "benchmarks" against which the performance of the system being considered can be judged.

2.5.1 **Answerman**

Answerman (1) was developed by the American National Agricultural Library (NAL). Their aim was to create a microcomputer based system that will guide an enquirer to a reference book likely to contain the answers to his question in the field of agriculture. Answerman was developed using an expert system shell called 1st-CLASS. Indeed the NAL were not concerned with, and did not seriously consider, building their own system from scratch and were primarily interested in a system shell that could easily be used by people with no previous expert system experience; 1st-CLASS proved to be well suited to their needs. It runs on an IBM PC (and compatible machines) and is an example based program that allows unsophisticated users to create a system using a framework much like a spreadsheet (for more detail on how an example based system shell operates, see section 4.1). The shell allows a single file to have only thirty one different responses, not nearly enough for any serious purpose, but this can be solved by creating a number of files linked together in a hierarchy guiding the user through in a general to specific manner. The shell also allows feedback on its use to be gathered since any system response can be linked to an external program. A member of the NAL staff wrote a BASIC program for gathering the responses of users to the system which easily integrated into the main program. After this had been added, the system was further expanded to include the page numbers of books where specific answers were to be found. It was found however that this required expanding the system to a vast degree, so much so that this was not practical for the whole system.

Other ways of increasing the capabilities of the system were however found. Certain expert system responses were linked automatically through a piece of communications software (Crosstalk) to certain files in either the Dialog or BRS on-line database systems. The system did not function as a translating intermediary since the host's own command language still had to be used. It was concluded that it would be possible to link the system to a commercial 'gateway' package that featured a single simple menu driven approach to searching in various systems. Finally Answerman was linked to a bibliographic file held on CD/ROM. The mass storage capacity of CD/ROM meant that the scope of the system proved to be expandable to an enormous extent.

2.5.2 CANSEARCH

CANSEARCH (38,39,40) was designed for use by doctors as an intermediary when searching for cancer therapy literature on MEDLINE. The system presents the user with various menus of Medical Subject Heading (MeSH) terms which he can select by using a touch sensitive screen. When the user has input sufficient terms to define his subject area the CANSEARCH system will generate a search strategy which is in the required form for the MEDLINE database.

2.5.3 EXPERT

EXPERT (41) is an experimental intermediary system described as an automated expert assistant for information retrieval. It uses production rules written in a LISP like notation and is controlled by a goal-directed forward chaining mechanism. Four areas of search strategy formulation are handled by EXPERT. EXPERT takes charge of the search strategy formulation and reformulation through a series of directed questions to the user who responds on the basis of a menu or a fill-in-the-form mode. The four areas handled by EXPERT are -

- Computer assisted database selection through a statistical transformation from menu-selected (standard) topic areas to a list of databases ranked by presumed relevance to the search topic
- Topic representation through a conceptual formulation of topic aspects and aspect search terms
- Automatic transformation of the conceptual formulation into a keyword/stem Boolean search strategy, followed by automatic execution of this strategy
- Computer directed relevance feedback type reformulation of the search strategy.

2.5.4 HEADS

HEADS (42) is an expert cataloguing advisory system developed at Teeside Polytechnic Library. One of the major problems for inexpert cataloguers is the varying requirements of different classes of document. Even when those requirements are understood there is then a vast array of rules which must be applied in order to ensure the consistency of style and standard that is necessary to make the catalogue as efficient an engine of information retrieval as possible. While it might be hoped that many rules are capable of being applied by an expert system there are many which could be difficult to implement in this way. For example, rule 21.15 in AACR2 (on the requirements for an added entry when the work is one of criticism) would be difficult to automate. The decision that any given work is of that nature must still be made by a human cataloguer. Where an expert advisor can come in though, is to inform the human user of this, and other rules that may, or must, be applied in certain circumstances. The expert system can therefore have a role in cataloguing in helping out in difficult cases and in training the inexperienced cataloguer. In some circumstances it can work not merely as an advisor but can determine the details of a specific example by working through the rule base which is applied to give the user an answer to a specific question. The system can, as well as enunciating those rules found in AACR2, be augmented by additional rules uncovered in developing the system, that is those rules which are applied in practice but are not explicitly written down. With these aims in mind two different versions of this HEADS system have been constructed using two different expert system shells, or inference engines, ESP/Advisor and SAGE. Neither of these uses the full set of cataloguing rules but both could be expanded to do so.

2.5.5 <u>INDEXES</u>

INDEXES (2) is an expert system designed to select between varying reference tools in engineering, depending upon user input. It was developed on the EXSYS expert system shell and uses production rules.

In the system the choices were limited to fifteen reference sources reflecting various levels of sophistication and containing three different types of engineering information -

- Factual/statistical (Handbooks)
- General overviews (encyclopaedias)
- Recent research (Periodical indexes)

When run, the initial screen describes the systems ability to determine the probability that different reference sources will adequately meet an information request. The system then runs through a series of interactive questions that seek to simply simulate a reference interview. These interactive question were determined by splitting the reference interview into three distinct sections -

- Identification of subject matter
- Identification of the user's background knowledge of the subject area
- Determination of the type of information needed

The system identifies the subject matter of the enquiry by asking for a choice of menu options, e.g. chemical engineering, civil engineering, and so on.

The user's background knowledge is assessed by asking if the user's knowledge is general or specific within the subject area.

Lastly the type of information is determined by asking which of the three types of knowledge the system covers is sought. In addition the user is asked if they want the Library of Congress Subject Headings (LCSH) covering the relevant areas to be displayed.

After these question have been asked, but prior to presenting the list of choices, the system details how the list of choices should be interpreted. The user discovers that the list has been hierarchically organised to reflect statistical probabilities. These probabilities indicate the likelihood that the type of information requested will in fact be found in the specified source.

The listing screen which identifies resources is ranked from highest to lowest with a score of 10 (high) to 1 (low) allocated to each item. Specific LCSH headings relevant to the chosen area are also shown at this point (or as nearly specific as LCSH provides).

According to the systems creator even a simple system such as this one can provide important insights into reference work and point the way to the future where expert systems will be a full participant in direct reference negotiations with users.

2.5.6 Intelligent Gateway Processor (IGP)

The IGP (34) at the University of California does not aid in the formulation of a search strategy but functions as an intermediary, standardising the command language that has to be used when searching on-line databases. It is attempting to make the use of a database 'transparent' or invisible to the user of the system. It is designed to provide agricultural information from a range of databases and networks and to eliminate the clerical effort in getting there. Telephone numbers, baud rates, identification procedures, passwords, differences in terminal keyboards, and display screens are all resolved by the gateway. Once the gateway has connected the user to a particular database it can be interacted with, either in its own command language or in a gateway overlaid common command language. Simultaneous connections to multiple hosts can be used to compare results, share files and verify findings, a feature only available through this type of system.

2.5.7 **MARKUP**

MARKUP (43) was devised at the Building Research Station (BRS). It is used for entering catalogue records into the BRIX database which is the main retrieval tool at the BRS Library. The main objective in designing it was in order to provide a training program for non-professional library staff on-the-job. This was done with the view that it would both save professional staff time and give clerical workers a deeper appreciation of the job in the context of library procedures as a whole. Since it was designed merely as an advisor it does not have a large enough knowledge base to allow non-professional staff to take over the cataloguing task entirely but new staff have found it useful as a training device and it could even prove useful to new professional staff as a way of acquainting them with local cataloguing practices and procedures.

2.5.8 MITINET/MARC

MITINET/MARC (44) is an expert system which can be used to aid in the creation of catalogue records in the MARC format. The system became available in 1986 and is microcomputer based, being usable on the Apple II and IBM PC microcomputers. The end product of the system is a complete MARC record on a floppy or hard disc in the standard US MARC communications format.

The system is designed to let non-expert, or even totally inexperienced users, create the same standard of record that it would have taken an expert to produce previously. The system is set up to cope with three different levels of use -

- New, or very inexperienced, users
- Experienced users
- Expert users

New users of the system are given a large number of prompts and instructions in the form of full English phrases and examples of how every part of the catalogue entry should appear. Users simply have to choose options and enter the data, the system can then do the rest.

Experience of the system in use has demonstrated that non professional and inexperienced users can be successfully trained to operate the system and can begin to produce records in less than two hours if they are guided through it by some one familiar with it. Even without someone helping them they can learn enough from the manual in about four to six hours to start operating the system.

Experienced users can choose to abbreviate the prompts given and can also choose not to have the examples appear on the screen.

Experts familiar with MARC standards have additional choices. These include being able to display and edit the indicators for each tag rather than just accept those automatically generated by the system and to have a good deal of control over the record being generated.

The system enables a MARC record to be created in five to eight minutes and the same standard of record will be created from the same data regardless of the experience of the user operating the system.

2.5.9 PLEXUS

PLEXUS (45,46,47,48,49,50) is an expert system that was developed by the Central Information Service of the University of London. Funding was provided by the British Library Research and Development Department (BLRDD). Its aim is to accept statements about specific problems faced by users and reply by providing information on reference texts that may provide the answer to their query, or persons or institutions who are knowledgeable in the relevant subject areas. The system does not aim to provide specific documentary references or to provide substantive answers to the question itself. It is intended to be used in a public library setting and the subject area of the initial system is limited to gardening.

The objectives behind the project were -

- To develop a generalised computer-based subject referral system that could be used to provide referral facilities to various types of information source
- To use the system to implement a particular referral system that could be of value to

libraries

- To demonstrate to libraries the nature and potentialities of expert systems

With these objectives in mind PLEXUS was designed with five main functions -

- Constructing a model of the user. This means that the system can become familiar with a user, or group of users, and is able to adapt its response accordingly
- Obtain a description of the user's problem
- Formulation, and where required reformulation, of a search strategy based on the information gathered in the first two stages
- Presenting the user with the results of the search in an informative manner and then obtaining the user's evaluation of how likely these item are to be helpful
- An explanation facility covering the system's capabilities, activities and the search outcome

Each of these functions has been produced as a separate system module with its own knowledge base as well as a general knowledge base containing information required by the whole of the system.

The system is capable of handling a vocabulary of several thousand words in a natural language manner. For example, a user could input a statement such as "How do I grow Peppers?". The system is capable of handling this and, after prompting the user for synonyms and any other information it thinks it requires, translating it into a search strategy in Boolean form for its own use e.g. Grow AND Greenhouse AND (Capiscum OR Vegetables). It will then present the user with one, or more sources that it considers to be suitable for the solution to the problem.

PLEXUS was written from scratch using PASCAL and is not shell based. As can be seen by the amount of literature generated by the project it has been one of the major research efforts directly into the use of expert systems in libraries. It is entirely possible that the relative success of this project will encourage and enable further systems to be developed in LIS.

2.5.10 POINTER

POINTER (51,52) was one of the first expert systems designed to answer the kind of questions a librarian would actually encounter at a reference desk. POINTER directs users to one or more of the fifty most useful reference tools in the US Government's Documents Department at SUNY-Buffalo. These sources were the result of an analysis of reference questions collected over an eighteen month period. The design is a simple menu-driven system that lists useful sources in the reference collection in response to user selections from fifteen categories of need, e.g. information contained in treaties. The system conducts the reference interview by means of menus. The user makes the appropriate choices from the menus and the system responds with the names of specific reference sources, call numbers, and brief instructions. POINTER was designed to run on a PC and though originally written in LISP was later converted to BASIC.

2.5.11 <u>Userlink/OASIS</u>

OASIS (53) is an on-line intermediary expert system that was developed from an earlier system, Userlink (54).

Userlink was a system based on production rules which aimed to help a database searcher by assisting him in choosing the database, the retrieval terms, automating the communications process, and some help in terms of the different command languages used. In a BLRDD research project this system was developed into OASIS. The OASIS unskilled user system is designed to simulate the operational knowledge and expertise of the search intermediary. In order to carry this out the main functions of the OASIS system are -

- A tutorial session to introduce the user to the concepts of on-line searching
- An interview and query formulation session to identify the search term which describe the user's search
- An automatic search formulation which creates the correct searching format for the system which is to be accessed
- A fully automatic search on-line which reports to the user the progress of each section of the on-line dialogue
- A reporting session which provides the user with the results of his search
- A review and revision session which allows the user to improve the search and carry out a further on-line search with a revised profile

The process is similar to the normal search process with the intermediary but the assessment of results and subsequent search modification are done off-line. The end user is likely to spend some time on these tasks and, if on-line, the costs incurred would be too large.

In the assessment stage of the project it was considered that users generally judged the system to be successful.

These are merely a selection of systems and research projects concerned with the development of expert system use in the LIS field. The number of such systems and their relative success hints at a fruitful future for expert systems in libraries.

3. An expert selector of Online Public Access Catalogues

3.1 Inexpert selectors

The most obvious use of accessing library catalogues over JANET is for users of those libraries to check whether the library has the work they want (5). In this case the user would be accessing one specific catalogue in order to search for an indication as to the presence of one particular known item. Assuming the user is able to use JANET to logon to that catalogue and is able to search that catalogue using the correct retrieval method this presents no great problems. The existence of that item in that particular library's catalogue will be confirmed or denied. It would be possible, for example, for a JANET user at RGIT to check (from their own office) whether Aberdeen University Library has a copy of a particular work they are interested in, to discover if the work is currently in stock and if it is therefore worth their while making a trip there in order to consult or borrow it.

If that user was making, or was willing to make, a trip further afield, then, still from their own office, they could conduct similar searches in the libraries of other academic institutions e.g. Glasgow or Edinburgh University libraries. In order to do this however the user would have to be aware of a greater range of NRS (Name Registration Scheme) addresses and other possible responses required to enter the selected catalogue(s).

An NRS address is a unique numerical identifier used to gain access to a particular remote database or service. For example the NRS address for the University of Lancaster library OPAC is 0000 1040 2000. Other possible procedures may be required also. For example, again at the University of Lancaster library, after the NRS address has been entered and the user has been connected to the University's computer they will be required to enter another identifier to gain access to the actual OPAC. In this case they user is presented with a computer prompt asking them to 'Logon'. At this prompt they must hit the Return key and when a second prompt 'Username' appears they must enter LANPAC. Most OPACs do not require such procedures to enter them but some are even more cumbersome in their protocols. The University of Kent, for example, requires the user to respond correctly to four separate prompts prior to them gaining access to the library's OPAC.

To an infrequent user having to remember, or being able to have access to documentation to such detail is a disincentive to use these facilities. Equally not all those people who would wish to be users of this service have ready access to computing facilities or to JANET. Lack of ready access to a terminal deters many, who then do not have the opportunity to practice and gain confidence (55). Availability of documentation, practice and training, such as that carried out by the JANET User Group for Libraries (JUGL) will help, but probably of greater value to the casual and infrequent user would be the removal of the necessity to bother with such details. After all what this type of user is after is quick and easy access to a particular OPAC for a particular purpose. If it is made difficult or time consuming for them it could be that they would abandon their use of the facility altogether.

The library at Dundee Institute of Technology, among others, has developed an uncomplicated program which simplifies the matter for the user. When accessing the OPAC on site in the Institute itself users are also presented with the opportunity to access other OPACs. A simple menu allows them to select from a range of eight other OPACs to search in. This system allows access to the OPACs of all the

Scottish university libraries and also to the OPAC of Cambridge University library. By simply selecting an option from this menu they will automatically be connected with that OPAC. They have no need to be aware of any of the procedures required to get them there.

Loughborough University Library have developed a system for aiding with access to up to forty two of the OPACs available through JANET (56). They considered that a simplified access method was needed in order to benefit fully from the opportunities provided by access to remote catalogues. A communications package (KERMIT) and a menu package (MENU) on an IBM PC were used here. Linked together, the communications package uses small program-like devices known as scripts to automatically link with the OPAC after the user has selected it through the menu system generated using the software package. By merely pressing cursor control, return and escape keys, any user can gain access to any one of the OPACs covered. The principles involved could readily be applied by any library with access to JANET at extremely low cost. The KERMIT package for example is public-domain software, MENU is of negligible cost and there are a number of freely available public-domain alternatives any one of which could have been suitable. At

Loughborough they consider that this has made a significant improvement in service to the general user.

These are examples of small in-house systems that require minimal resources and are easily set up by librarians with only very limited experience in computer use.

As can be seen in the following section, more ambitious projects have also been carried out leading to the creation of more complex and co-ordinated systems.

3.1.1 SALBIN

The Scottish Academic Libraries Bibliographic Information Network (SALBIN) was a project on library co-operation in Scotland that included, among other things, the proposed production of a user interface to allow general user access to SALBIN facilities (57,58). This interface has now been produced and is currently available to the library community for use on IBM PC's connected to JANET.

The interface has been designed as a resilient and user friendly menu driven system. It is in essence similar to the simpler systems described in the previous section but offers a greater number of features and a greater degree of sophistication. This was possible partly since the system was not constructed from already existing pieces of software but was designed exclusively for this purpose from the start. It is designed to be made available in public areas for self-service by users. It can cope with unexpected inputs and responses by distant hosts and has built into it defaults and time limitations to cope with indecisive users and with those who do not finish their session in the approved manner. It attempts to protect itself against accidental or deliberate tampering by disabling the route to DOS.

Upon entering the system the user can gain access to some general help and information screens, or from the main menu, select an OPAC in which to conduct a search. When the system has connected, the user then has the choice of printing the search session, downloading it onto floppy disk or searching without either option. A help screen is available giving further information on this.

The user is then free within the selected OPAC to conduct the desired search. This search is of course subject to all the advantages and disadvantages of the selected OPAC. There is a facility in the system for site-specific help screens to be inserted and academic libraries are being encouraged to create an appropriate help screen for remote users. Such a screen could be made available on the JANET bulletin board and downloaded into a site's SALBIN system.

When the user logs off they will be returned to the system but if a user finds that they are unable to leave the OPAC, for whatever reason, then the system incorporates F10 as an emergency escape route to return to SALBIN.

The system as available is not particularly expensive (£75) and does offer some major advantages for the inexperienced user. While many of the same things can be done with a simple menu-communications system created from standard software it would be difficult to make it as resilient and friendly as SALBIN.

3.2 Expert selectors

What then is the difference between an expert selector and an inexpert selector of OPACs?.

An inexpert selector, as described in section 3.1, is a device that simplifies the communication and logon procedures for the range of OPACs it covers. It may provide help screens and the like, but essentially the system leaves the choice of which OPAC to go into entirely up to the human user, who may or may not have expert knowledge of these OPACs.

An expert selector would also be able to automate the communications and logon procedures but in addition would either advise the user, or select for the user, as to the OPAC which the system, according to its knowledge, considered to be the most suitable for that particular search.

This kind of system would prove to be most useful where the user was concerned with discovering items on a subject basis. As pointed out, in section 1.1, the collections of academic libraries vary. There are libraries in which there are more items of more relevance to a particular topic than others. If a user is looking for the existence of a particular item in a particular collection then such a system will be of no value, but to a user searching for items (on a subject area) of which he has no previous knowledge then such a system could prove of use by increasing the probability that the user will find something of value. There would of course be no guarantee of success when using such a system, the system would recommend the OPAC of the collection it considered to be most fruitful for that type of search but this would surely be preferable to a selection by a user that was at random or effectively no more than a guess.

In order to be considered as an expert selector however a system would be distinguished by having certain system objectives as looked at in the following section.

3.2.1 **System objectives**

The basic objective of an expert OPAC selector would be to be able to recommend to the user the OPAC most suited to searching for information on the topic selected by the user.

In order to fulfil this basic objective the system would have a series of tasks and objectives which it would have to carry out. Some of these would be general objectives of the system as a whole and others would be specific tasks that the system would be required to carry out at certain points during its functioning. The general system objectives, which are often common to any piece of software, include -

- ease of use. At each stage in the procedure it should be clear to the user what is expected of them. The user should not be expected to make any long or complex entries into the system. It is supposed to make the process easier, not to add complexity.
- helpfulness. At any stage in the process where the user has to make a choice there should be help available to make it clear what effect this choice will have.
- robustness. The system should cope with the user hitting the wrong key. If the user makes an inappropriate entry the system should not crash, freeze or throw him out. It should offer them the opportunity to re-enter in a correct form, ideally advising that a mistake had been made and what form that mistake took.

The more specific system objectives are -

- to simulate a dialogue with the user in order to determine the subject area of interest to them.
- to determine any other parameters that

would affect the outcome of the search. These would be conditions that the nature of the OPACs themselves would put upon the search. For example, not all the OPACs allow searching by class number. If the user wished to search using this form there would be no point in recommending an OPAC that did not allow it.

- to select an OPAC for the search. This would be a ranked list of OPACs where there was no single outstanding option.

- to carry out automatic communication and logon to the selected OPAC if the user wishes it, or return to query the expert system.

- allow the user to search the OPAC and download their search onto floppy, or hard, disk.

- after the user has logged off from the searched OPAC allow him to end the session, search another OPAC (where there was a ranked list of options), or return to query the expert system again.

The system to carry out these objectives would require an expert system component and a communications component. The expert system would deal with the selection of the OPAC and the activation of the communications software. The expert system component for this system is an expert system shell called Crystal. More details of expert system shells and of Crystal, and its alternatives, are given in sections 4.1 to 4.4.

4. Expert system shells

4.1 What is a shell?

Expert system shells, sometimes referred to as knowledge engineering languages (10), are specialised programming environments tailored for expert system development.

Originally each expert system had to be individually written. They had to be custom built in a computer language suitable for the task, common languages being PROLOG and LISP which are dedicated artificial intelligence (AI) languages. Expert system shells provide an outline mechanism for applying a set of rules and a framework in to which purchasers can put their own knowledge base. In theory all users have to do, is enter their particular set of data into the shell, and it will undertake the rule formulation for them (7). They are in effect expert systems with the domain-specific knowledge removed, leaving only the inference engine and support facilities (11).

A typical expert system shell is a package of two components (59) i.e. a compiler to translate the rule base (expressed by a language particular to that system into an internal representation) and a runtime system, or user interface, to apply the knowledge base when queried by a user. These facilities, of course, vary from system to system but among the more sophisticated packages, features in the user interface often include -

- expansion and explanation on the meaning of a question. This would have had to be specifically placed into the knowledge base of the system by its developer.
- ability for the user to alter previous answers to a question asked by the system and to have the effect of this new response propagated through the previously derived conclusions.
- a provision for the user to volunteer answers to a question before the system asks them. This would normally be done by having the user type in an option at the keyboard which would bring up a menu of questions for selection. It is likely that in the future the development of enhanced natural language interfaces will enable this type of dialogue, i.e. where things can be user initiated, to take place more naturally and easily.
- allowing the user to inspect the working database, enabling them to see the answers given previously in the session and the derived values for other goals.
- an explanation facility. This allows the

user to see the path that his enquiry has taken through the knowledge base. During a session it allows the user to know which rule it is that has been triggered to ask him that particular question at that time. In other words it enables the user to find out "why am I being asked this question?." Some shells, after an enquiry has been completed, are able to provide to the user a detailed picture of the path the entire query took through the structure of the knowledge base. This facility, though it can in some cases be useful, has its limitations, one of the major ones being that this explanation is usually provided in a totally mechanical way. This, although it will accurately reflect the structure of the knowledge base, need not reflect the true conceptual structure of the problem domain in that the expert system or consultational program may itself not reflect this structure in an accurate manner (60).

Most shells fall somewhere between applications programs and programming languages and like languages may be used to develop applications which meet a specific developers needs. Shells may be used efficiently by those with limited programming experience but there is, however, a trade off between ease of use and flexibility. Capabilities are sacrificed with most shells, which make them somewhat more than simple applications programs but less than full blown programming environments. An expert system shell is to expert system applications what a programmable database manager such as dBase is to database management applications (61). This does not however mean that they are inferior. Each problem area must be looked at to see if a shell would be a suitable tool. Some of the reasons why are looked at in the following section when the use of programming languages to create expert systems is looked at.

4.2 Shells vs. languages

The alternative in expert system development to using one of the shells available, is to use a programming language to develop the system from scratch. When using a language it is possible to use either a conventional language or a specific AI language.

Of the conventional, or problem orientated languages the use of languages such as C and PASCAL is common. PASCAL was in fact the language used to design the expert system PLEXUS (see section 2.5.9) and C with its speed and flexibility has become increasingly popular as an expert system building tool.

The AI field itself has two major languages in use. These are LISP and PROLOG. Of these LISP is by far the older and better developed (62) and is the more common of the two in the US, although in Europe PROLOG is more popular (10).

LISP is a flexible and extendible system, which means that if you don't like some of the facilities is provides, you simply write your own and incorporate them into your personal version of LISP without a loss of understandability or effectiveness (63). These facilities arise from the very structure of the language itself. It operates in small, easily understood chunks and the spaghetti-junction-like constructs familiar to users of languages such as BASIC or ALGOL are extremely hard to replicate in LISP. It is however, not particularly good at handling numbers and therefore not good for mathematical problems. However, it is (and this is what makes it most suitable for expert systems applications) a symbol manipulation language.

The task of writing a PROLOG program is not like specifying an algorithm in the same way as in a conventional programming language (64). Instead the PROLOG programmer asks more what formal relationships and objects occur in his problem, and what relationships are true about the desired solution. It can thus be viewed as a descriptive language as well as a prescriptive one. The approach is rather to describe known facts and relationships about a problem, than to prescribe the sequence of steps taken by the computer to solve a problem. Using this language the actual way the computer carries out the computation is specified partly by the logical declarative semantics of PROLOG, partly by what new facts can be inferred, and only partly by the explicit control information supplied by the system designer. PROLOG programming consists of -

- declaring some facts about objects and their relationships
- defining some rules about objects and their relationships
- asking questions about objects and their relationships

A program consists of a set of clauses, where each clause is either a fact about the given information or a rule about how the solution may relate to or be inferred from the given facts. Thus PROLOG may be seen as a first step towards the goal of logic programming. Logic programming is characterised by the view that knowledge, rather than data, is the essential raw material to be processed (65).

In the end the decision on whether a shell or a programming language is to be used for the development of any particular expert system will be influenced by exactly the same kind of factors that influence the decision in any application vs. language choice. The applications program that at first glance seems more expensive, will almost always end up being cheaper when the cost of development, debugging and maintenance are considered. For a non-programmer the use of applications software is usually considerably easier than learning not only a programming language but how to program effectively. Provided a shell is capable of supporting the features required by the expert system being considered it is probably a better choice for many applications. In fact given the sophisticated user interfaces available on some shells it may in some ways prove to be a superior option to the use of a language. In the end the decision will rest on the nature of the problem and the expertise of the system designer. In the case of the proposed expert selector of OPACs a shell is the preferred choice due to the relatively uncomplicated nature of the problem and the ease of use offered by a shell. The choice now of course is which shell?.

The first shell considered was Superexpert since copies of this shell were already available within the department. The consideration and rejection of this shell is looked at in the following section. Since Superexpert had been rejected another shell had to be found and after a period of considering reviews and recommendations the Crystal3 expert system shell was looked at. Reasons for the choice of this shell are given in section 4.4.

4.3 SUPEREXPERT

SUPEREXPERT is the shell that was first considered for use in this project. SUPEREXPERT was produced by Intelligent Terminals Ltd. (66) but is however no longer produced or supported by them so no technical assistance is available for this product. Produced in 1986 it is now in computing terms somewhat out of date anyway and incorporates many of the common limitations found in such shells, particularly the older ones.

The system is one that uses rule induction. That is it uses an induction algorithm, in this case the "Analogue Concept Learning System (ACLS) (67), to generate a set of decision rules from an example base. An example base is a set of information which serves as a specification of the decision-making process, it provides the information from which the system creates rules containing the knowledge in the examples. This method of rule production has its limitations since a decision rule induced by the system may not operate in the way intended by the designer. This is particularly true when considering the order of the questions the user will be asked. The system will induce rules to find an answer asking the fewest possible questions. Unless they are split into categories of questions very carefully by the designer they could be asked in an order that would be without meaning to the user of the system.

Response time would be slow since, in order to deal with a large problem, the enquiry would have to be broken down into a number of sub-problems chained together. This is due to the fact that each problem allows only thirty-two possible solutions and due to the nature of the system some breaking down of the problem would be essential to create a meaningful structure for the user. These sub-problems would have to load one at a time in sequence as the problem progressed and the user would be unable to return immediately to the problem level preceding the current one without repeating the entire query from the beginning.

The system lacks the explanation facility which would demonstrate to the user how any particular decision was reached, or at least which rules were activated in the reaching of that solution.

The user interface is highly limited. Dialogue with the user must be conducted via a series of menus. The scope for customisation of these menus is limited and the fact that only seven, one-line, options can successfully be displayed on screen at one time without scrolling being necessary is irritating to users, particularly when they discover that having scrolled down through a menu of choices they are unable to scroll upward back through those choices. The only input that will be accepted by the system is the number corresponding to a menu choice, the user is unable to enter anything else, except for quitting, which by this time they may consider to be the most attractive option of all!.

The shell was designed as a stand alone package and whilst able to exchange problem information with databases and spreadsheets (dBase and Lotus 123) lacks the facility to link with communications software and so allow it to directly link the user to any OPAC it may have been able to select.

In an attempt to overcome these difficulties, particularly the latter, the use of the programming language Turbo-C was considered in this work, in order to provide a link between SUPEREXPERTs screen only output and other files. It seemed unlikely that this could be done however, without enlisting considerable outside expertise in the programming field. Also considered was the use of a menu driven procedure to link the shell and communications package and perhaps provide a more attractive front-end to the software. This could have been done using either simple DOS programming or more attractively one of the many available menu packages (e.g. Automenu). Neither of these options provided any clear solution to the limitations of this product for this purpose and so it was abandoned as unsuitable for further consideration.

Also briefly considered were the shells ExpertEase and XI Plus. These seemed to suffer some of the same limitations as Superexpert and although both are more sophisticated products neither of them seemed to offer any obvious solutions to the problems raised by Superexpert. Consideration of these shells was quickly abandoned however when the shell Crystal became available since it became apparent that Crystal was a far superior product which offered a range of features and ease of use far outstripping any shell so far considered. Crystal is looked at in more detail in the following section.

4.4 **CRYSTAL**

Since the use of SUPEREXPERT had proved untenable, another expert system shell was required. CRYSTAL, produced by Intelligent Environments, was the one selected (after consideration through reviews (69)) due to the range of features it offered, its ease of use, the user support available, and its price compared to its alternatives.

Crystal (70) itself is menu operated, providing the user with a series of screens on which the commands, knowledge base, and query system can be easily built up. These screens (each screen in effect corresponding to a rule) can easily be linked together, in a variety of ways, to create an ordered structure of the designers choice.

This allows the designer, in a way that was not possible with Superexpert, to specify in what order the rules will operate, how the user will be queried, and in what order these questions will be asked. This can be done using the range of commands available. The "shape" of the knowledge base can be determined in the way the commands are linked together. Four commands are available to the designer, IF, AND, OR, and NOT. To these command can be linked questions and conditions. A user can be questioned and the command then tested to see which selection the user made and thus which commands should be followed to continue the query down a particular line of enquiry. How this is done is looked at in more detail in section 6.1 which examines how the rules used for this system were created and implemented using Crystal.

The user interface is far more sophisticated than with Superexpert. Questions can be asked of a user in many ways. Menus can be used, yes/no choices offered, sliding scales, and direct user input can be accepted. Text and help screens can be made available to the user and with Crystals "paint" function these screens can be made up as the designer requires. Since these screens can be customised they can be made as complex and as colourful as the designer wishes. It is even possible to import graphics created using other programs in order to provide the user with detailed graphic images.

This shell like Superexpert is also able to exchange date with other programs, such as dBase and Lotus, but more importantly has the ability to be used to interact with other programs via DOS. Using the command DOS($,$) Crystal can be caused to run another program. For example the command DOS("xtalk","\") would cause the communications program Crosstalk to be run from the root directory. This will allow a system to be created which will be able to select an OPAC for use and then automatically link the user to that OPAC via communications software.

Given these advantages Crystal was selected as the shell for use in this project. The next stage of the project was to acquire the knowledge that could be used to build a system with Crystal. This is detailed in the following section.

5. Knowledge acquisition

Knowledge acquisition can be seen as a crucial stage in the development of any expert system. An expert system can ultimately only be as good as the knowledge contained within it. If that knowledge is faulty then the expert system will never be able to correctly achieve its aims. It is therefore vital to look at the knowledge acquisition process and the problems that can be associated with that process, prior to looking at the knowledge that is required for this expert system.

5.1 Problems of knowledge acquisition

Knowledge acquisition, sometimes referred to as knowledge engineering or elicitation, involves eliciting, analysing and interpreting the knowledge which a human expert uses when solving a particular problem (71). This process can be seen as a critical first step in the construction of any expert system as it will end up influencing both the final performance and the cost-effectiveness of the system (72). This process is normally carried out by someone described as a knowledge engineer although there are attempts to automate this process. Some consider in fact that it is preferable to have a program rather than a person extract an expert's knowledge (62). It may be that the next major breakthrough in the construction of expert system development techniques may come in the automation of knowledge acquisition. The technique of building rules by induction or generalisation over examples of expert decision making is an example of automated knowledge acquisition of sorts (73). For example, in an imaginary system, to be able to classify animals on the basis of their colour and shape, the expert may type in -

- grey, big and elephant
- brown, medium and deer
- grey, small and tortoise

The induction algorithm within that system would then, from these examples, be able to generate rules such as -

If the animal's colour is

(a) brown, then it is a deer

(b) grey, then if the animal's size is

(i) big, then it is an elephant

(ii) small, then it is a tortoise

A system like this operates as SUPEREXPERT (see section 4.3) does with the precise form of the rules being determined by the system and the algorithm it uses to generate its rules.

Computer based automation of the full knowledge acquisition process, however, may be more difficult than simple rule induction and may, in the foreseeable future, be an unrealisable dream. In the view of some (74) the automatic construction of rules by algorithmic or heuristic methods is some way from routine practical applicability and the attempt to automate the process of knowledge base construction may be based on a somewhat premature rejection of current knowledge elicitation techniques.

The expert system's knowledge can be gained from a variety of sources. The classical source is, of course, a human who is particularly skilled and knowledgeable in a particular field. Other sources may include books, manuals, reports, databases, journals, case studies and in fact any collection of knowledge or information that contributes to the solving of the type of questions likely to be asked of the expert system being designed.

Attention in this area is mainly focused on the obtaining of knowledge from human experts however, and much work has been carried out in this area from a variety of perspectives -

psychological, sociological, cost-effectiveness and so on, in order to get around what seems to be the bottleneck of knowledge acquisition.

The early stages of knowledge acquisition are in the main about obtaining an overall structure. The purpose at this time being to gain an overview of the problem area which can be seen as being a large scale map of the subject domain (75). This means that initially the knowledge engineer must look at the broad issues and should not become submerged in excessively detailed information. It would be an extremely difficult procedure to separate out the primary issues from among a sea of detail and it is therefore vital to gain appreciation of the subject as a whole. As the process of knowledge acquisition progresses further and the overall structure has been clarified then the attention can be switched to a more detailed view of the problems. This will involve mapping out the main tasks, sub-tasks and fine detail involved in the subject.

In the next step, which is establishing a more detailed knowledge of the domain, more attention has to be paid to the types of knowledge which will be encountered. It is common for knowledge engineers to distinguish between two types of knowledge. These are declarative and procedural knowledge which can be seen as corresponding roughly to facts and rules respectively. Declarative knowledge may be viewed as assertive, for example a database of propositions. Procedural knowledge, on the other hand, is a knowledge of how things are done, that is rules, heuristics, algorithms and so on (76). It may be that individuals at different levels of expertise in a domain posses different type of knowledge (72). Novices tend to rely more on declarative forms of knowledge and exercise a significant degree of conscious control over their problem-solving processes. Experts however are more likely to use automated sequences of behaviour and procedure that are not consciously controlled. These experts may not actually be aware of much of the actual procedure used or the knowledge involved in carrying out these automated procedures. Experts may also find it harder to verbalise this automated procedural knowledge.

There are a variety of techniques available for use when attempting to elicit knowledge from human experts. Problems however still can arise. Two common causes for problems in these procedures are (75) -

- the task is being discussed with the wrong expert
- the expert's pronouncements do not fit with

what the knowledge engineer is hoping to hear

As ways to avoid having problems caused by these two factors it is suggested that -

- a suitable expert is one who solves the problem on a day-to-day basis and is motivated to help with the project

- knowledge engineers should believe that experts probably do solve problems in the way that they say they do. If this does not fit with the knowledge engineer's ideas then those ideas will have to be changed

Assuming it has been possible to avoid the encountering of these two problem areas and a suitable expert has been selected (or experts since in many cases there may be no one expert who can build the conceptual representation of knowledge required) (77), it is necessary to consider the possible techniques that may be used to gain the knowledge of that expert. The reasons why the methods of eliciting knowledge are often vague or unstated is because in many cases they are ad-hoc and unscientific. As the importance of this area has come to be recognised more however, the techniques have become better defined. Some of the main techniques include (73)-

- interviewing. This can be used to select an overview of the domain of expertise and an understanding of the expert's jargon. It is possible during an interview to ask more direct and probing questions of the expert in order to gain access to the store of procedural and declarative knowledge he possesses.

To this end a range of interviewing techniques have been developed which can aid the knowledge engineer in structuring the interview in order to best be able to acquire the knowledge required from the expert. These include (71) -

a) distinction of goals. This involves the expert being presented with a specific goal and being asked what evidence is needed to tell this goal from the alternatives.

b) reclassification. The expert is asked to work backwards from the goal he has identified by elaborating on the decisions which supported it.

c) dividing the domain. The expert starts with a set of facts and forward chains through successive sub-goals until the final goal is arrived at.

d) systematic symptom-to-fault links. A list of all possible faults and all possible symptoms are presented to the expert who is asked to say which faults will produced which symptoms.

e) intermediate reasoning steps. Similar to d) but asks for the evidence that links the steps.

f) twenty questions. A structured technique aiming to quickly and efficiently identify the major characteristics in a field from a collection of case studies. The expert is presented with a problem and has to obtain from the interviewer, who is only permitted to answer yes or no, the information he requires to solve the problem.

g) laddered grid. Produces a map of the domain by asking for examples to discover super and sub classes within a domain.

h) critical incident. In this technique the expert is asked to provide the most memorable examples of problems providing evidence of specific events rather than recounting routine cases.

i) forward scenario simulation. The expert is asked to describe how he would deal with a hypothetical problem.

Even experts however, may have trouble in accurately verbalising the way in which they actually solve problems and there may be difficulties in using knowledge gathered in this manner into the rules and control structures of an expert system and there are other ways of gaining such information.

- protocol analysis. This can be defined (78) as the procedure used by cognitive scientists and psychologists in order to understand the human problem solving and decision making process. The procedure consists of giving individuals real or simulated tasks to perform and asking them to "think aloud" as they work. The comments such individuals make are referred to as verbal protocols, and the techniques used to analyse these comments are termed, collectively, "protocol analysis". The main purpose of a protocol analysis is to identify structures and patterns rather than to simply look at the content and it is important to collect the verbalisations concurrently rather than retrospectively with the interpretation of these think aloud protocols reflecting the products of cognitive processes. Verbalisations provide information about the information attended to or needed by the subject at a given point and reflect the current contents of the subject's short term, or working, memory. Three levels of verbalisation can be identified (76)-

1) in which information would be produced in the form in which it is needed

2) in which the information would be needed but would normally be encoded in a non verbal

form which must be translated for verbalisations

3) where the subject is asked to report selectively or induced to attend to information that would not normally be needed

Once the protocols have been gathered however, what methods should be used to analyse them?. Several knowledge extraction techniques have been derived using protocol analysis including the three phase methodology (79), Delphi (80), and the Crawford slip method (81). The common denominator underlying all of these techniques is an emphasis on structuring and systematising the knowledge elicitation process. Despite its weaknesses, it has a serious incompleteness problem, in that certain types of thinking lead to little or no verbal expression. Protocol analysis has been used in eliciting knowledge in a majority of expert systems (78) although due to this limitation it may prove to be much more valuable in some knowledge domains than in others.

- observation. This could be regarded in many ways as a process of non-verbal protocol analysis. In this type of analysis it is not purely a verbalisation that makes up the protocol, but rather includes a series of actions and processes carried out by the expert during the course of his problem solving behaviour.

- multidimensional techniques. These types of techniques are those commonly used to form a map of a subject domain by eliciting structural criteria which the expert can use to organise concepts in a representation which may be awkward to verbalise. Three of the most used forms of this technique are -

1) card sort. A set of cards is prepared each of which bears the name of one concept. The expert is then asked to group them into piles according to any criteria he chooses. This procedure is repeated for all other criteria by which the expert perceives the concepts as varying.

2) multidimensional scaling (MDS). Each concept in a domain is compared to all others and some estimate of their similarity is arrived at. This is a complex technique and is only really of use where the concepts are liable to vary according to only a few criteria.

3) repertory grid. A key aspect of this method is that it is idiographic, i.e. it measures attributes of individual subjects (76). The grid is constructed as a matrix of constructs and elements. The elements (domain concepts) are used by the expert as a basis for generating the constructs, which is the dimension of difference between elements.

Subsequent statistical analysis of the grids by various methods can reveal clustering and factoring which can be used in rule generation. Repertory grids are a general purpose knowledge elicitation technique used in other fields, such as personality theory, and will not provide much information on procedural knowledge.

These are illustrations of the most common methods of knowledge acquisition for expert systems. The sources of knowledge for an expert system to automatically select an OPAC are somewhat different however. Most expert systems focus on acquiring knowledge directly from an expert or experts. For this system however direct use of knowledge possessed by humans is of less usefulness. This is due to several factors, firstly the difficulties involved in identifying who would be considered as an expert in this problem domain.

The criteria for determining who would be considered an expert in the selection of an OPAC would be:-

- someone who uses JANET regularly to access OPACs
- someone who uses a wide range of OPACs on a regular basis

- someone who searches by a range of different methods e.g. subject, ISBN etc.

- someone who searches over a large range of topics

In order to be conducting such searches on a regular basis the expert would be unlikely to be searching for information for purely personal use, they would be conducting a range of searches on different topics for different people. They would therefore almost certainly be librarians/information workers most likely in one of the UK academic institutions with access to JANET, conducting searches as an intermediary for staff and/or students of those institutions. It seems unlikely that experts of such a type actually exist, at least in any numbers at any rate. If a search for information were to be carried out electronically it may well be carried out on a more specialised database than an OPAC and if it were to be carried out on an OPAC via JANET it may well be handed over direct to the end-user of the information by some means such as SALBIN. OPACs available through the JANET network are too inconsistent a source of information to make it likely for a wide range of them to be used regularly by any one person over a range of subject areas.

Even if such a person, or persons, could be identified as an expert the second problem area raises itself. This problem is the vast range of the knowledge base. With forty plus OPACs available and hundreds of possible subjects to choose from, one expert would not be enough to define the knowledge needed for such a system. Even a small group of experts would be unlikely to have an opinion, or even an educated guess, as to the suitability of each individual OPAC in relation to each individual subject area, even on a broad classification basis.

Thirdly, the knowledge of such experts, whilst based on experience of use of these OPACs, would be unlikely to be based on definitive experience. By this I mean an expert's judgment may run something like this-

- "in the past I have searched OPAC A and B and C for information on subject X using subject keywords. OPAC A provided 12 references, OPAC B provided 4 references, and OPAC C provided only 1 reference. Therefore OPAC A is best."

Such a judgment may while still accurately reflecting the expert's experience not consider other factors that may play a part in the true situation. If, for example, the expert had searched the OPACs using classification numbers rather than subject keywords the result may have been different. The expert's knowledge while accurate is only partially true.

Also the time it would take to gather such knowledge would have been large. Whether using observation, interviewing or any of the other knowledge elicitation techniques, the nature of the task would mean that a large amount of knowledge would have to be gathered.

Expert knowledge however can be available in forms other than in the direct use of human experts. The sources of knowledge considered for use in this project are looked at in sections 5.2 to 5.5.

5.2 Conspectus

The Conspectus system, which was developed by the Association of Research Libraries' (ARL) the Office of Management Studies (OMS) and the Research Libraries Group (RLG) in 1983, is a method of measuring collection levels in libraries. Using coding mechanisms, it permits the assessment by subject of a library's existing collections and current collecting policy, on a scale of comprehensiveness ranging from zero to five. A separate set of indicators is used to indicate language coverage (82).

The subjects assessed are keywords based around the Library of Congress classification system. For example within the broad area of library and information science is the area of the history of books. Within that area is the subject the history of writing (among others). It is this specific subject, the history of writing, which is then coded. The broader subject areas, of which it is a subsidiary part, are not.

The fact that the subject area is then coded is the vital part. It is this information on a collections strength in a particular subject area which can be used to make judgments as to a collections suitability to provide information on a specific topic. The linkage of these factors (subject to code) that will allow this information to be used in the creation of an expert system which will then make a judgment on the suitability of an OPAC for searching on a subject area.

Conspectus has attracted much attention since its inception, some good and some bad. In its favour it has been said that the efforts of RLG and ARL have gone a long way towards realising the dream of many librarians for a description of existing research collections nation-wide in areas of interest to scholars all over the nation (83). The development of a national conspectus is a project of potential significance approaching that of the development of the National Union Catalogue of past generations. Through this effort, scholars and librarians everywhere will have a better concept of the location of major research collections. By understanding existing patterns of strengths, and by distributing responsibility on the basis of collaborative self-interest, the research libraries of the nation may develop even stronger research collections with less undesirable redundancy and unnecessary expenditure. Conspectus represents an insurance policy against future uncertain times. This new vehicle should provide the means for improved service, as well as enabling trade off and reallocation of resources in ways that will result in local economies. The project will give to librarians and scholars a bibliographic research tool on a grand scale that will make efforts at bibliographic access and rapid delivery more efficient and effective than ever before.

Another positive view of Conspectus is that (84) the completed conspectus constitutes a detailed overview or map of a collection that is different from, and complementary to, title based inventories, such as that of OCLC. The data permits detailed comparison of local collections with collections in other libraries where librarians have also completed the conspectus. Statistical reports and comparative studies are generated even more efficiently and effectively when the data is entered into a common database shared by a group of libraries. The detailed comparative reports which are then possible to produce are essential to librarians who have, or plan to have, co-ordinated, co-operative collection development arrangements with other libraries. Whether or not an automated database is established, the completed conspectus provides librarians with the data necessary to respond systematically to budget reductions or increases and changes in institutional roles and missions, curricula and also user patterns. It gives libraries a rational basis for initiating and reacting to change. At the local level perhaps the most important outcome can be measured in human terms. The skills of the librarians who participate in the conspectus process are upgraded. Librarians communicate more knowledgeably about their collections, are better prepared to make selection and management decisions, and generally take better control of the collection under their stewardship. The conspectus assessment process can encourage bonding between teaching staff in an academic institution and librarians, and reassures teaching staff that the collecting goals of the librarians correspond to their needs and the needs of their students. The completed conspectus is an excellent tool for communicating with college administrators and funding agencies. It can demonstrate powerful arguments for improved budgets and assures those to whom we report that we understand, control, and accurately predict our programs, in effect, that the optimum use is being made of the funds available.

These arguments in favour have converted many to the cause of conspectus, particularly in the United States. There, more than 80% of the 107 ARL member libraries are involved in, or are considering, the application of conspectus (85). The first country to adopt conspectus outside the USA was Canada (86). The Canadian Association of Research Libraries (CARL) and the National Library of Canada are working collaboratively with the ARL on Conspectus within the North American Collections Inventory Project (NCIP) (87). Canadian involvement was seen in many ways as the gateway to much broader international application of the conspectus. This was partly because the conspectus was translated into French for use in certain areas of Canada and many sections of the RLG conspectus which had a marked United States bias were revised. During 1986-87 the National Library of Canada developed an online system which was designed to store and provide access to coded collections' data and free text comments generated within the conspectus activity (88). It is intended that Canadian conspectus online will greatly facilitate co-operative collections development by Canadian libraries by making data on collection strengths readily available.

Other countries have now started to implement conspectus programs as well. The British Library has used Conspectus on its own collections (89,90) and other countries, some of them prompted by encouragement from LIBER (which now has its own conspectus group) (91,92), have conspectus programs. These countries include Scotland, Holland, France, the Scandinavian countries, Spain, Austria, Switzerland, Australia, and Malaysia. Conspectus has not so far however found much favour among library services in England, where librarians have tended to look at its disadvantages and criticisms of it rather than at its positive side.

Conspectus has perhaps as many critics as it does supporters. Many of these are based on the system of numerical indicators that it uses to rank the collection strengths and collection policies. These ratings run on a scale of 0-5 as illustrated (83) -

0 out of scope: the library does not collect in this area.

1 minimal level: a subject area in which few selections are made beyond very basic works. For foreign law collections, this includes statutes and codes.

2 basic information level: a collection of up to date general materials that serve to introduce and define a subject and to indicate the varieties of information available elsewhere. It may include dictionaries, encyclopaedias, selected editions of important

works, historical surveys, bibliographies, handbooks, a few major periodicals, in the minimum number that will serve the purpose. a basic information collection is not sufficiently intensive to support any courses or independent study in the subject area involved. For law collections, this includes selected monographs and loose leafs in (American) law and case reports and digests for foreign law.

3 instructional support level: a collection that is adequate to support undergraduate and most graduate instruction, or sustained independent study; that is, adequate to maintain knowledge of a subject required for limited or generalised purposes, of less than research intensity. It includes a wide range of basic monographs, complete collections of the work of more important writers, selections from the works of secondary writers, a selection of representative journals, and the reference tools and fundamental bibliographic apparatus pertaining to the subject. In (American) law collections, this includes comprehensive trade publications and loose leafs, and for foreign law, periodicals and monographs.

4 research level: a collection that includes the major published source materials required for dissertations and independent

research, including materials containing research reportings, new findings, scientific experimental results, and other information useful to researchers. It is intended to include all important reference works and a wide selection of specialised monographs, as well as a very extensive collection of journals and major indexing and abstracting services in the field. Older material is retained for historical research. Government documents are included in (American) and foreign law collections.

5 comprehensive level: a collection in which the library endeavours; so far as is reasonably possible, to include all significant works of recorded knowledge (publications, manuscripts, other forms), in all applicable languages, for a necessarily defined and limited field. This level of collecting intensity is one that maintains a "special collection"; the aim, if not the achievement, is exhaustiveness. Older material is retained for historical research. In law collections, this includes manuscripts, dissertations, and material on non legal aspects.

Attached to these codes on collection strength are codes reflecting language coverage. These language codes are indicated by letters added after the number. These codes are:-

E for primarily English language material

F for selected foreign language material

W for wide selection of foreign language material

Y for material primarily in one foreign language

These codes combine to give a rating for current collection strength and current collecting intensity. This will provide, for each subject, a code such as:-

4W/3F

Where the first number/figure is the current collection strength and the second number figure reflects the current collecting intensity of the institution in that subject area.

Some (93) have criticised Conspectus as being no more than a set of codes for the subjective evaluation of parts of collections, assessed by a number of individuals in groups. Or (94), as little more than an extravagantly designed and assiduously propagated bushelful of best guesses - those guesses not being aided by the difficulties involved in interpreting the categories. In particular, phrases like "endeavours, so far as is reasonably possible" have been especially criticised. It is claimed that the qualifiers in such a phrase rob the definition of the rating of its capacity to convey the same meaning to different readers. A recent analysis of the Conspectus results for twenty collections in French language and literature suggest that assuming objectivity for the figures contained in them would be difficult (95). Although the study claims to have detected certain broad correlations in the results, this is only mildly true for the distribution and quite untrue for the range. Five of the twenty libraries which awarded themselves scores of four were discovered to possess 62%, 62%, 60%, 52% and 31% respectively of the titles in the bibliographies against which they were compared. Another library that claimed a rating of three/three (past/present) holdings owned only 17% of the titles - a level which some consider to be alarmingly modest to be able to fill the category description of being adequate to support undergraduate and most graduate instruction, or sustained independent study.

To get around this problem some conspectus areas have made modifications to the basic weighting structure in order to achieve greater precision of definition. The Pacific North-west conspectus modified the six RLG levels into a total of ten levels (96). These are -

0 library does not collect
1a minimal, uneven coverage
1b minimal, well chosen
2a basic information
2b augmented information
3a basic study (undergraduate)
3b intermediate study
3c advanced study
4 research level
5 comprehensive

Despite the fact that these levels allow for a greater amount of differentiation between collection levels, the validation of the results still remain a matter of concern. A concern of many is that the collection level evaluations be based on objective elements to the greatest possible degree while also allowing for the evaluator to exercise his/her professional judgment. In order to ensure that this is in fact the case it must surely be possible in some way to verify the results obtained. In the case of the Pacific North-west librarians who have rated their collections at the same level, they have been asked to describe those collections in quantitative and qualitative terms. This, after review, is hoped to provide descriptions of sample collections to aid other libraries in gauging the accuracy of their assessments. Exactly what describing the collections in terms of quantity and quality means or what form the review will take is not spelt out however and this is a common difficulty with such efforts to establish authenticity of results.

Particularly within the UK, other reservations also apply to the possible implementation of Conspectus. The subject areas within the system are based on the Library of Congress (LC) classification scheme and therefore reflect the inadequacies and pro-American bias of that scheme. The cost of implementing Conspectus can also be considerable. At a time when there is major concern over funding levels for public and academic libraries in a wide range of areas, libraries may find it difficult to justify the cost elements associated with conspectus: the staff time involved, which is the major cost, documentation, training, and keyboard inputting of data. Another concern and a reason for not wishing to implement Conspectus is a fear that its results may be used for cost cutting purposes and that its existence may encourage arbitrary and adverse decisions in the present funding climate within research libraries in the UK (86).

Despite these flaws and reservations Conspectus has a large number of positive proponents willing to declare that (97) the Conspectus is a better map of research library collections and collecting strengths than we have ever had and that much benefit has been derived from the very process of making the map. Local use of the Conspectus information has proved of value to some libraries already (98) and potential benefits have led to many adopting the Conspectus including the Scottish Conspectus which was initiated by the Working Group on Library Co-operation and co-ordinated by the National Library of Scotland (NLS). Eleven Scottish libraries were involved in the project. These were the eight university libraries, two major public libraries (Edinburgh Central Library and the Mitchell Library, Glasgow), and the NLS (99).

This has led to the production of the Scottish conspectus database. This contains all the information derived from the Scottish conspectus program and also the British Library's conspectus data on its own collection. It is from this database that the information has come which has been used to create the expert system for OPAC selection. This database however is very large, taking up some 13Mb (Megabytes) of disk space. This is an extremely large amount of data and in a project of this scope it would be impractical to attempt to produce a system that uses it all. There are 26 main subject divisions in the initial level of the database. This is shown in the following list. Also shown here are the LC classes on which the conspectus subject divisions are based:-

Philosophy and religion	B-BX
Auxiliary sciences of history	CB-CT
History	D-F, JK-JN
Physical geography and Earth sciences	GA-GF, QE
Cartographic materials	G
Anthropology and recreation	GN-GV
Economics and sociology	H-HX
Political science	JA-JX
Law	K-KT
Education	L-LT
Music	M-MT
Art and architecture	N-NK
Linguistics, languages and literatures	P-PZ
Physical sciences	Q-QD

Natural history and biology	QH-QR
Medical and health services	R-RZ
Agriculture	S-SK
Technology	T-TX
Military science	U-UH
Naval science and shipbuilding	V-VM
Library and information sciences	Z
Government documents	[J]
East Asia	
South Asia	
Scottish studies Update	

From these main divisions three were selected to be looked at for use in the demonstration system. These are Library and information sciences, Psychology, and Education. These areas were selected because within them there were a manageable number of subdivisions with which to operate and provided a range of subjects in different areas. For details on how the subject ratings obtained were converted into rules within the expert system see section 6.1. Other sources of information that were considered for use in the creation of the expert system are looked at in the following sections, 5.3 to 5.5.

5.3 University Funding Council report

In 1989 the University Funding Council (UFC) produced a report on their research selectivity exercise, which was carried out in that year (100). This exercise was carried out to assess the quality of research done within each department, or "cost centre", in all UK universities. The eventual rating would help to determine the grant level that would be allocated to that university in the future. This perhaps goes some way to explain concerns that conspectus ratings might also be used for a similar purpose by funding bodies.

The ratings produced were on a scale of 1-5 as follows:-

1 Research quality that equates to attainable levels of national excellence in none, or virtually none, of the sub-areas of activity.

2 Research quality that equates to attainable levels of national excellence in up to half of the sub-areas of activity.

3 Research quality that equates to attainable levels of national excellence in a majority of the sub-areas of activity, or to international level in some.

4 Research quality that equates to

attainable levels of national excellence in virtually all sub-areas of activity, possibly showing some evidence of international excellence, or to international level in some and at least national level in a majority.

5 Research quality that equates to attainable levels of international excellence in some areas of activity and to attainable levels of national excellence in virtually all others.

These ratings were applied to research in all areas of university activity. One hundred and fifty seven areas of research, based around "cost centres" e.g. clinical dentistry, cell and structural biology, or corrosion science and engineering, were given ratings for each university that carried out activities in that area.

It was considered trying to use these ratings to provide information for the expert system. This would have been done on the basis that those departments with a high research rating would have required a comprehensive and high level collection of material in their parent university library. In those institutions with a low rating in an area, and presumably a lower priority on that area, the library collection would be less comprehensive.

This is not necessarily the case however. There is no evidence to indicate whether this hypothesis is in fact true or not. Given this, and the fact that the conspectus database became available, it was decided not to pursue the use of this data for construction of an expert system. Conspectus, although covering only Scottish institutions, does provide data which is direct evidence of the strengths of library collections. The UFC data however, although it may provide an indication as to collection strengths, cannot be shown to have a direct link to library collections.

5.4 University Grants Committee report

The University Grants Committee (UGC) (precursor to the UFC) in conjunction with the Committee of Vice-chancellors and Principles produced annual reports on university management statistics and performance indicators covering universities in the UK. Among many other items these reports carried information on university libraries expenditure. This included information on:-

Library expenditure as a percentage of general expenditure

Publications as a percentage of library expenditure

Pay expenditure as a percentage of library expenditure

Library expenditure per Full Time Equivalent (FTE) student

Library expenditure per FTE academic staff

Expenditure on books per FTE student

Expenditure on books per FTE student

These type of figures allow a general view of library expenditure to be arrived at. For example in the period 86-87 it could be seen that Oxford and Cambridge university libraries spent about double the national average per FTE student, whereas Cardiff and Heriot-Watt spent only 3/4s of the average per FTE student (101).

These reports also included figures, similar to these produced for library expenditure, for "cost centres" within universities. The UGC reports however only included 37 of these "cost centres" rather than the 157 the UFC required for their report. It was considered using these figures on expenditure and income related to "cost centres" e.g. physics, computing, or law, as a potential rating for the institutions library collection in that area. That is, if expenditure was higher than average in a particular area then it was likely that more was spent by the institutions library to support the work of that department, or group of departments. Higher spending on that area by the university as a whole, per FTE staff/student, would reflect a high priority on that area which was liable to be reflected in the library collection, since the library would also give a high priority to that area.

As was the case with the ratings produced by the UFC (section 5.3) there is no hard evidence that this is in fact the case. On the figures produced in these reports there is no way to directly determine likely areas of subject strength in library collections. Again, as with the UFC figures, when the conspectus data became available it was decide to abandon these figures as a possible basis for an expert system to select OPACs through the JANET network.

As detailed in section 5.1 however the classic source of information for expert systems is human experts in a subject area. By means of questionnaires it was hoped to obtain information which could be of use, directly from the users of JANET OPACs. Details of this are given in the following section.

5.5 Questionnaires

5.5.1 Design and response

In an attempt to gather some information from persons who actually make use of JANET to search in OPACs, a questionnaire was devised (for details of this see appendix A). Copies were distributed to all Universities and Colleges in the UK who possessed an OPAC available through JANET. The intention was that the questionnaire would therefore reach the persons who were most familiar with the use of OPACs through JANET, those people being the most likely to have valid or "expert" opinions on the OPACs they use. Despite the distribution of questionnaires to over 40 establishments the return rate was disappointing. Follow up copies of the initial questionnaire were dispatched to all those bodies who had not responded but still the response was low. In total 12 returned and completed (or partially completed, in some cases) questionnaires were obtained out of the 42 organisations they were sent to.

This return rate is too low to be able to come to any valid conclusions, especially since most of them were not fully completed. A brief summary of some of the results (those for which there were most responses) may prove useful by giving some indication of user's opinions at the time the questionnaires were distributed (February - May 1990).

On question 4:- For how many years have you been using JANET in order to search the OPACs of other academic institutions?. Some 67% of respondents had been using JANET in this way for more than one but less than three years.

On question 5:- On average how frequently do you search OPACs using JANET?. 50% searched OPACs only once per fortnight or less frequently and 50% used the service once a week with 16% of those using it once per day.

On question 6:- 50% of respondents were searching exclusively for items that were known i.e. those where the author, title or ISBN, etc. was known to them beforehand. In other words they are looking for evidence of the existence of specific items occurring within a collection.

For question 7:- Asking how they searched for the existence of a known item within a collection the vast majority searched by author and title, some way behind were those searching by author or title alone with very few supporters of searching by ISBN (no one rated searching by ISSN).

On question 8:- Asking how they searched in order to produce a list of documents on a subject, almost all searched by a subject or title keyword as their primary method. Next most popular was subject heading with very few opting for class number, and nobody choosing this as their primary option.

On question 9:- Asking how many OPACs they would expect to have to consult in a search, 50% expected to have to look at 2 or 3 different OPACs when looking for a specific known item. When looking to produce a list of documents on a subject area however, 50% would only expect to look at 1 or 2 different OPACs. Only one respondent indicated that he would look at more, specifying that sometimes up to 10 were consulted by him.

Question 10:- Asking if there was an OPAC that would be used in all or most enquiries, generated a list of OPACs that were mentioned only once; London, Aston, Aberdeen, Glasgow, Stirling, Dundee, Oxford, Edinburgh (2 mentions) and Cambridge (5 recommendations). Comments included Cambridge as "the best", "the biggest collection" and "very full". Other OPACs were listed as "local library" though a few were given for coverage of "relevant subject area".

In summing up, in a general fashion, it could be said that the expert user of OPACs through JANET fit the following pattern. They have been using JANET to search OPACs for about two years. They use it for this purpose about once a week in order to check the existence in a nearby library of a specific known work they wish to locate. It can be seen that the "user model" generated from the returned questionnaires is rather general and does not really cast any useful light on the problem area except to suggest that there is plenty of scope for the use of these OPACs to be expanded.

An example of the questionnaire used is provided in Appendix A. As so few were returned completed no detailed analysis of the results have been attempted as this would be unlikely to yield any useful information.

5.5.2 Dundee Institute of Technology

The library of Dundee Institute of Technology, as previously mentioned, (section 3.1) has developed a simple program to ease the matter of user access to certain OPACs. This program is named the Remote Library Catalogue Access Facility. Using this facility, by means of a simple menu system, users are allowed to gain access to not only the Institute's in-house OPAC but also the OPACs of the Universities of Aberdeen, Dundee, Edinburgh, Glasgow, St. Andrews, Stirling, Strathclyde and Cambridge.

In order to analyse the extent to which this service was being utilised and its effectiveness, the library decided to carry out a survey by means of leaving questionnaires beside the computer terminals in the hope that users would complete them and hand them in to the library staff. In the end 130 questionnaires were completed and returned to the library. An example of the questionnaire used can be seen in Appendix B. As collaborating body in this study the Institute's library kindly granted permission to examine the returned forms to see if any useful data could be gathered from them for use in this study. No detailed analysis is provided here as the Institute intends to publish an article detailing the results of the work themselves.

Some broad figures, however, may be of relevance in determining the types of user that avail themselves of such a facility when it is available and what they use it for.

Perhaps not surprisingly some 65% of the use of this facility was made by students rather than staff, and of those, the largest single group was from the Business Studies course which had 25% of the total use.

80% of the users had consulted the Institute's own OPAC before going on to search an external one through JANET. Of those, some 50% went on because the Institute's own library did not possess the item(s) they required (or they could not find them in that catalogue). A further 25% went on because although the Institute's OPAC showed that they possessed the item it also showed that the item was currently unavailable. Some 20% were carrying out a literature search on a subject area and searched other OPACs to provide a wider range of items.

The searches conducted were largely by title or title keyword, about 60 % of users searching in this manner. 40% searched by author and 20% by subject with no users searching by classification notation.

Most users were seeking to find books (80%) with only 30% interested in finding journals. This is perhaps a reflection of the coverage of the OPACs available where compared to coverage of book stock, the coverage of journals tends to be less both in terms of the amount of the journal stock covered and in the detail of the entries given for journal material.

As to the use of the information obtained from the OPACs, only 15% of users considered that the information obtained was of no use at all to them with the majority (70%) finding the information moderately useful or better. 50% considered the information either useful or very useful. Fewer of them actually did anything with the information however. Some 25% attempted to obtain the book direct from the relevant library. In this case the source library was most often Dundee University library due to its proximity. 20% used the information for "reference/bibliographic" purposes and only 10% used the information to generate an interlibrary loan request.

Of the range of OPACs available to be searched there was, perhaps not surprisingly, a definite bias towards the geographically closest, this being Dundee University library. 80% of users searched this OPAC, indeed some 50% of all searches were carried out here. The rest did relatively poorly only being searched by about between 5-20% of users.

From those that did use other OPACs however, when it comes to ease of use there were two clear favourites. Of those that searched them (70% and 65%, respectively) found the OPACs of Strathclyde and Cambridge easy to use. These are by far the best figures for ease of use with the other OPACs being considered easy to use by only 20-30% of their users. When it came to OPACs considered difficult to use most were so considered by 10-20% of users. The only exception to this was Aberdeen University's OPAC which was considered difficult to use by 40% of those who used it. This is perhaps somewhat surprising as Aberdeen University library use a Dynix OPAC which is the same type as used by Dundee Institute of Technology. It might have been thought that familiarity with their own institution's OPAC would have made it easier, or at least less difficult, for them to use another OPAC of the same type. In this case however this does not appear to be true.

When asked what it was about an OPAC that made it difficult or easy to use, not many of the users found it possible to specify. Only around a third of users expressed an opinion on this and of those that did express a view there was a fairly evenly spread between the options offered. The options were -

screen layout

clarity of prompts

help screens

speed of response

variety/lack of different types of search facility

ability to switch easily between different searches

Only a few users expressed other concerns. One was concerned by a lack of familiarity with OPACs, one considered the log off procedure to be unclear, another thought that subject searching did not allow him to be selective enough, and one more wanted a colour screen on the VDU.

On the whole not a great deal can be drawn from either questionnaire that is of direct relevance to this study. They do however raise some interesting points. For example it seems that in many cases geographical location will outweigh any other consideration as to which OPAC to search. The fact that it may be possible to remotely search an OPAC is more attractive if, in the case of a search being successful, it is possible to gain access to that item almost immediately in another library. This option however has been selected by users largely without information on subject strengths of collections and it may be that had they possessed that type of information, they would have been more willing to search OPACs outwith their immediate area.

6. System construction

Once it had been decided what information would be used as a basis of an expert system to select OPACs through the JANET network (section 5) and what tools, Crystal (section 4), would be used in its construction it was necessary to combine the information with the tools to produce a working system.

6.1 Explanation of Conspectus

The conspectus database provides codes (section 5.2) for each of the libraries for a particular subject area. For example, the subject of automatic indexing provides the following codes:-

Aberdeen university library	1E1E
British library	4W4W
Dundee university library	0*0*
Edinburgh university library	1E1E
Glasgow university library	3E3E
Mitchell library	1E2E
St. Andrews university library	2E1E
Stirling university library	0*0*
Strathclyde university library	1E1E

These codes indicate, first the current

collection strength, followed by current collection intensity. A code of 2E1E therefore indicates a current collection strength of 2 and a collection intensity of 1.

The codes are set out in more detail in section 5.2 but briefly summarised they are:-

0 Out of scope
1 Minimal
2 Basic information
3 Instructional support
4 Research
5 Comprehensive

The letter accompanying each rating code is a language indicator. Again these are detailed in section 5.2 but are briefly:-

E primarily English material
F selected foreign material
W wide selection of foreign material
Y primarily in one foreign language

It can thus be seen that a rating of 2E1E shows a basic informational collection primarily in English where the current collecting intensity is only at a minimal level in again, primarily English language material.

If one excludes the codes for the British library and the Mitchell library, since this system is concerned with selecting university library OPACs only, it can be seen that on the subject of automatic indexing Glasgow university library has the highest rating, 3E3E. In some subjects however there is not always so obvious a choice and some procedure must be arrived at which allows the consistent selection of a rating as the "best" rating for that subject.

6.2 **Procedure for OPAC selection**

The selection of an OPAC as the "best" OPAC within that subject area is dependent on a judgment being made using the conspectus ratings. The OPAC with the best conspectus rating is judged as being the best OPAC to search in for material on that subject. The procedures used to derive the judgment as to the which OPAC is best are set out below. In drawing up the criteria for judging which rating is to be preferred a largely intuitive method has been used. In effect my own judgment has been used. I am the expert in this case, applying my own knowledge in an intuitive fashion to the data provided by conspectus. This judgment can be summarised by saying that the rating with the highest current collection strength in the widest range of languages indicates the best OPAC in which to search. The order of priority for selecting a rating as the best is shown below, priority 2 only coming into effect if priority 1 does not produce a clear choice. Priority 1 selects an OPAC on the basis of current collection strength. With ratings of 3E2E and 2E2E this would select the first rating as best since the current collection strength for that subject is higher. Priority 2 selects on the basis of current collecting intensity. Ratings of 3E3E and 3E2E would not be differentiated by priority 1 since the current collection strength rating is the same. If that rating is the same then priority 2 is used to select an OPAC. In this case the rating of 3E3E would be preferred to 3E2E since the current collecting intensity is greater. These are the main priorities for selection. The rest of the priorities reflect the difference in the language codings. These prefer first, a collection with a wide range of foreign language material, then a collection containing selected foreign material, then a collection primarily in English with little foreign material in addition, and lastly selected would be a collection primarily in one foreign language.

For example if four ratings are looked at:

1	3E2E
2	3W2E
3	2E1E
4	3E1E

Priority 1 would remove the third rating as having the lowest current collection strength. Priority 2 would remove rating 4 since it has a lower current collecting intensity. This leaves ratings 1 and 2 which are identical in collection strength and intensity. Rating 2 would however be selected as the best rating out of these since priority 3 would differentiate between them on the basis of the language code on the current collection strength on rating 2. The code indicates that the collection has a wide range of foreign language material within it as opposed to the collection represented by rating 1 which is primarily in English.

1 The highest rating for current collection strength is preferred so long as the language code is not Y.

2 The highest rating for current collecting intensity is preferred if the language code is not Y

3 Language code W on current collection strength is preferred

4 Language code F on current collection strength is preferred

5 Language code E on current collection strength is preferred

6 Language code W on current collecting intensity is preferred

7 Language code F on current collecting intensity is preferred

8 Language code E on current collecting intensity is preferred

9 The highest rating for current collection strength is preferred even if the language code is Y

10 The highest code for current collecting intensity is preferred even if the language code is Y

These priorities for selection are those used in this project to assess which conspectus ratings to prefer. They are the ones which are applied within the demonstration system. They are derived from the conspectus ratings in an explicit fashion. By this I mean that nowhere in conspectus does it state that a rating of 3E2E is "better" than a rating of 3E1E. That judgment has been made separately, for the reasons above, and it is the fact that this judgment has been made that renders the information "expert". These are not however the only priorities that could be used. By shifting the emphasis on to non English language materials and changing the above procedures different result could be obtained. Such results would be no more or less valid than the results obtained by the procedures above, they would merely reflect a different set of priorities.

In some cases even following the full set of selection priorities will not produce a single

OPAC. The demonstration expert system selects one OPAC to represent the subject. In a fuller system however the user could either be presented with a choice of OPACs available or other selection procedures could be taken into account. For further development of other selection procedures see section 7.2.

6.3 Rules used in the expert system

Once a rating, and therefore an OPAC, has been established as being the best choice within a subject area this information must somehow be converted into rules that can be used by an expert system. In a procedural way a rule can be established for the area of automatic indexing. The rule would be:-

IF (subject is) automatic indexing THEN (OPAC choice is) Glasgow university

However since the conspectus data is structured in a hierarchical form the rule must be more complex, in order to reflect this. For example:-

IF (main division is) library and information science
AND (category is) library science, information science
AND (subject is) automatic indexing
THEN (OPAC choice is) Glasgow university

This form of rule reflects the way in which the data is structured and provides an obvious way in which a path can be traced through it, from the general to the specific. Crystal however does not use procedural rules in a form such as this. The structure of the rule can nevertheless be followed. In Crystal each of the screens provided to the designer of a system can be regarded as an element of a rule. The initial screen can be regarded as the IF element of the rule with subsequent, linked, screens being the AND, AND, and THEN sections. This can be seen in Appendix C which provides screen shots of Crystal rules illustrating the above example. Screen 1 is the introductory screen for the system. Screen 2 is the main menu, corresponding to the IF part of the rule, which allows selection of the main subject division, library and information science. Screen 3 is the first AND element which allows selection of the subject category of library science, information science. Screen 4 is the second AND condition allowing the final element of the subject, automatic indexing, to be selected. This leads on to the final illustration which is the THEN element, the subject having been selected, the system recommends its choice as Glasgow university.

In this way the rules and the knowledge required for the selection of the OPAC are built into the structure of the system. The inference engine and the knowledge base are thus combined in this structure. The rule used in any particular case is used because that rule also contains knowledge. Possible separation of the rules and knowledge is looked at in section 7.2.

By means of the menus presented to the user at each stage of the process this structure can be extended to cover other possible paths leading to a wider range of subjects. As mentioned in section 5.2, in order to keep the size of this demonstration system manageable only certain subject areas have been included. These are:-

Library and information science

Psychology

Education

Each of these is in itself a large domain and so not all the areas within these broad subject divisions have been included. Only a large enough portion to provide several options at each choice the user has to make and so give a general feel of what a larger, more comprehensive, system would be like. The choice of subjects is also limited by those available in conspectus. The possible main divisions, categories and subjects in conspectus are based on the Library of Congress classification (with some modification) and only the subjects are rated. It is thus not possible to be either more general or more specific than the subject ratings in conspectus allows.

Once an OPAC has been selected by the system the next task it should carry out is to automatically link the user to the selected OPAC, in order that he may carry out a search for the subject of his choice. The procedures for carrying out this function are looked at in the following section.

6.4 Communications

Assuming software is available that is able to link with, or activate, other software the communications is not a significant problem. There are many widely available communications packages which would be suitable for a task such as this, Crosstalk being a commonly found example.

Like many such programs it allows the creation of command files, small program-like devices, which can be used to store the required information to access, among other things, OPACs via JANET. These command files can be triggered, either from within Crosstalk while already running this program, or from outside the program by using DOS batch files. A typical command file might look like this:-

```
Name        Glasgow.library
Number      0000 7110 0005
Speed       1200
Go          R 45
```

In this example the first line identifies the OPAC as that of Glasgow university library. The second line is the NRS address required to call that OPAC. The third line refers to the baud rate at which data will be received and the final line instructs Crosstalk to call every 45 seconds until a connection is made. These files can easily be created using a text editor in much the same way as DOS batch files are created (102).

A DOS batch file is a small program created in DOS using a text editor. Suitable editors for easy creation of these files are EDLIN or RPED which are included as a part of MS-DOS. When these files are created and named they are given file extension of ".bat". This is what signifies them as batch files. A typical batch file might look like this:-

```
CD\Xtalk
Xtalk OPAC1
```

The first line causes the computer to change directory to a directory called Xtalk which is a sub-directory of the computers root directory. The second line causes the Crosstalk program to be run and triggers a Crosstalk command file to automatically cause Crosstalk to link to a predetermined remote host, in this case it would be an OPAC via the JANET network. If this file were given the designation "Xtalk1.bat", then when run, by entering "Xtalk1", it would run Crosstalk and cause it to link to the designated target.

Crystal allows the use of a command, DOS($,$), which can be used to run a batch file in DOS. To run the batch file in the previous example the command DOS("Xtalk1.bat") would be used. If this command is used in Crystal as the final outcome of a decision making process then at the end of that process the batch file can be run. Crystal will end running and the new program, in this case Crosstalk and its associated command file, will begin. By creating a range of batch files, each one connecting to a specific Crosstalk command file, the OPAC selected by the expert system can be automatically linked to if the user desires.

Since, firstly it cannot be assumed that a user wishing to look at the demonstration expert system will have access to Crosstalk, and since secondly complications can arise depending on the procedures require to gain access out of the LAN of a particular institution, this part of the procedure has not in fact been implemented in the demonstration system. Instead at the point where communication would be triggered a message confirming this will be presented to the user.

The system will thus only function as a guide as to how a more complete system would operate. It does however serve to illustrate some of the problems encountered and solutions required in the construction of such a system.

7. Conclusions

7.1 Evaluation of the project

Experts systems now have a place within LIS. That is the first conclusion that can be drawn from the work carried out for this project. As demonstrated (section 2) there is a wide range of possible applications for expert systems and an increasing number of situations in which expert system solutions are being applied. With the increase in choice and availability of expert system shells for system development the level of computer expertise required to produce an expert system is decreasing (section 4). This now makes them applicable by a range of people, to a range of problem areas, in a way that was not possible previously. It is therefore likely that their use in LIS will increase to the point where they become commonplace features in libraries and other information providing services.

Secondly it can be seen that the OPACs available through the JANET network are potentially a valuable source of bibliographic information (section 1). They are not however being used to anywhere near their full potential at the moment. Difficulties in accessing them and the problem of selecting which OPAC is worthwhile searching in for information on a particular subject area mean that they are not seen as easy to use, or reliable sources of information. If access to them could be provided on a more consistent and coherent basis it is possible that they would be increasing used. The information contained in the OPACs would therefore become available to a wider range of users and their utility value would increase. One way of attempting to do this is by making it more possible to gain useful information from them. If a user wishes to gain bibliographic information on a particular subject area he is more liable to use an OPAC to gain that information if he knows that the OPAC is likely to contain the type of information he wants. To do this he must have some knowledge of the subject strengths of the material listed in the OPAC. Such information had not been available until the advent of conspectus. The Scottish conspectus project however has produced a vast amount of data. Indeed, without simplification, probably too much information to easily be of use for selecting an OPAC.

The third conclusion that could therefore be made, is that an expert system that can use the available information to simplify the selection of an OPAC would be a valid solution to this problem. The expert system produced as a part of this project is a very simple model designed only to show some of the possibilities available, but the ideas (see following section) could be used to lead on to a larger, more comprehensive, and more useful system.

7.2 **Future developments**

In developing an expert system to select OPACs available on the JANET network on a subject basis further, both its size and complexity could be increased.

The size can be increased by enlarging the range of subject areas covered by the system. The full range of subjects covered by conspectus could be included. This would lead to a large and comprehensive system capable of selecting an OPAC for a vast range of subjects. This would make it of use to any enquirer wishing to access bibliographic information on a subject basis. If such a large system was not considered desirable then systems of intermediate size could be constructed. These could cover a specific range of subjects of special interest to a particular user group. Increasing the size of the system would not be a difficult task. The knowledge the system would use already exists in the form of conspectus, all that would be necessary would be to incorporate that information into the system in the required manner.

The complexity of the system could also be increased. This could be done by a combination of changing the way in which the system operates and by changing the user interface to ask more questions enabling the process of OPAC selection to be more complex by taking in to account a greater range of factors.

A major way in which the operation of the system could be altered would be to separate the rules from the data. At present the rules and the knowledge are combined. If they were separate then a relatively small number of rules could be used to operate on a large amount of data. This would make it easier to use a larger amount of conspectus data without having to duplicate rules and simulate the full structure of the conspectus database. Crystal, by means of various interface programs, allows data to be used in problem solving that is stored in other forms of program, databases or spreadsheets, etc. The conspectus database operates by means of a database management system called Tinman. With a suitable interface, data could be imported and exported to and from this system. The basic procedural form of visualising the rules could be maintained as in section 6.3 but the knowledge would not be contained implicitly or explicitly within those rules. when a rule was being tested it would be tested against data stored outside the structure of its rules.

The complexity of the system could also be increased by increasing the number and type of questions the system asks of its user. At present the demonstration system uses a series of menus which allow the user to specify a single subject area if a certain path is followed through the menu structure. Within the limitations of the expert system shell Crystal there are many ways in which this can be altered. As well as using menus it would be possible to allow direct user input into the system, i.e. when prompted the user would enter information directly using the keyboard rather than selecting an option. Sliding scales could be used to differentiate between two options and determine which was more important, and crucially how much more important one was than another. The user interface could be improved by making it resemble more of a hypertext or perhaps a Windows system. The provision of a more graphically oriented interface and the ability of the user to skip and move around in the enquiry process may make it easier for a user to utilise the system in the way they find easiest and which benefits them most. At present the rule structure and depth of the system is not great enough for this to of any benefit but this can be changed by taking other factors rather than just the subject into account. The basic rule structure could be altered from what it effectively is at present:

IF (subject is) A THEN (OPAC is) B

to something like:

IF (subject is) A AND (user priorities are) B THEN (OPAC is) C

User priorities are the additional factors here. These can be any factors other than subject information which could give guidance as to which OPAC to select. They could thus be thought of as forming a model of the user and of the purpose behind his enquiry. They could include things such as questions on the users geographical position. Users in a particular part of the country may choose to prefer a local OPAC at the expense of a "better" but more distant OPAC. The purpose of the enquiry can be important. Are they seeking to compile a list of bibliographic information or do they wish to locate a specific item and then to obtain that item by interlibrary loan?. If so then an OPAC that allows the electronic issue of an interlibrary loan request may be preferred. The kind of search the user wishes to make can be considered, by subject, by author, by classification number. Again the user priorities as well as subject information are being used to determine the OPAC. If a user was interested in a combination of two or more subject areas, an OPAC could be recommended that had the highest combined rating for these areas. By using the users defined priorities about a particular search, this rating could be biased towards certain subject areas or OPACs depending on the circumstances. In order to complement this increased complexity, inbuilt help screens could be provided to ensure the user is presented with enough information to use the system effectively. The complexity of the system in this respect is, in many ways, only limited by the conspectus data on which the system is based. Subject enquiries could not be more general since in conspectus the main subject divisions, and categories (see section 5.2) are not given ratings and the depth of the system is limited by the depth of the conspectus subject areas. Even within this there is considerable scope for making the system more complex and sophisticated. All these options for increasing the sophistication and complexity of the system were considered for implementation in the demonstration system. For the benefits of simplicity and clarity and to ensure that the demonstration system could be completed without the requirement for additional resources they have however not been carried through into the demonstration system.

The scope of the system may also eventually be increased. At present the system is based around the information in the Scottish conspectus. If institutions elsewhere in the UK were to carry out a conspectus project then they could also be covered by the system. If it could be shown that other information (such as that in sections 5.3 or 5.4) could be related to collection strengths, then that could be used to increase the scope both in geographical terms and in terms of the range of subjects and level of subjects covered. As pointed out in sections 5.3 and 5.4 however it would require further work to determine if other sources of information would be valid for inclusion as the basis for such an expert system. Since JANET (see section 1) also allows access to other networks in the UK and overseas it is possible that in the long run such a system could even extend Europe-wide although that is a distant prospect at present.

With further development a system such as this could be produced to a level of size, scope, and complexity that would make it a useful tool to a wide range of users, enhancing the eventual possibility of universal bibliographic access.

Bibliography

1. Waters, Samuel T.

Answerman, the Expert Information Specialist: An expert system for retrieval of information from library reference books

Information Technology and Libraries, 5, (3), Sept '86, p204-212

2. Carande, Robert

INDEXES: A micro-based expert system

Small Computers in Libraries, Dec '88, p26,27

3. Guide for libraries on JANET

Project Jupiter, Glasgow university library, 1990

4. Wells, Mike

JANET - the United Kingdom Joint Academic Network

Serials, 1, (3), Nov '88, p28-36

5. Buxton, Andrew

JANET and the librarian

The electronic library, 6, (4), Aug '88, p250-263

6. JANET.NEWS

(Bulletin board available on the JANET network)

7. Heeks, Richard

Expert Systems

New methods and techniques for information management, (Ed) Feeny, Mary

Taylor-Graham, Oxford, 1986, p227-232

8. Tsotsos, John K.

Expert system overview: Where does the phrase fit?.

CIPS Review, 9, (5), Sep/Oct '85, p12-14

9. Ercegovac, Zorana

Artificial intelligence and cataloguing: Reasoning, promises and pitfalls

Proceedings 7th National online information meeting, Learned Information, 1986, p109-117

10. Fenly, Charles

Expert systems, Concepts and applications

Advances in Library Information Technology, 1, '88

11. Sharif, Carolyn A.J.

Developing an expert system for classification of books using micro based expert system shells

British Library research paper No. 32, BLR & D Dept., 1988

12. Shoval, Peretz

Principles, procedures and rules in an expert system for information retrieval

Information processing and management, 21, (6), '85, p475-487

13. Anderson, P.F.

Expert systems, expertise, and the library and information professions

Library and information science review, 10, '88, p367-388

14. Hawkins, Donald T.

Artificial intelligence (AI) and expert systems for information professionals- Basic AI terminology

Online, 11, (5), Sep '87, p91-98

15. Barr, Avron; Feigenbaum, Edward A.

The handbook of artificial intelligence:vol 1

Pitman, 1981

16. Vickery, Alina; Brooks, Helen
Expert systems and their applications in LIS
Online review, 11, (3), Jun '87, p149-165

17. Addis, T.
Definition of an expert system
AISB Quarterly, 40/41, Spring/Summer '81, p39

18. Expert systems in British industry
BBC (video), Open university productions, 1985

19. Addis, T.R.
Expert systems: An evolution in information retrieval
Information technology research and development, 1, '82, p301-324

20. Clarke, Anne; Cronin, Blaise
Expert systems and library/information work
Journal of librarianship, 15, (4), Oct '83, p277-292

21. Parrot, James R.
Expert systems for reference work
Microcomputers for information management, 3, (3), Sep '86, p155-171

22. Canisius, P.
Knowledge based systems, chances and threats: A challenge to information specialists
International forum on information and documentation, 13, (2), '88, p5-6

23. Gibb, Forbes
Expert systems: An overview
Expert systems in libraries/(Ed) Gibb, Forbes , Taylor Graham, Oxford, 1986, p3-21

24. Davis, Roy; James, Brian
Towards an expert system for cataloguing: Some experiments based on AACR2
Program, 18, (4), Oct '84, p283-297

25. Hjerppe, Roland; Olander, Birgitta
Cataloguing and expert systems: AACR2 as a knowledge base
Journal of the American Society for information science, 40, (1), Jan '89, p27-44

26. Enser, P.B.G.
Experimenting with the automatic classification of books
Informatics 8: Advances in intelligent retrieval, Aslib, 1985, p66-84

27. Meder, N.

Artificial intelligence as a tool of classification

International classification, 12, (3), '85, p128-132

28. Burton, Paul F.

Expert systems in classification

Expert systems in libraries/(Ed) Gibb, Forbes , Taylor Graham, Oxford, 1986, p50-66

29. Brooks, H.M.

Expert systems in reference work

Expert systems in libraries/(Ed) Gibb, Forbes , Taylor Graham, Oxford, 1986, p36-49

30. Thompson, Don

AI - tailormade for librarians

Canadian library journal, 45, (2), Apr '88, p73-75

31. Krulee, Gilbert K. ;Vrenios, Alexander

An expert system model of a reference librarian

Library software review, Jan/Feb '89

32. The Datacomms book

VNU Business Publications, London, 1989

33. The computer users yearbook 1991
VNU Business Publications, London, 1990

34. Burton, Hilary D.
The intelligent gateway processor: Effective computer access to distributed information sources
IAALD Quarterly bulletin, 30, (3), '85, p55-60

35. Aragon-Ramirez, V; Paice, C.D.
Design of a system for the online elucidation of natural language search statements
Informatics 8: Advances in intelligent retrieval, Aslib, 1985, p163-190

36. Williams, P.W.
The design of an expert system for access to information
Proceedings 9th International online information meeting, Learned Information, Oxford, 1985, p23-29

37. Morris, A.; Tseng, G.; Newham, G.
The selection of online databases and hosts - an expert system approach
Proceedings 12th International online information meeting, Learned Information, Oxford, 1988, p139-148

38. Pollit, A.S.
An expert system as an online search intermediary
Proceedings 5th International online information meeting, Learned Information, Oxford, 1981, p25-32

39. Pollit, A.S.
A search statement generator for cancer therapy related information
Proceedings 6th International online information meeting, Learned Information, Oxford, 1982, p404-413

40. Pollit, A.S.
A front-end system: An expert system as an online search intermediary
Aslib proceedings, 36, (5), may '84, p229-234

41. Marcus, Richard S.
An automated assistant for information retrieval
Proceedings 44th annual ASIS meeting, Learned Information, 1981, p270-273

42. Black, W.J.; Hargreaves, P.; Mayes, P.B.
HEADS: A cataloguing advisory system
Informatics 8: Advances in information retrieval, Aslib, 1985, p227-239

43. Prangnell, R.D.
MARKUP: An expert system for marking up bibliographic data
Program, 21, (2), Apr '87, p174-182

44. Epstein, Hank
An expert system for novice MARC cataloguers
Wilson library bulletin, Nov '87, p33-36

45. Vickery, Alina; Brooks, H.M.; Vickery, B.C.
Developing an expert system for referral
Proceedings 2nd International expert systems conference, Learned Information, 1986, p285-301

46. Vickery, Alina; Brooks, H.M.
PLEXUS- the expert system for referral
Information processing and management, 23, (2), '87, p99-117

47. Vickery, A.; Brooks, H.M.; Vickery, B.C.
An expert system for referral: the PLEXUS project
Intelligent information systems: Projects and progress, p154-183

48. Vickery, Alina; Brooks, Helen; Robinson, Bruce

A reference and referral system using expert system techniques

Journal of documentation, 43, (1), Mar '87, p1-23

49. Wade, Stephen; Willet, Peter; Robinson, Bruce; Vickery, Brian; Vickery, Alina

A comparison of knowledge-based and statistically-based techniques for information retrieval

Online review, 12, (2), '88, p91-108

50. Vickery, A.; Brooks, H.M.; Robinson, B.A.; Stephens, J.

Expert system for referral

Library and information research report No.66, BLR & D Dept., 1988

51. Riggs, Donald E.

Productivity increases in public services: are expert systems the answer?

Computing, electronic publishing and information technology, 1, p89-99

52. Smith, Karen F.

Robot at the reference desk?

College and research libraries, 47, Sep '87, p486-490

53. Williams, Phil

User trials of the OASIS search system

Proceedings 10th International online information meeting, Learned Information, 1986, p437-452

54. Williams, P.W.

The design of an expert system for access to information

Proceeding 9th International information meeting, Learned Information, 1985, p23-29

55. Palmer, Judith; Yeadon, Janice

A training workshop in the use of JANET

Program, 22, (2), Apr '88, p182-183

56. Burton, Jim; Newport, Jose; Robinson, Everard

OPACs and JANET: a simple technique for easy user access

Program, 23, (3), July '89, p257-268

57. Andrea, Klitos; Duncan, Charles
SALBIN, report to the British Library, 1988

58. Ralls, Marion
SALBIN: freedom of access for users unlimited
Vine, 75, '89, p28-31

59. Black, W.J.
Software for developing expert systems
Expert systems in libraries/(Ed) Gibb, Forbes, Taylor Graham, 1986, p22-35

60. Clancey, W.J.
The epistemology of an expert system
Artificial intelligence, 20, '83, p215-251

61. Alberico, Ralph
Software for expert systems: languages versus shells
Small computers in libraries, 8, (7), July/Aug '88, p4-12

62. Sowizral, Henry A.
Expert systems
Annual review of information science and technology, 20, '85, p179-199

63. Hasemer, Tony
A beginners guide to LISP
Addison Wesley, Wokingham, 1984

64. Clocksin, W.F.; Mellish, C.S.
Programming in PROLOG (2nd edition)
Springer Verlag, Berlin, 1984

65. Addis, T.R.
Knowledge for a new generation of computers
Expert systems for decision making/(Ed) Baird, Patricia, Taylor Graham, Oxford, 1987, p18-32

66. Intelligent Terminals Ltd.
George House
36 North Hanover St
Glasgow, G1 2AD

67. Superexpert user manual
Intelligent Terminals Ltd., Glasgow, 1986

68. Shen, Stewart N.T.
An evaluation of expert shell ExpertEase
Library software review, 6, (9/10), '87, p303-311

69. Ford, N.J.

Crystal

International journal of information management, 7, '87, p171-172

70. Crystal user manual

Intelligent Environments, 1990

71. Neale, I.M.; Morris A.

Knowledge acquisition for expert systems: a brief review

Expert systems for information management, 1, (3), p178-192

72. Garg-Janardan, Chaya; Salvendy, Gavriel

The contributions of cognitive engineering to the design and use of expert systems

Behaviour and information technology, 7, (3), '88, p323-342

73. Wright, George; Ayton, Peter

Eliciting and modelling expert knowledge

Decision support systems, 3, (1), '87, p13-26

74. Myers, C.D.; Fox, J.; Pegram, S.M.; Greaves, M.F.

Knowledge acquisition for expert systems: experience using Emycin for leukemia diagnosis

Proceedings of expert systems '83, 1983

75. Barrett, M.L.; Beerel, A.C.

Building effective systems

Expert systems in business: a practical approach, Ellis-Horwood, 1988, p134-192

76. Evans, J.St.B.T

The knowledge elicitation problem: a psychological perspective

Behaviour and information technology, 7, (2), '88, p111-130

77. Kinnell, Susan K.; Chignell, Mark H.

Who's the expert?: conceptual representation of knowledge for end user searching

Proceedings 8th National online information meeting, Learned Information, p237-243

78. Johnson, P.E.; Zualkernan, I.; Garber, S.

Specification of expertise: knowledge acquisition for expert systems

Proceedings of the knowledge acquisition for knowledge based systems workshop, Banff: Canada, 19.0-19.37.

79. Grover, M.D.

A pragmatic knowledge acquisition methodology

Proceedings 8th international joint conference of artificial intelligence, Karlrhue, Germany, 1983, p436-438

80. Jagannathan, V.; Elmaghraby, A.S.

MEDKAT: multiple expert Delphi-based knowledge acquisition tool

Technical report, Louisville, University of Louisville, Dept. of engineering, maths, and computer science

81. Rusk, R.A.; Krone, R.M.

The Crawford slip method as a tool for extraction of expert knowledge

Human computer interaction, (Ed) Salvendy, G., Elsevier, New York, 1984, p279-282

82. Matheson, Ann

The planning and implementation of conspectus in Scotland

Journal of librarianship, 19, (3), July '87, p141-151

83. Gwinn, Nancy E.; Mosher, Paul H.
Coordinating collection development: the RLG conspectus
College and research libraries, 44, Mar '83, p128-140

84. Oberg, Larry R.
Evaluating the conspectus approach for smaller library collections
College and research libraries, 49, May '88, p187-196

85. Ferguson, Anthony W.; Grant, Joan; Rutstein, Joel S.
The RLG conspectus: its uses and benefits
College and research libraries, 49, May '88, p197-206

86. Matheson, Ann
Conspectus in the United Kingdom
Alexandria, 1, (1), '89, p51-59

87. Buckingham, Jeanette
NCIP, conspectus methodology and Canadian health science collections
Bibliotheca Medica Canadia, 8, (3), '87, p136-139

88. Crandlemire, Debbie
The development of Canadian conspectus on-line
The Canadian journal for information science, 13, (3/4), Dec '88, p110-119

89. Hanger, Stephen
Collection development in the British Library: The role of the RLG conspectus
Journal of librarianship, 19, (2), Apr '87, p89-107

90. Conspectus in the British Library
British Library, London, 1986

91. Pringle, R.V.
Conspectus in Scotland
Liber news sheet, 23, '88, p5-17

92. Reed-Scott, Jutta
he conspectus in north America and western Europe: a report on the program session at the ACRL/WESS
LIBER news sheet, 24, '88, p14-20

93. Allen, G.G.
A case against conspectus
Australian library journal, Aug '89, p211-218

94. Henige, David
Epistemological dead end and ergonomic disaster?. The North American Collections Inventory Project
Journal of academic librarianship, 13, (4), '87, p209-213

95. Larson, Jeffry
The RLD French literature collection assessment project
Collection management, 6, Spring/Summer '84, p97-114

96. Forcier, Peggy
Building collections together: the Pacific northwest conspectus
Library journal, 15th Apr '88, p43-45

97. Abel, Millicent D.
The conspectus: issues and questions
Proceedings of the association of research libraries conference, means to an end, 1986, p26-30

98. Pinnell-Stephens, June
Local conspectus applications
PNLA quarterly, 53, Spring '89, p22,23

99. Milne, Ronald

Conspectus at the coal-face

British journal of academic librarianship, 3, (2), p89-98

100. Report of the research selectivity exercise

University Funding Council, London, 1989

101. University management statistics and performance indicators in the UK

Committee of Vice-chancellors and Principles and University Grants Committee, London, 1988

102. Crosstalk user manual

Microstuf Inc., Roswell, 1984

103. Buxton, A.B.

International gateways

Standards for the exchange of bibliographic information, papers presented at a course at UCL, 1990.

Library Association, London, 1991

Questionnaire

An expert selector of Online Public Access Catalogues

Questionnaire

A. Introduction.

This questionnaire is intended to discover and collect the human knowledge and expertise which has been acquired by those people who use the Joint Academic Network (JANET) regularily in order to search the Online Public Access Catalogues (OPAC's) of other academic institutions. This information will be used to help in the design of the knowledge base of an expert system. This system is intended to enhance the useability of JANET by aiding the decision and selection making process of experts and non-experts alike.

JANET is a network linking computers in organisations in the UK concerned with research. This includes most of the UK academic institutions whose libraries have OPAC's. However since the provision of access to these OPAC's was not one of the primary objectives of JANET they do not present themselves as a coherent or consistent source of information to a JANET user. Since this is the case some assistance needs to be give to the user of the system to enable best use to be made of it.

An expert system situated on a JANET users intelligent terminal (microcomputer) could simplify the task of using these OPAC's to provide a list of documents on a subject basis. It would do this by automatically reccommending the most suitable OPAC for use on any particular occasion. That is it would reccommend the OPAC in which the likelihood of finding material relevant to the particular enquiry was the highest. This would be done in an "expert" manner, that is, a decision would be made by the system based on accumulated knowledge and opinion which may greatly exceed that of the individual user.

B. **Information about respondant.**

1. Please provide a brief description of your job. (eg general reference librarian, cataloguer, etc)

2. In your duties do you deal with the literature of:

   a) a particular discipline (eg in a physics subject department) : ___ :

   b) a group of disciplines (eg a humanities library) : ___ :

   c) all, or most, disciplines : ___ :

3. In regard to the disciplines whose literature you deal with, please list these below along with the level of the literature concerned. If you deal with a particular discipline or group of disciplines please list all of these. If you deal with all, or most, disciplines please list those you most commonly encounter along with the level dealt with and give a general level for the other material.

   In order to indicate the level of the material you deal with, please use the following scale:

   1 = Minimal level. Few items beyond very basic works.

   2 = Basic informational level. Up to date, general material, defining and introducing a subject.

   3 = Instructional support level. Capable of supporting under graduate and most graduate instruction as well as independent study.

   4 = Research level. Major source material required for independent research.

   5 = Comprehensive level. All, or almost all, significant works in a field, in applicable languages.

	Discipline	Level
eg	Inorganic chemistry	4

C. Information on your use of JANET

4. For how many years have you been using JANET in order to search the OPAC's of other academic institutions?.

   a) up to one year :___:

   b) more than one but less that two years :___:

   c) more than two but less than three years :___:

   d) more than three but less than four years :___:

   e) four or more years (please specify) :___:

5. On average how frequently do you search OPAC's using JANET?.

   a) once per day (if more often please specify) :___:

   b) once per week (if more often but less than daily, please specify) :___:

   c) once per fortnight :___:

   d) once per month :___:

   e) less frequently (please specify) :___:

6. When searching an OPAC do you:

   a) look for known item(s) of information. ie those where author or title or ISBN, etc is known beforehand. :___:

   b) look for a list of documents on a subject hoping to retrieve one or more previously unknown items. eg by subject heading, keyword, class no. etc. :___:

   c) both the above. :___:
   If so can you show which is the most common. Please specify a ratio of a to b. eg 1:2, or 1:1.

7. When searching for known items how do you most commonly search?. Please number the following options in ranked order of the frequency with which you use them (one being most common).

   Author Title Author and title ISBN ISSN

   Other (please list and rank any other options you use. eg a conference name, OCLC no. etc)

8. When searching for a list of documents on a subject how do you most commonly search?. Please number the following options in ranked order of the frequency with which you use them (one being most common).

Subject heading Title/subject keyword Class no.

Other (please list and rank any other options you use. eg name as subject etc)

8. When searching for either a) or b) below how many OPAC's would you expect to have to consult?. (please circle correct number)

a) known item(s)

1 2 3 4 more (please specify)

b) a list of documents on a subject

1 2 3 4 more (please specify)

9. Is there any OPAC that you would tend to use in all or most enquiries? If so please say which one this is and why you use it so commonly.

10. If you search for a list of documents on a subject are there any OPAC's that you would tend to associate with the subjects with which you are often involved. If so please list these frequently used search subject(s) and the OPAC's you would associate with them along with the level of information that you consider the OPAC to hold on that subject. In order to indicate the level please use the scale of 1 to 5 as defined below.

1 = Minimal level. Few items beyond very basic works.

2 = Basic informational level. Up to date, general material, defining and introducing a subject.

3 = Instructional support level. Capable of supporting under graduate and most graduate instruction as well as independent study.

4 = Research level. Major source material required for independent research.

5 = Comprehensive level. All, or almost all, significant works in a field, in applicable languages.

	Subject	Associated OPAC(s)	Level
eg	Architecture	Edinburgh univ.	3

D. **Information on the last occasion you searched a JANET OPAC.**

11. How many days ago did you last search a JANET OPAC?. If you cannot remember excactly please give an approximate figure.

12. For what subject or item was the search carried out for?. (eg books on X-ray diffraction, published after 1985, in english, suitable for under graduates)

13. Please give a ranked list of the OPAC's consulted for this enquiry, the reasons why these particular OPAC's were selected and the level of satisfaction with the results obtained.

Rank	OPAC	Reason for choice and satisfaction with result
1.		
2.		
3.		

(if required please continue on another sheet)

Thank you for your cooperation in this venture.

Appendix B

Dundee Institute of Technology survey

Questionnaire

DUNDEE INSTITUTE OF TECHNOLOGY
Library

|___|___|___|___|

After you have used the Remote Library Catalogue Access facility, we would be grateful if you would please complete this brief questionnaire. The purpose of the questionnaire is to give library staff a better understanding of what library users require from this facility, thus enabling us to improve it.

1. Please indicate your status in the institute by ticking the appropriate box 5

   - Staff (DIT) [ ]
   - Student (DIT) [ ]
   - Research Student (DIT) [ ]
   - External Borrower [ ]

   If you are an external borrower, please proceed to Question 3.

2. In which department do you teach or study? Please tick the appropriate box. 9

   - Civil Eng., Surveying & Bldg. [ ]
   - Electronic & Electrical Eng. [ ]
   - Mechanical Eng. [ ]
   - Accountancy & Economics [ ]
   - Business Studies [ ]
   - Mathematics & Computer Sciences [ ]
   - Molecular & Life Sciences [ ]
   - Central Services [ ]
   - Administration [ ]
   - Other. Please specify [ ]

3. *Did you consult the DIT on-line catalogue for an item/ items before searching on the Remote Library Catalogue facility?*	Yes	|___|	19
	No	|___|	
4. Why did you use this facility? Please tick as appropriate.	D.I.T. Library did not have required copy/copies	|___|	21
	D.I.T. Library copy/copies unavailable	|___|	
	For bibliographic details for references purposes, e.g. publication details	|___|	
	For a literature search to obtain further references, e.g. to find what books are available in your subject field.	|___|	
	Curiosity	|___|	
	Other. Please specify.	|___|	

5. *What type(s) of search did you conduct?*		27
Author		___
Title/Title Keyword		___
Author/title		___
Subject		___
Class		___
Other (Please specify)		___

6. What type of material were your searching for? Please tick as appropriate. 33

- Books |___|
- Journals |___|
- Other (Please specify) |___|

7. How useful was the information you found? Rate 1-5, 1 = very useful, 5 = useless |___| 36

8. *If the information was useful, what did you do with it?* Please tick as appropriate. If it was not particularly useful, leave blank. 37

- *Use for reference/bibliographic* purposes |___|
- Ask for a book or journal through interlibrary loan |___|
- Apply direct to the relevant library |___|
- Submit a 'Book suggestion' order form |___|
- Other. Please specify. |___|

9. Which catalogue(s) did you search? Please tick as appropriate. 42

- Aberdeen University |___|
- Dundee University |___|
- Edinburgh University |___|
- Glasgow University |___|
- St. Andrews University |___|
- Stirling University |___|
- Strathclyde University |___|
- Cambridge University Library |___|

If you searched only one catalogue, please finish here and give the completed questionnaire to an assistant. Thank you for your co-operation.

10. Tick the catalogue(s) you found easiest to use. 49

Aberdeen University	[ ]
Dundee University	[ ]
Edinburgh University	[ ]
Glasgow University	[ ]
St. Andrews University	[ ]
Stirling University	[ ]
Strathclyde University	[ ]
Cambridge University Library	[ ]
None	[ ]

11. Tick the catalogue(s) you found most difficult to use. 57

Aberdeen University	[ ]
Dundee University	[ ]
Edinburgh University	[ ]
Glasgow University	[ ]
St. Andrews University	[ ]
Stirling University	[ ]
Strathclyde University	[ ]
Cambridge University Library	[ ]
None	[ ]

12. *If you found differences in the ease of use between the catalogues,* was it because of one or more of the following. Please tick as appropriate.

Screen layout	[ ] 65
Clarity of prompts	[ ]
Help screens	[ ]
Speed of response	[ ]
Variety/lack of different type of search facility, e.g. subject, class mark searches	[ ]
Ability to switch easily between different searches	[ ]
Other. Please specify	[ ]

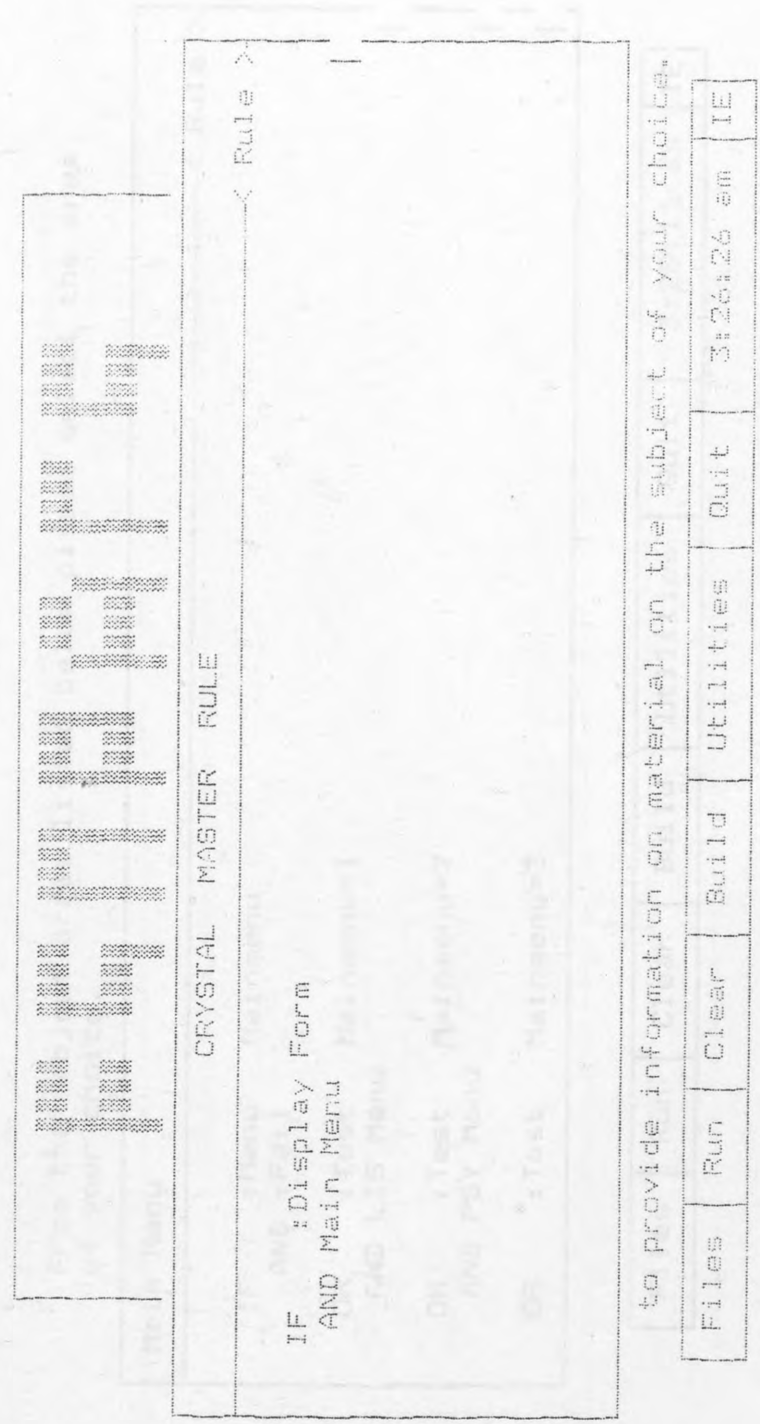
<u>Please give the completed questionnaire to an assistant. Thank you for your co-operation.</u>

<u>September 1989</u>

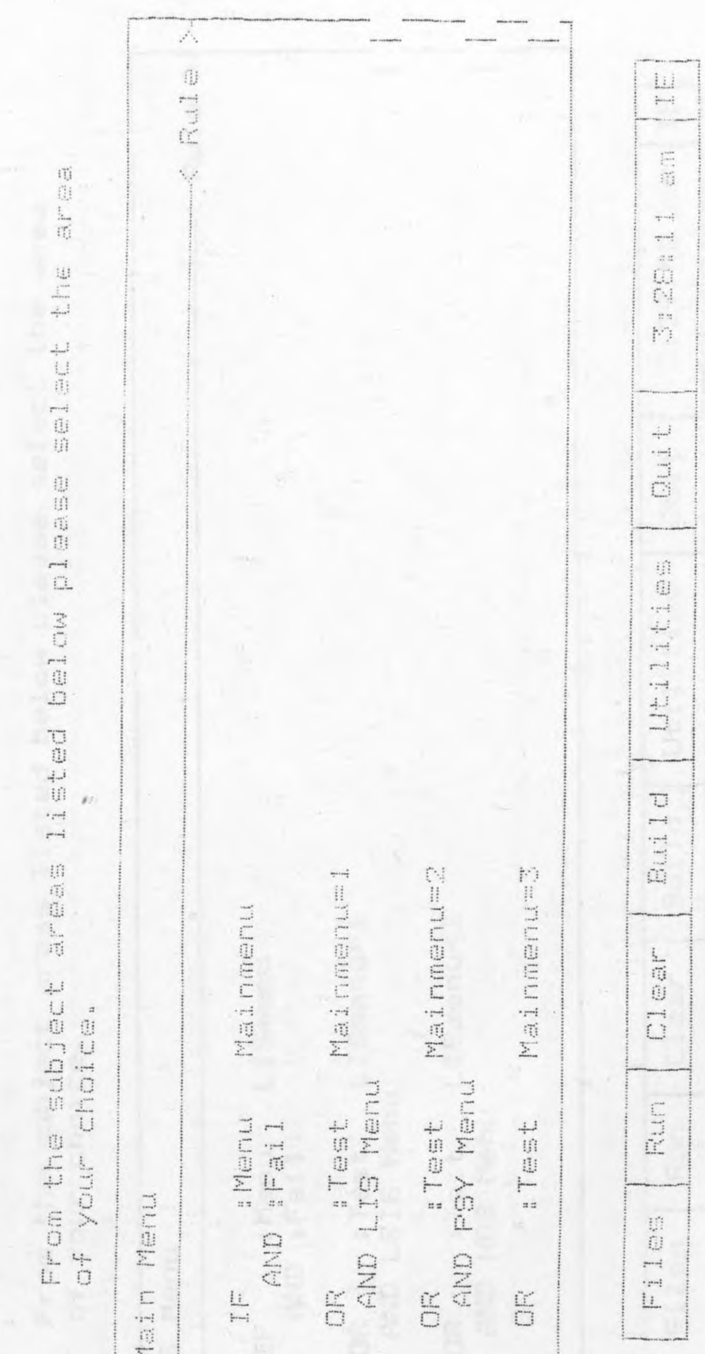
Appendix C

Screen shots of the demonstration expert system

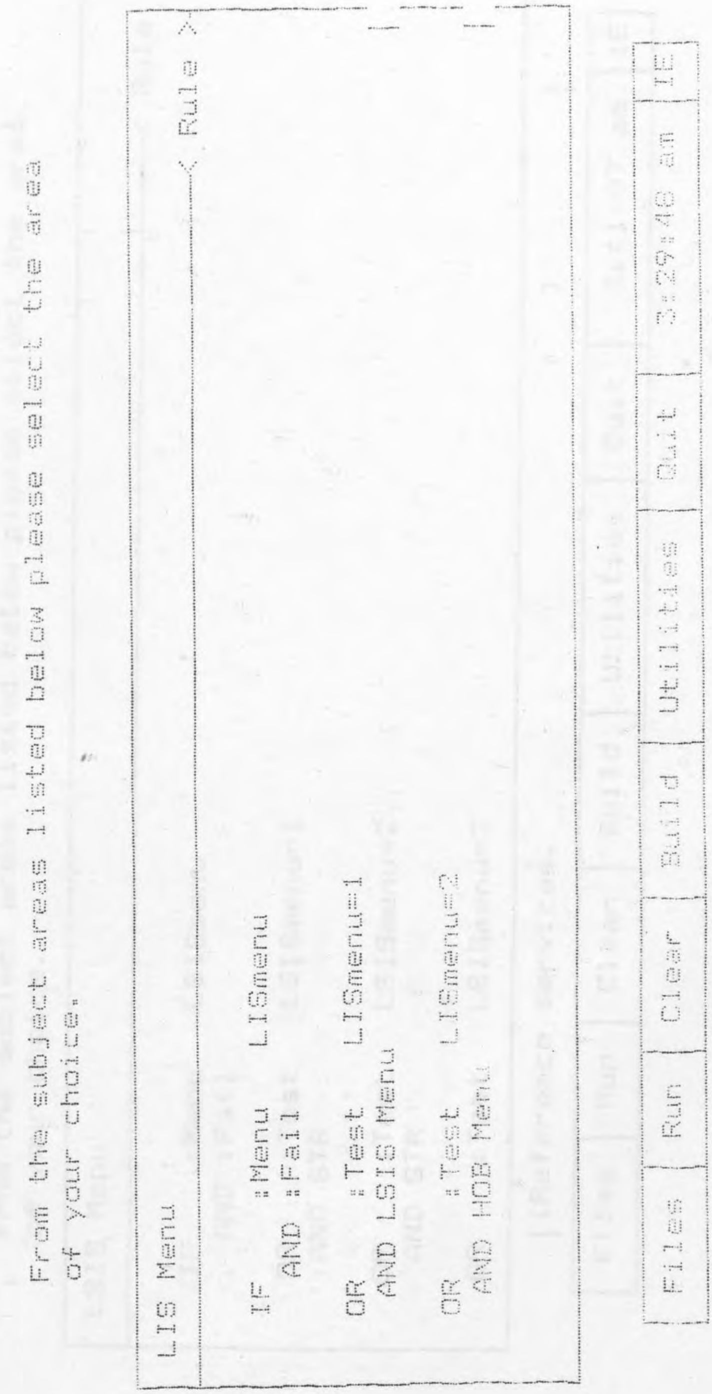
Screen shot 1



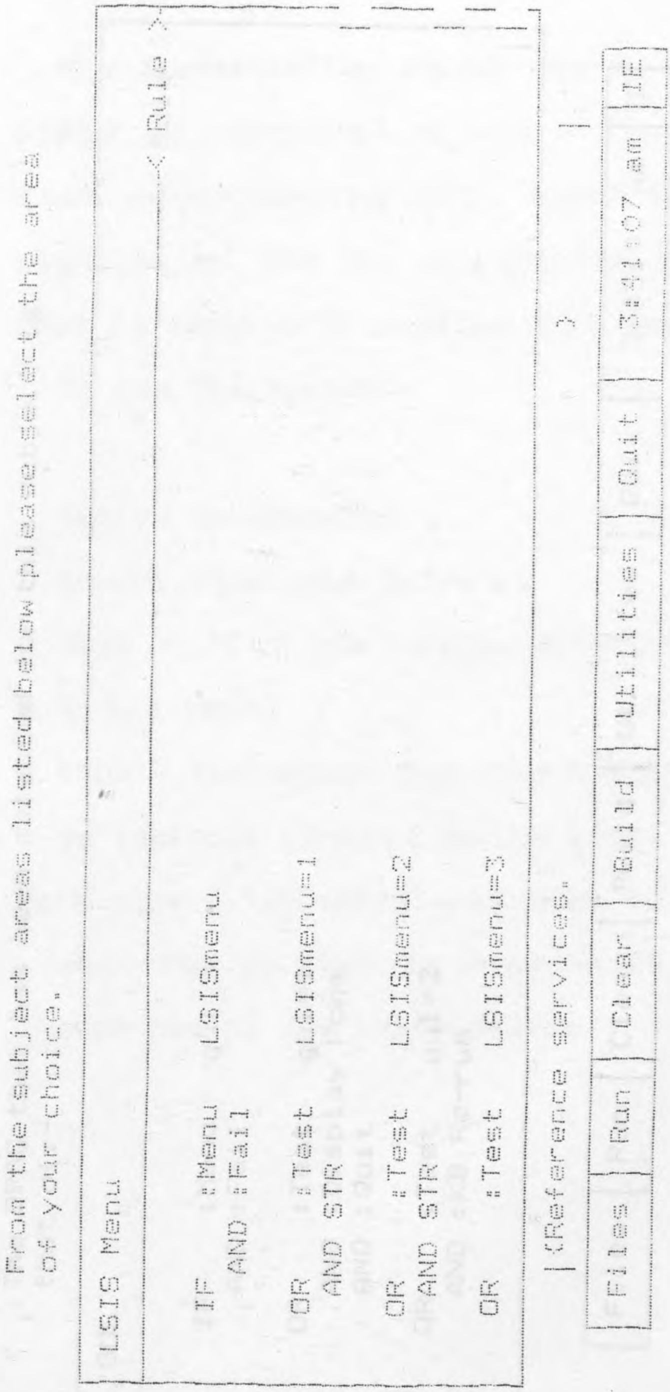
Screen shot 2



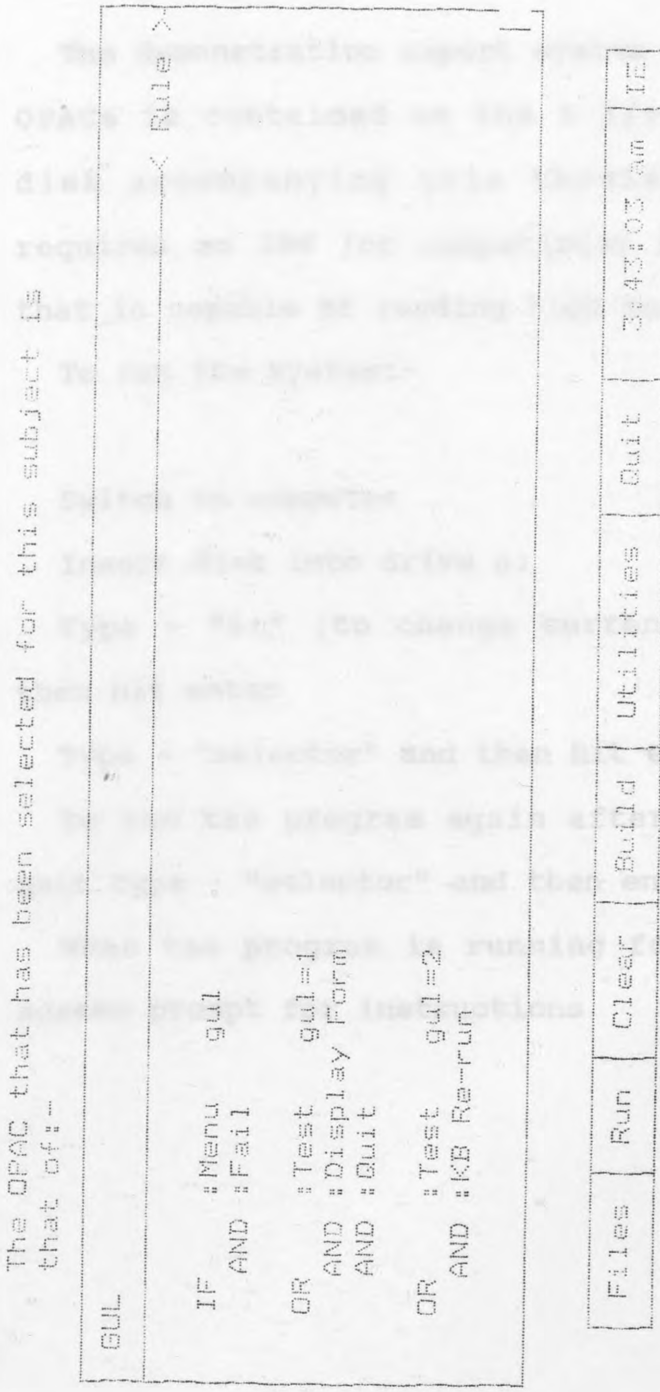
Screen shot 3



Screen shot 4



Screen shot 5



Appendix D

How to run the demonstration expert system

The demonstration expert system for selecting OPACs is contained on the 5 1/4 inch floppy disk accompanying this thesis. To run it requires an IBM (or compatible) microcomputer that is capable of reading high density disks.

To run the system:-

Switch on computer

Insert disk into drive a:

Type - "a:" (to change current drive) and then hit enter

Type - "selector" and then hit enter

To run the program again after it has been quit type - "selector" and then enter again

When the program is running follow the on-screen prompt for instructions