EXPERT SYSTEMS IN MAP DESIGN

Gary Robinson
Michael Jackson

Thematic Information Service
Natural Environment Research Council
Holbrook House
Station Road
Swindon SN1 1DE

BIOGRAPHICAL SKETCHES

Dr Gary Robinson leads the digital cartography research section of the Thematic Information Service, NERC and is the convenor of the British Cartographic Society (BCS) special interest group in digital cartography. His interests cover all aspects of digital cartography, computing, expert systems (especially the man-machine interface), Geo Information Systems and error analysis. He graduated from Imperial College, London with BSc and PhD degrees in Physics and Astronomy in 1976 and 1981 respectively.

Dr Mike Jackson is Head of the Thematic Information Services (TIS) NERC. He worked with UK Department of the Environment and Atomic Energy Research Establishment, Harwell before joining NERC as head of the Experimental Cartography Unit (now incorporated within TIS) in 1980. He is Chairman of the European Association of Remote Sensing Laboratories Working Group on Integrated Geo-information Systems and Principal Investigator to NASA for the Landsat 4/5 LIDQA Programme as well as sitting on UK and International Committees related to digital cartography and remote sensing. His research interests relate particularly to the integration of digital cartography and remote sensing data and technology and the development of a knowledge-based integrated GIS.

Both authors are current members of BCS Council.

ABSTRACT

Increasingly spatial data is being collected, analysed, modelled and manipulated using digital computers. The users of such systems include planners, social and environmental scientists, businessmen and engineers. Where they once had recourse only to printed maps and plans they are now able to combine their data sets with spatial data and in the process generate new maps tailored to their specific needs. However, such people rarely have either the cartographic expertise or access to guidance from a professional cartographer. The result can be poorly presented maps and graphic material which fail to impart the information intended, or worse, mislead. What is required therefore is a means whereby the computer can itself act as a friendly cartographic 'advisor'. How can this be achieved? The paper presents one possible solution, using a so-called 'expert', or Intelligent Knowledge-Based System. It outlines the nature of such systems, discusses other potential applications within the field of cartography and describes a collaborative project in the United Kingdom between the Thematic Information Services of the Natural Environment Research Council, Aberdeen and Glasgow Universities and Kingston Polytechnic which involves the investigation and eventual implementation of such a system.

INTRODUCTION

The past decade has seen a rapid rise in the use of computer-based tools such as data-bases and mapping packages by a wide range of people, from scientists and engineers through to planners and administrators. Whereas these people originally had access to fairly limited data sets which they manipulated and displayed by relatively simple programs they are now able to combine disparate data sets such as digital vector and raster data with non-spatial information such as statistics. This all too often results in a product which fails to impart the information intended, or worse still misleads (Carter & Meehan 1984). Some may argue that 'bad workmen blame their tools', but if the workmen don't know any different then perhaps some responsibility should be accepted by the people who develop 'Geo-information Systems' and other mapping packages which are used by cartographically ignorant users.

A typical user rarely has the services of a professional cartographer with the necessary knowledge of map design and production techniques necessary for optimal display and presentation of data. This may be because he may not realise that one is needed or if he does then there may be insufficient cartographers or funds (or both) to go around. A partial solution is to provide cartographic training to all users. In general this is infeasible because of limited resources, cost, aptitude etc. What is therefore essential is that the user realises that a problem exists in the effective display of spatial data.

Probably the user most at risk is the one who produces maps or other graphical output for his own use or for limited circulation. There are several facets to the overall problem. Firstly the user, in designing his output, will often use an interactive graphic facility and therefore he needs to optimise the information appearing on the screen appropriate to his particular expertise. Secondly, the final product may appear on a totally different medium, e.g. paper, which leads to further problems. Thirdly the producer must take into account who the final product is intended for, and for what purpose.

These problems are ones of 'information transfer'. One solution which has been around for several years and is gaining increasing attention, particularly due to the Japanese 'Fifth Generation Project' is the so-called 'expert system' in which the knowledge and skills of one or more experts are encapsulated in the form of 'rules' which are capable of being manipulated by computers.

Do expert systems provide a suitable mechanism for solving the problems encountered in map design? To answer this question it is useful to outline what expert systems are, how they work, which fields they are currently employed in and describe a few example 'rules' involved in map design. Subsequent sections describe a collaborative project in the UK to implement an expert system-based map design system.

EXPERT SYSTEMS

Definition

A formal definition of an expert system, approved by the British Computer Society Specialist Group on Expert Systems (Naylor 1983) is:

> 'An expert system is regarded as the embodiment within a computer of a knowledge-based component from an expert skill in such a form that the system can offer INTELLIGENT ADVICE or take an INTELLIGENT

DECISION about a processing function. A desirable additional characteristic, which many would consider fundamental, is the capability of the system, on demand, to JUSTIFY ITS OWN LINE OF REASONING in a manner directly intelligible to the enquirer. The style adopted to attain these characteristics is 'RULE-BASED PROGRAMMING'.

How They Work

An expert system or 'rule-based program', contains three basic components:

assertions are statements of fact which are known or unknown and attain logical, numerical or textual (i.e. character string) values. Examples are:

'number_of_different_areal_symbols' (1)

'user' (2)

'map_is_confusing' (3)

rules are used to combine assertions together in such a way that the values of further assertions can be determined. For example:

'if number_of_different_areal_symbols 20 and
user = "tourist" then map_is_confusing = TRUE' (4)

the inference mechanism determines the order in which rules and assertions are processed. There are two fundamental ways of doing this which the preceding examples may be used to illustrate:

(i) To determine the consequences of showing a map containing 22 different areal symbols to a tourist the inference mechanism would apply rule (4) to assertions (1) and (2) to deduce using assertion (3) that it is probably too complicated for his use.

(ii) Conversely, one may want to know under what circumstances the map would be difficult to understand. In this situation the inference mechanism would use rule (4) to decide which assertions ((1) and (2)) are necessary to make assertion (3) (the consequence) true.

In either case the expert system would ask the user for the values of unknown assertions, unless a rule can be used instead. These approaches are known as 'forward chaining' and 'backward chaining' respectively. Which is used depends on whether the application concerned is predominantly 'data-driven' or 'goal-driven'.

Characteristics

The collection of assertions linked by rules resembles a network, or to use a better analogy a 'directed graph'. The way in which the inference mechanism navigates this graph varies from system to system and is the major factor governing the behaviour of individual expert systems.

Expert systems differ from ordinary computer programs in that the rules and assertions are treated as data, rather than being 'hardwired' or built into the code. This enables several different 'topics' to be accessed by a single inference mechanism. Also, most systems are able to dynamically update their rule-bases as new information becomes available, for example during an inter-active session with a user.

An additional feature of backward chaining systems is their ability to list the path currently being traversed in the rule-base, giving the impression of justifying the line of reasoning.

Fields Currently Using Expert Systems

The number of fields in which expert systems are currently employed is rapidly increasing. The originals chemistry, geology and medicine are still the major areas though, with the chemical analysis system DENDRAL reputedly being the oldest, dating back to 1965. Other well known expert systems are (Michie 1979, Hayes-Roth et al 1983): MYCIN which is used to diagnose medical ailments (specifically relating to blood disorders); PUFF (based on MYCIN) is used to diagnose breathing problems; and PROSPECTOR which assists in the location and evaluation of potential mineral ore deposits.

Expert Systems in Cartography

Some areas of cartography and digital mapping where expert systems could be of benefit are:

Manual and Automated Map Design. The knowledge of map design, in the form of rules and heuristics seems reasonably well suited to expert systems and forms the basis of this paper. Considerable advances have been made in the area of automatic name placement (Freeman & Ahn 1984).

Digital Data-base/User Interface. The rapidly increasing complexity of systems, diversity of data sets, usage and the increasing knowledge required of their users means that this area is a prime candidate for exploitation by expert systems. Several workers are already active in this area (e.g. Peuquet 1983, Bouille 1983).

Cartographic Education and Training. Computers are playing an increasing role as teaching aids at all levels of education. Although this is looked upon with mixed feelings it is clear that expert systems will have a major impact in this field, and the teaching of cartography is no exception.

Spatial Data Error-train Analysis. Systems holding rules about these could assist a user in overlaying and combining different data sets by indicating likely sources of error and which statistical techniques are most appropriate.

Data Capture and Storage Standards. Since expert systems can be used as computer-based repositories of facts they are ideal for on-line storage of data standards and digitising conventions, with the possibility of allowing inconsistencies between different standards to be weeded out.

Data Format and Transfer Standards. Expert systems offer a way of holding different digital data formats in a self-contained manner and could therefore provide a consistent mechanism for data transfer and conversion operations.

Replacing Cartographers. As expert systems are intended to substitute for human experts then conceivably they could replace all cartographers. However, this is not really practical since:

(i) It is unlikely that any human expert could (or would) formalise all his knowledge in such a form that a computer could take over his job (although there is one recorded instance of this happening, Feigenbaum & McCorduck 1983).

(ii) Cartography, like most subjects is evolving, especially in the light of new technological developments. Cartographers would therefore still be required to undertake research work, even if it was only to up-date cartographic expert systems.

(iii) Most importantly, cartography involves artistic elements which would be (as yet) impossible to capture in an automated system.

MAP-AID: A UK PROJECT

The subject of map design is very complex, particularly in view of the subjective decisions made in determining what constitutes a 'good' map. However, it is essential that an effort is made to solve some of the problems outlined in the introduction, which have arisen from the rapid increase of computing power, cheaper colour output devices and wider access to spatial data banks and other data sets.

Project Origins and Progress

The Thematic Information Services (TIS) of the Natural Environment Research Council (NERC) incorporates the Experimental Cartography Unit (ECU) which was a pioneer in digital cartography. Over the years the Unit acquired several hundreds of megabytes of structured digital map data contained within its Mk 1 database (Jackson et al 1983). In 1981 TIS became the centre for image analysis for NERC and rapidly acquired large quantities of airborne and satellite remotely sensed data. To exploit these data sets powerful interactive image analysis systems were purchased. The equipment includes one I^2S Model 70 and two I^2S Model 75 processors with host computers plus video and digitiser table input and colour raster output. Software was developed to allow the integration of the map and image data (Jackson 1984). In 1984 with the growing acceptance of these facilities as a standard research tool for geologists, ecologists, etc the process of providing local facilities was commenced and the transfer of the facilities into the NERC Computing Services began.

The above developments, whilst welcomed by users, increased the dangers indicated in the introduction, that is, of poor visual representation of data in the interactive and final presentation stages of a project with the consequent loss of information. The impetus was thus provided to transfer the cartographer's map planning, design and production skills to the user through the embodiment of his knowledge in an expert system.

The multi-disciplinary nature of the TIS staff with professional cartographers, geographers, mathematicians, physicists, etc, many with considerable computing experience, provided a sound basis for the research. In addition specific theoretical and practical cartographic expertise was sought from UK universities and the project team now includes J Keates of Glasgow University and M Woods of Aberdeen University.

Finally, independent computing expertise was incorporated through the participation of scientists at Kingston Polytechnic (G Wilkinson and P Fisher).

The Project

The early stages of the project have been concerned with:

Design and Planning. A project plan involving the authors and above identified collaborators in a 3-year programme of work has been prepared. The first significant working test system is planned for mid-1985. The use of the System Designers Limited Poplog program development environment which includes POP11, Prolog and a limited LISP compiler is proposed, supplemented by code written in FORTRAN and C. The GIMMS mapping package (Waugh 1980) is to be used for graphical output. More extensive graphic, map and image options will be added later. The computing environment comprises a 4 Mbyte VAX 11/750 running under VMS supported by an I^2S Model 75 image analysis system, a Britton-Lee IDM relational database machine and two ICL PERQ microcomputers with associated colour raster displays.

Increasing the Group's Knowledge of Existing Expert System Tools. A preliminary evaluation of languages and packages has been carried out to assess their suitability for map design applications. The SAGE package (SPL Int.1984) was selected for an in-depth three month assessment using a trial map data set. The necessarily limited purpose of this study was to identify suitable symbolisation for the different features to be included in the trial map. Conventions were incorporated into the rule base together with rules covering the relative importance of different features and what constituted acceptable combinations of symbols. The user then had the choice of picking from an available range of symbols, with the system providing defaults or warning when conventions were broken or poor combinations selected. The experience gained from this limited exercise suggested that many of the available commercial systems available were going to be too limiting for the complex task of map design. In particular groups of symbols or other features need to be selected together in parallel if frustrating and time consuming iterations are to be avoided. This parallel mode of symbol or feature presentation and selection will require powerful real-time graphics facilities. The hardware is available within low cost systems but software development is necessary. Because many of the decision making processes occur in parallel or near-parelle1 mode Prolog appears to be the best logic programming language for future developments of the MAP-AID system as it naturally lends itself to implementation on parallel processing systems (Clocksin & Mellish 1979).

Systems Design. The MAP-AID system (figure 1) is composed of three elements linked at the system level: the expert system; data-base system(s) and graphics package(s). Communication between these is via inter-process links (provided by most computer operating systems such as VMS and UNIX) on the same computer system, physical links (possibly over a wide-area network) between different computer systems, or a combination of both.

(i) Expert System. The expert system divides into four logical sections: the 'core' contains the map design rule-base and other information held as rules in a knowledge-base; the user module through which the user controls the entire system and interacts with the knowledge base; a set of data-base system modules (one per data-base) and a set of graphics package modules (also one per package).

(a) The Core. This contains the map design knowledge-base and is independent of the format of the user, data-base system(s) and graphics package(s) modules. To allow communication between the

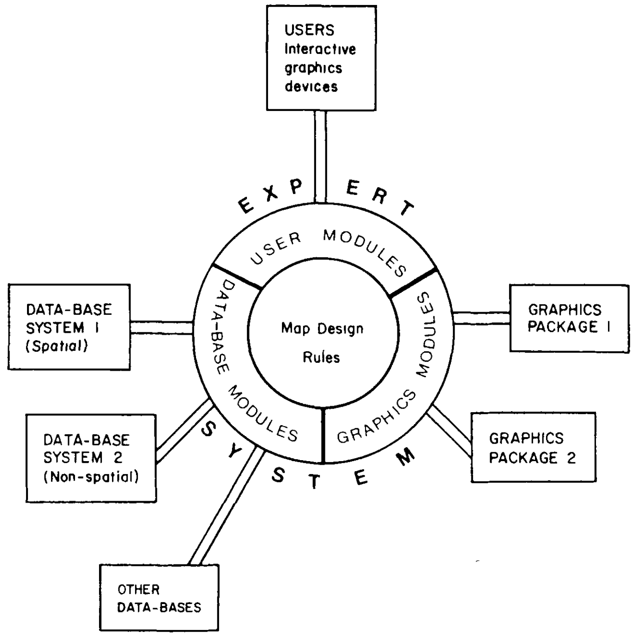

Figure 1 STRUCTURE OF THE MAP-AID EXPERT-SYSTEM

core and the surrounding modules, a uniform set of standard interface procedures is defined, each of which must have a corresponding implementation in every user, data-base and graphics package module.

For example, suppose a particular design rule needed to know the scale of an input map. The core would invoke a sub-goal of the form:

'return the scale of the data set in the selected data-base'.

The selected data-base would be known (from a previous sub-goal or query) and the next sub-goal to be tested would be triggered in the rule-base of the relevant data-base module.

(b) User Module(s). The user module(s) operates as an intelligent or 'user-friendly' interface by converting what the user types in at a keyboard (or enters via other input devices) into a format suitable for processing by the core. Similarly, when the core requires information to be supplied by the user, because for example the data-base system is incapable of giving it, the user module will convert the query into an English-style question(s).

(c) Data-Base Module(s). To satisfy any queries asked by the core each module must be capable of either:

generating a query in the syntax of the relevant data-base system, issuing it, receiving the response, decoding it and then reacting accordingly.

or returning an 'unknown' response, in which event the system could ask the user to supply the answer via the user module (the user module resembles a data-base module in this respect).

The procedures that create, issue and receive data-base queries are written in 'C' or FORTRAN 77 for efficiency.

(d) Graphics Package Module(s). In a manner analagous to the data-base modules, the core has a uniform set of defined graphical procedures. Each graphical package module has a corresponding implementation, if applicable.

To enable the map design process to operate efficiently requires the core to interrogate the graphics packages as to what line styles, colours area patterns etc are available. This is similar to the approach used in graphics systems such as Graphical Kernel System (GKS). GKS appears to be a useful starting point for defining the set of uniform procedures necessary in defining the interfaces between the core and the graphics package module(s). A GKS-based graphics package could even be used as one of the mapping systems. A similar approach could be used in the definition of the standard data-base procedures.

Each module may be viewed as comprising: an interface to the external system (the user, data-base or graphics package); internal rules to govern the functions performed by the module; and an interface to allow communication with the core.

(ii) Data Base System. The data-base(s) hold the spatial and associated non-spatial data in the form most suitable for the application in hand, and can be proprietry systems, specialised or house systems or even a combination thereof.
An essential feature of any data-base selected is that it has a well-defined interface at the system level, i.e. what the user would see if it were used interactively.

(iii) Graphics Package. Similar remarks about the data-base system apply to the graphics package. Communication between the data-base and the graphics package (via the expert system) is slightly easier to implement if the latter has some date-base capabilities, as GIMMS has.

Rule Identification. The rule identification stage raises fundamental questions in cartography and constantly tempts one to be side-tracked from the initial objectives. Thus, a simple sounding rule as expressed by the design cartographer such as 'don't use too many strong colours on complicated data' leads one into questions of perception of colour, measurement of colour, spatial interaction of colour etc. In trying to provide guidance in the use of colour the following parameters (and more) may need to be considered:

(a) the number of classes to be depicted (and number of colours to be used).

(b) the total area of the final map.

(c) the amount of overprinting (by lines and text)

(d) the mean size and variance of polygons

(e) the autocorrelation function

(f) the number and ratio of classes to sub-classes

(g) the degree of contortedness of the polygons

(h) whether the map contains polygons of discrete variables (e.g. land use) or continuous variables (e.g. contours)

(i) whether the map is intended for experts

(j) the relationship to conventions, etc

Rather than proliferate rules for the rule-base initial selection of the more general and powerful is sought even where full quantification cannot be achieved. i.e. the system will prompt the user for inputs on a nominal scale based on subjective assessment where precise values cannot yet be calculated.

Model testing. The cartographic model and rule-base are being interactively developed with theoretical evaluation of the model. The model is stepped-through conceptually to test for integrity and shortcomings. This process is useful in the early learning stages and allows more advanced concepts to be tested than in the implementation stage. The process is itself interactive with increasingly complex models being evaluated theoretically before test implementation. Results of the implementation using Poplog will be reported in a forthcoming paper.

CONCLUSIONS

Expert systems can play a useful role in many aspects of cartography and more work should be undertaken in this direction by cartographers and computer scientists. This is particularly important since it is pointed out (Taylor 1984) that unless cartographers get more involved in the 'New Cartography' they will be supplanted by computer graphic designers with inadequate knowledge of map design techniques but who will be responsible for producing the output for systems such as Telidon.

The direction of computer assisted cartography has over the years tended to move away from the original aim of producing 'look alike' versions of manually produced maps and instead focus on the information contained in the data. Many reasons have been presented for this, the most quoted being that computers are inefficient at the simple replication of manual techniques but are better at other tasks. The application of expert systems to map design may lead to this trend coming full circle.

ACKNOWLEDGMENTS

This paper is published with the permission of the Director, Natural Environment Research Council Scientific Services, United Kingdom. Our thanks go to Julie Swann for patiently typing many drafts and to Darrell Smith for producing the figure.

REFERENCES

Boulle F, 1983, A Structured Expert System for Cartography Based on the Hypergraph-based Data Structure, Proc 6th Int Symp on Automated Cartography, Vol II, pp 202-210.

Carter J R and Meeham G B, 1984, Austra Carto One, Perth, pp 259-276.

Clocksin W and Mellish C S, 1981, Programming in Prolog, Berlin, Springer-Verlag.

Feigenbaum E A and McCorduck P, 1983, The Fifth Generation, London, Pan Books.

Freeman H, Ahn J, 1984, Autonap - An Expert System for Automatic Map Name Placement, Proc Int Symp on Spatial Data Handling, Zurich, Vol II, pp 544-569.

Hayes-Roth F, Waterman D A, Lenat D B, 1983, eds, Building Expert Systems, Reading, Massachusetts, Addison-Wesley.

Jackson M J, Bell S M B, Diaz B M, 1983, Geographical Data-base Developments in NERC Scientific Services, Cartographica, Vol 20, No 3. pp 55-68.

Jackson M J, 1984, A Methodology for an Integrated Spatial Data Model, Proc of EARSeL Symposium on Integrative Approaches in Remote Sensing, Guildford UK, 1984.

Michie D, 1979, ed, Expert Systems in the Micro-electronic Age, Edinburgh University Press.

Naylor C M, 1983, Build Your Own Expert System, England, Sigma Technical Press.

Peuquet D, 1983, The Application of Artificial Intelligence Techniques to Very Large Geographical Data-bases. Proc 6th Int Symp on Automated Cartography, Vol I, pp 202-210.

SPL International, 1984, SAGE User Manual, SPL International, Abingdon, England, UK.

Systems Designers, 1983, Poplog User Guide, System Designers Ltd, Camberley, Surrey, England, UK.

Taylor D R F, 1984, Vol I, 12th Int Conf ICA, Perth, pp 455-467.

Waugh T C, 1980, The Development of the GIMMS Computer Mapping System, in Taylor D R F, 1980, ed, The Computer in Contemporary Cartography. New York, Wiley, pp 219-234.