EXPERT SYSTEMS IN STATISTICS

D.J. Hand

Biometrics Unit
Institute of Psychiatry
De Crespigny Park
London SE5 8AF

ABSTRACT

Statistical expert systems are attracting increasing attention as a possible way to alleviate the shortage of expert consultant statisticians. This paper summarises the requirements of such systems, showing how the demands of data analysis are different from those of other fields, and describes some recent work.

INTRODUCTION

A great deal of attention and effort has been focussed on developing expert systems for the medical domain. Reasons for this include the basically ill-structured nature of the diagnosis/treatment problem, the time it takes to train experts in the area, the scarcity of such experts, and the expense of employing such experts. Of course, many other fields are also experiencing a rapid growth in expert systems development, but one in particular has many similarities to medicine. This is the field of statistics.

Similarities to medicine include the ill-structured nature of choice (though here it is choice of statistical technique rather than identifying the afflicting disease), the fact that it typically takes around ten years (see Hand, [5]) to train an expert consultant statistician, the consequent fact that such experts charge fees comparable with doctors in private practice, and the simple fact that the demand for such experts becomes ever greater in all aspects of an increasingly complex society. Examples of the latter include quality control in industrial production, the design and analysis of drug trials, interpreting data for administrators and governments, the analysis of traffic flow in urban planning, and the assessment of alternative agricultural practices.

However, although some of the motivating reasons for developing statistical expert systems may parallel those for medical expert systems, there are also major differences between the two domains and their requirements. One very important one, of which we will say more later, is that a statistical expert system will typically not only interact with the user of the system but will also interact directly with the data. It is as if a medical system itself had instruments plugged into the patient and itself injected the patient with various medications.

This paper explores the requirements of statistical expert sytems and surveys recent work. We start by considering what people will expect of statistical expert systems.

WHAT WILL STATISTICAL EXPERT SYSTEMS DO?

An obvious place to begin is by asking what statistical expert systems will be used for. We can best answer this question by taking a step backwards and asking what consultant statisticians do. The answer is that they talk to clients and:

(a) refine the research objectives and questions,
(b) choose appropriate statistical techniques to address the questions,
(c) apply the techniques,
(d) interpret the results.

The linear nature of a-b-c-d is an artefact of the written mode of communication — and has probably been the cause of much of the misunderstanding of the aims and scope of statistics. First, statistics as a discipline is a structure, not merely a collection of recipes. No statistical technique stands in isolation, but there are complex interrelationships between statistical tools. Second, statistical practice is an iterative and cyclic process: research questions are ofen refined in terms of a suggested technique, attempted application may cause alternatives choices to be tried, thoughts on interpreting the results to a non-statistical audience might cause one to change one's mind about appropriate techniques.

Bearing these points in mind, the fact is that when considering building statistical expert systems we must start small. And for convenience this means that attention is often focussed on just one of the above four areas.

For (a) the most obvious choice of an area in which expert systems could be extremely useful is some aspect of experimental design: what is the most efficient way to organise and structure one's research so that highly accurate answers can be obtained for least cost? Are specific comparisons between different treatment groups of interest or does one wish to make overall statements? A simple example of this is in sample size determination: how many subjects would a clinical trial require in order to reveal whether or not the drug was effective?

All too often a researcher has already collected the data before seeking statistical advice. This can cause problems (in the worst case it can mean that the questions the researcher would like to answer are unanswerable). In other cases it means that the consultant has to tease out the structure and objectives from the researcher. A simulated example of how an expert system designed for this role might go about it is given in Figure 1, which shows the beginning of a session. This session clearly requires a certain level of technical statistical expertise from the user. This is an issue we return to below.

FIGURE 1:A simulated example of determining the structure of a researcher's experiment so that an appropriate multivariate analysis of variance can be performed.

1. HOW MANY BETWEEN SUBJECTS FACTORS ARE THERE?
   2
2. GIVE THE NAME OF FACTOR 1.
   sex
3. GIVE THE NAME OF FACTOR 2.
   treatment
4. HOW MANY LEVELS DOES SEX HAVE?*
   2
5. HOW MANY LEVELS DOES TREATMENT HAVE?
   4
6. ARE YOU INTERESTED IN PARTICULAR CONTRASTS ON THE LEVELS OF TREATMENT?
   yes
7. PLEASE GIVE THE CONTRASTS
   1 -1 0
   1 0 1

   . . . . . .

24. HOW MANY VARIABLES ARE MEASURED ON EACH SUBJECT?
   6
25. HOW MANY FACTORS DESCRIBE THE WITHIN SUBJECTS STRUCTURE?

   . . . . . .

*a system with a little substantive knowledge could appear more intelligent.

For (b) one might have in mind a global system, which contained expertise about a wide range of very different methods, so that use of it might culminate in the recommendation to use a broad class of techniques such as descriminant analysis or correspondence analysis, for example. Or one might have in mind a narrower system, which chose one method from the class of two-sample non-parametric tests, or one method from the collection of twenty-two tests of equality of proportions listed in [10].

For (c) one knows which method is to be used and now has the task of applying it: does the data satisfy any assumptions that the method makes? Are there outliers? Are transformations needed? Is there missing data? How many factors should be included in a model? Can these interactions be neglected? And so on.

(d) presents a special class of problem, on which there has been relatively little work. Interpretation of the results of a statistical analysis is a translation from the technical language of mathematics and statistics into the substantive language of the researcher's speciality. This last stage makes it particularly clear why an effective consultant statistician must have some knowledge of his client's specialist domain.

This is a convenient point to return to a remark made at the beginning of the discussion of the four stages above. We said there that statistical consultants 'talk to clients' in order to carry out tasks (a) to (d). This means that we must address the question 'to whom will statistical expert systems talk?'. The question is important because there is a continum of possible users of such a system, at the extremes of which are:

(a) professional statisticians
(b) statistically naive researchers.

(The term 'naive' here is not intended to be deprecatory; it is merely meant to reflect the fact that although they may be internationally leading researchers within their own discipline, they have limited statistical knowledge — which is why they are seeking statistical advice in the first place.)

In most other application areas of expert system technology there is an implicit assumption that one or other of these extremes is the primary target population for use of the system. For example, most medical applications assume that an expert, or at least someone familiar with the technical medical terminology, will use the system.

In most cases there is little question of a choice of which type of population the system should be aimed at.

In statistics, however, this choice does exist, and the requirements of the two extreme types of population are very different. Statistically naive researchers will typically wish to proceed slowly, with much explanation from the system regarding why it is doing what it is doing. They may require the facility to be shown examples of some statistical phenomenon or concept and they may not be certain of the appropriate answer to any particular question. Above all, they will expect the system to guide them. Expert statisticians, on the other hand, have complementary requirements. They will not wish to be bored by lengthy and tedious explanations of something they already know. Rather, they will simply want to be reminded of a detail they have forgotten, or prevented from making some oversight. In brief, they will use the system in the same way as a medical consultant will use a medical expert system — to provide a second opinion. And, in particular, the use of the system as a guide will be much less important: they must have the upper hand and may want to insist the system takes some course of action, even though the system recommends against it.

Two potential types of user were identified above, but we did remark that they were extremes of a continuum. In general, it is necessary to consider where on this continuum it is most effective to aim the system at — or, better still, the system must be made adaptable, so that anyone can use it to advantage.

STATISTICAL EXPERTISE

If one looks at current statistical software, one sees that it contains *arithmetic* and *algebraic* expertise. That is, the user provides the number and the system carries out the matrix manipulations, the numerical optimisations, and so on, and produces an answer. More complex packages (such as SPSSX, SAS GENSTAT, GLIM, etc) require that the user possesses some *programming* expertise in order to be able to use them. None of these systems, however, contain *statistical* expertise. For example, one can provide such systems with integer codes indicating the religions of the people within a town, and then ask for the sum of these numbers. The programs will duly give you the sum — even though it is, of course, meaningless. Statistical expertise is knowing when it is appropriate to use a technique, when it should not be used, what one must do to the data to make it legitimate to use the technique, and how to draw sensible conclusions from the application of the technique. Existing programs rely on the user possessing that expertise.

Some programs (the MANOVA program in SPSSX, for example) are particularly vulnerable to this kind of criticism by very virtue of the fact they they have been made easy to use. Easy use implies easy misuse, and the problems typically tackled by MANOVA require a good grasp of the theoretical background — and of the way this particular program tackles things — before one can be confident of avoiding mistakes. Other systems (eg. MULTIVARIANCE and GENSTAT) are vulnerable to the opposite kind of criticism: these are not very easy to use. One has to be something of an expert oneself to use them at all. Since only relative experts will use them, there is less chance of them being misused. (Smith Lee, and Hand [27] see below, have more to say about MULTIVARIANCE in this context).

There is clearly something of a dilemma here, and the way out seems to be via expert systems technology. Such an approach will permit us to apply statistical techniques readily (the system will show us how) while at the same time protecting us from making mistakes in application. That, at least, is the idea.

The fact that statistical software exists — that computers are the primary tool used in exploring data — leads to one of the major differences between medical and statistical expert systems. This difference has already been referred to above. In the former, the system collects its information via the user of the system. In the latter the system will get its information both from the user and directly from the data. The user will supply (probably in response to queries from the system) such information as that a variable has a nominal

measurement scale, or that interest lies chiefly in patterns of change over time. The data themselves will reveal that one observation is an outlier or that there is a nonlinear relationship between variables. A more sophisticated system will detect that some variable only takes values 0, 1, or 2 and, alerted by this, will ask the user if it is in fact an interval variable or if it is really merely nominal. It is important to note here that the 'data' is part of the presenting problem. Thus a medical system may also extract information relevant to a problem from a database of records of similar patients, but this represents a constant source of knowledge within the expert system. It serves an analogous role to the system's rule base. In contrast, the information that the statistical expert system extracts from the data presented to it is part of this single problem, and is not part of the system's background knowledge.

This dual source lends an extra complication to statistical expert system design.

It will be useful here to list some of the major attributes that statistical expert systems should possess. Not all of these are common to other disciplines. (See [6] for a more detailed discussion.)

— Statistics is a young and dynamic discipline. It is growing rapidly, driven in part by the number crunching power of the computer, with major new developments regularly appearing. Some of these represent extensions of classical methods, freed from computational time constraints (eg. multivariate analysis of variance, multidimensional scaling, log-linear modelling) and others represent radically new departures, perhaps not even imagined before the power of the computer became apparent (eg. bootstrap and other resampling schemes, interactive graphics). It is thus absolutely essential that any such systems should be easy to modify and extend.

— Not only must the system be able to explain its reasoning processes (taken by some to be a defining characteristic of expert systems) but a statistical expert system must also be able to define and clarify technical terms. At the least this must consist of canned definitions, and ideally it should consist of the ability to generate examples and indicate the links and connections to other concepts. This requirement of being able to explain technical terms arises because of the potential for being used by non-experts. While an expert might require an occasional reminder (for which a brief potted definition will probably be adequate) for a non-expert the capacity for detailed explanation is essential. It is clear that, when being used by a non-expert, there are elements of the tutorial role.

— It is rare that a consultation involves a single question leading to a recommendation to use a single technique (unless the 'technique' is a method of analysis, implicitly involving many decisions). Much more common is that several research questions are posed. Typically these are related in some hierarchical structure (if the anser to A was 'yes' then ask B). Often the structure is dynamic, changing as more is discovered about the data. An expert system must not only address all the relevant questions but, more important and more difficult, it must know when no further questions remain. One way in which this might be accomplished is by having the system present a summary of its conclusions, which the researcher can study to see if everything is covered.

— The system must be able to handle misunderstandings of technical terms by the user. It must be able to backtrack and correct any conclusions it has drawn. Such a pattern stands out distinctively in analyses of consultancy sessions [15]. It emphasises the nonlinearity of stages a-b-c-d referred to above. Detecting misunderstanding is a difficult problem. A contradiction in the information the user has provided might mean there is some misunderstanding but it need not. (The researcher might express an interest in all contrasts between groups and in certain specific contrasts — it's simply that he has a greater interest in the specific ones.) Rectifying a detected misunderstanding is also not an easy problem, although often the researcher becomes aware of an increasing tension between his conceptual model of what is going on in the statistics and what the numerical results seem to be saying.

— Some authors maintain that expert systems are best modelled on humans, although not all hold this view (see [9] for a discussion). While a statistical expert system following this principle might appeal to an expert user, it is not clear that it will appeal to a statistically naive client — after all, sometimes even a human statistician's conceptualisations do not appeal to such clients!

REPRESENTING STATISTICAL KNOWLEDGE

Knowledge representation is, of course, often regarded as the fundamental issue in artificial intelligence research. Representing statistical knowledge is a particularly challenging task because of the many different kinds of knowledge that have to be represented. Thisted [25] describes the complete expertise of an expert data analyst as encompassing such areas as "mathematical statistics; techniques of graphical display and analysis; rules of thumb for judging the importance of apparent indications; copious examples of bad or misleading analyses (coupled with a catalogue of common errors made by novices, the avoidance of which is essential to respectability); methods, both ad hoc and those thoroughly grounded in theory, for basic

operations such as smoothing, assessment of variability, and model building; and — perhaps most important — knowledge of how and when to elicit specific subject matter information from a scientific collaborator." He goes on to add that "it is unlikely that an expert system could be built today that could incorporate this vast range of different kinds of knowledge".

The problem is compounded by the fact that very little research has focussed attention on how statisticians themselves represent these different kinds of knowledge, although some small areas have been examined briefly (see, for example, Hand, [4]). Of course, one can take advantage of work done in areas which parallel any of the areas mentioned above. For example, the question of choice of statistical technique has parallels with choice in other domains — such as medical diagnosis. This means that production systems are an obvious starting point for systems concerned with selecting methods. Having said that, they are not the only possible approach, and are not necessarily the best. Another contender is the classical decision tree approach, although its short-comings are well known. Hakong and Hickman [21] describe an alternative based on intersecting sets of properties of techniques.

Other standard knowledge representations also seem appropriate for some sub-domains in the above. Ellman [24] for examples, uses 'frame-like structures' to build a system which can reason about statistical computations. Frames or scripts appeal to the present author as a way of thinking about particular analytic techniques. This leads us to the notion of strategy in statistics [18]. In this article, by a 'statistical strategy' we mean a formal description of the choices, actions, and decisions to be made whilst using statistical methods in the course of a study. Since, as we have already noted, statistical consultancy work does not simply consist of a list of things to be tried, one after the other, we note that "a strategy is not simply a sequence of actions but a prescription for choosing an action in response to any possible external event such as an outcome of a test, . . ., or the output of a complicated computational procedure." [8]. It is not just a recipe — it also says what to do if the pan boils over. Stimulated partly by the interest in statistical expert systems researchers have focussed attention on developing — or at least attempting to make explicit — the strategies that statisticians follow when performing an analysis. A major consequence of this is that the importance of the concept of a statistical strategy in its own right is now beginning to be recognised. Apart from their obvious crucial relevance to expert systems, strategies are important to clarify what is known and what needs to be known (attempting to write down an explicit strategy soon shows where the important holes are) and to aid in educating statistical consultants.

It is also perhaps just worth remarking that building statistical expert systems is particularly challenging because of the multiplicity of knowledge representations which must be used. Not only must one represent statistical expertise — using rule-based systems, frames, scripts, logics, or whatever seems most appropriate — but one must also represent the mathematical structure of statistics itself. For example, it is necessary to have some kind of algebraic representation for possible crossings and nestings in an experiement.

SOME SYSTEMS

Strategy

The first observation to be made about systems for guiding the application of a chosen technique is that even something so simple as a checklist can be extremely effective in ensuring that no steps are omitted and that no departures from assumptions are overlooked. However, checklists cannot themselves look directly at the data, and so provide only a sort of crutch to assist the researcher. Neither do checklists guide a researcher in understanding what is going on.

Perhaps the most famous system for guiding statistical analysis is the REX demonstration system built at Bell Laboratories [12, 19]. REX "allows a novice to use advanced regression techniques safely by systematically checking the assumptions of the techniques. It provides guidance to what tests need to be done and when, interpretation of the results of tests and plots, and instruction in statistical concepts." From a statistician's viewpoint REX is especially interesting because it embodies a particular view of data analysis, namely that data analysis consists of successively detecting and resolving problems until none remain. (Pregibon [18] describes data analysis as ". . . an iterative process where one successively tries to remove structure from the residuals." And Gale, [11] "After the initial knowledge is acquired, a data analysis is in one of two states: trying to detect a possible problem, or trying to fix a detected problem. When no more problems can be detected, the analysis is complete.") Not all statisticians would accept this philosophy. It seems particularly alien to work in the behavioural sciences. This, of course, prompts the observation that the advent of statistical expert systems is forcing us to consider the foundations of approaches to statistical consultancy, not something which has been closely examined before.

Inspired by REX, the team at Bell Laboratories are now developing 'Student', [11]. Student is a program which monitors how a consultant statistician carries out analyses in a chosen domain (the analysis being performed on the statistical system S) and which then synthesises a strategy for those kinds of analyses. That

is, it builds systems like REX. Ultimately Student should be able to function as a tutor, a consultant, or as a student. In Student, problems which must be overcome before the method can be validly applied (that is, assumptions of the method of analysis being modelled) are represented by frames, which are organised into a tree, with internal nodes representing classes of contraventions of assumptions and leaf nodes representing particular contraventions. For example, an internal node might be 'problems with residuals' with a dependent leaf node 'heteroscedasticity'. Each leaf of this assumptions tree is a root of a tree of possible actions to make the assumptions true (eg. transform the data). Following the earlier work on REX, and in keeping with their general philosophy of data analysis, Student places much emphasis on graphical plots.

Student is clearly a very ambitious project and whether it suceeds or not its results will have important implications for both statistical consultancy practice and for knowledge engineering research in general.

REX and Student use knowledge engineering and artificial intelligence approaches, but it is not clear that this is necessary. We have already remarked that emulating the way humans do things might not be the best approach. This is illustrated by a series of commercially available user-friendly programs built by G.B. Wetherill's team at the University of Kent which use more conventional methods with great success (eg [20]).

Selecting techniques

Hand [22] describes two systems for selecting appropriate statistical tools. The first is based on classic decision tree methodologies and was intended to serve as a baseline against which the performance of later systems could be assessed. The system was tested on both consultant statisticians and on their clients. Its chief disadvantages were found to be the difficulty of consistently modifying it (well known to be a weakness of decision trees), its poor ability to explain technical terms (which used a traditional lexical approach), its inability to cater for a 'don't know' response, and occasional conceptual mismatches between the system and researcher (eg. regarding two groups as two groups or as two different scores on a nominal variable). One very striking point which emerged forcefully during the construction of this system, was how very domain dependent were the various classes of statistical technique. Not only is there emphasis on different statistical tools in different ground domains (eg. emphasis on nonparametric methods in the behavioural sciences which is not reflected in engineering disciplines), but a single tool may be used in different ways — to answer different kinds of questions. Indeed one also finds the same name being used for entirely different statistical techniques, since they are used in different kinds of areas (for example, reliability theory in engineering refers to how reliable a machine is — whether or not it will break down — whereas in the behavioural sciences it refers to how reliable a test is — if it is repeated, will one get the same results?)

The second system described in Hand [22] was intended to overcome the difficulties encountered by the first. It was based on a structured production system, regarding the collection of statistical techniques as consequents of a base level of rules and with clarifications and simplifying concepts forming the rules feeding them. Both this and the decision tree approach clearly used a basically cookbook approach and thus were fundamentally limited in the kind of manipulations about statistical knowledge which could be performed using them. The development of this system led to the suspicion that the deep deductive capacity of production systems is, in general, an irrelevant and misleading property for statistical expert systems concerned with selecting methods. Instead, a broad shallow set of conditions, some of which are more important than others, which must be satisfied for a system to be appropriate and legitimate, seems to better represent the way humans view things. (For example, when asked 'why do you recommend this?' a human does not embark upon a deep chaining explanation, but instead picks out the few most important points and expands by giving additional, less important points).

Other work on choice of technique includes that of O'Keefe [23] on choice of non-parametric technique and Hakong and Hickman [21]. The latter work has already been referred to above, and is interesting because it does not use a production system approach. In a sense this might be seen as working in the opposite direction from the system in Hand [22], in that it lets the conditions (assumptions, restrictions) drive the choice process, rather than the methods themselves.

Interfaces

A vast amount of effort has been expended over the last few decades in building packages for statistical computation. Writing effective such programs is clearly not a trivial undertaking, involving experts in statistics, numerical analysis, and computer science. In view of this, researchers concerned with putting statistical (as well as arithmetic, etc) expertise into statistics programs have naturally considered the possibility of concentrating on the interface between man and existing packages. Some existing statistical programs make this easy (eg. S, which runs under Unix, and so makes it easy to switch between the module containing the statistical expertise and S, which contains the arithmetic expertise). Others, however, were built for use by the non-expert and already contain user interfaces. These interfaces, which may be an

intrinsic part of the program, not readily separated from it, constitute an obstruction in the path of the expert system interface writer. For example, the only mode of output might be in the form of formatted sheets for a printer, complete with headings, titles, and other comments. These must be stripped away and relevant numbers isolated for the new interface to use. Nonetheless, with a premium on achieving results with a minimum of investment, this is an approach which has been explored.

Smith, Lee, and Hand [27], for example, built an interface for the MULTIVARIANCE program. MULTIVARIANCE is a powerful program for generalised univariate and multivariate analysis of variance, covariance, regression, and repeated measures analysis. At one time it was very heavily used at the Institute of Psychiatry, but, as remarked above, it is a difficult program to use and required a considerable time investment by the consultant statisticians at the Institute. The interface, BUMP, interrogated the user using canned natural language questions, constructed a MULTIVARIANCE program on the basis of the responses, combined it with the data, and sent it off for processing. Several points are worth making about BUMP. First, since the target data structure had a well specified form (namely a MULTIVARIANCE program) knowledge representation was not a key issue. Secondly, it was received enthusiastically by researchers at the Institute. Thirdly, it was quite limited: it could not, for example, explore the data to see if what the user was saying matched the data, it could not caution the user about false assumptions, and it could not in any way examine the output — once it had dispatched the program for processing it was finished.

Finn and Bock [26] have also built a MULTIVARIANCE interface: MULTISTAT. This is a keyword driven system (cf. SPSS) which considerably eases the problems of correctly entering programs: MULTIVARIANCE was very much card oriented and used numerical code input.

Other work on interfaces includes that by Nelder (personal communication), who is constructing a GLIM interface.

Other work

Clayden (personal communication) is conducting further basic work on how human statisticians do their job. It is clearly important that such work as this should be done at an early stage of research on statistical expert systems: whether or not one decides to emulate the human approach it will certainly produce ideas.

Pregibon [18] argued that in building strategy systems one should concentrate on small areas. Oldford and Peters [16] have done this, building a system for exploring multicollinearity.

Sundgren [28] describes a system, CONDUCTOR, for data manipulation.

CONCLUSION

Statistical expertise is a scarce and expensive resource. Because of this there is a tendency for researchers to attempt to apply statistical techniques with an inadequate grasp of their limitations and suitability. Hooke puts it thus: "Use (of statistics) has been replaced by overuse and misuse. Regression is being used in foolish ways in the neighbourhood of almost every computer installation." Worse is to come, since the number-crunching power of computers means that more sophisticated techniques are slipping into regular use by researchers who have a weak grasp of the theory. If this trend continues, not only will research in all disciplines suffer, but statistics itself will (as it has in the past) receive the criticism which should be better directed at the misuse of its methods. Putting statistical expertise into a computer is the only way to avoid this. It is a good omen that more and more researchers are becoming interested in this exciting application of knowledge engineering work.

REFERENCES

The references below have been grouped into sections, but publications often discuss several issues, so that often they should ideally lie in more than one class.

General references

[1] Chambers, J.M. (1981) Some thoughts on expert software. Proc. Comp. Sci. and Statistics, 13th Symposium on the Interface.

[2] Chambers, J.M., Pregibon, D., and Zayas, E.R. (1981) Expert software for data analysis — an initial experiment. *Proceedings of the 43rd session of the International Statistical Institute,* Buenos Aires.

[3] Hahn, G.T. (1985) More intelligent statistical software and statistical expert systems: future directions. *The American Statistician, 39,* p1-16, with discussion.

[4] Hand, D.J. (1984) Statistical expert systems: design. *The Statistician, 33,* p351-369.

[5] Hand, D.J. (1985) The role of statistics in psychiatry. *Psychological Medicine, 15,* p471-476.

[6] Hand, D.J. (1985) Statistical expert systems: necessary attributes. *Journal of Applied Statistics, 12,* p19-27.

[7] Hand, D.J. (1985) *Artificial Intellligence and Psychiatry,* Cambridge University Press, Cambridge.

[8] Pearl, J. (1984) Heuristics: intelligent search strategies for computer problem solving. Addison-Wesley, Massachusetts.

[9] Slatter, P.E. (1985) Cognitive emulation in expert system design. *The Knowledge Engineering Review, 1, (2),* p28-40.

[10] Upton, G.J.G. (1982) A comparison of alternative tests for the 2x2 comparative trial. *Journal of the Royal Statistical Society, 145,* p86-105.

Strategy

[11] Gale, W.A. (1986) Student phase 1 — a report on work in progress. In *Artificial Intelligence and Statistics,* ed. W. Gale, Addison-Wesley.

[12] Gale, W.A. and Pregibon, D. (1982) An expert system for regression analysis. Proceedings of the 14th Symposium on the Interface, ed. Heiner, Sacher, and Wilkinson, p110-117, Springer-Verlag, New York.

[13] Hajek, P. and Havranek, T. (1982) GUHA-80 — an application of artificial intelligence to data analysis. *Pocitace a umela inteligencia, 1,* p107-134.

[14] Hajek, P. and Ivanek, J. (1982) Artificial intelligence and data analysis. *COMPSTAT-82,* Physica-Verlag, Vienna.

[15] Hand, D.J. (1985) Patterns in statistical strategy. In *Artificial Intelligence and Statistics,* ed. W. Gale, Addison-Wesley, Massachusetts.

[16] Oldford, R.W. and Peters, S.C. (1984) Building a statistical knowledge based system with mini-Mycin. Proc. of the ASA: Statistical Computing Section.

[17] Oldford, R.W. and Peters, S.C. (1985) Implementation and study of statistical strategy. In *Artificial Intelligence and Statistics,* ed. W. Gale, Addison-Wesley, Massachusetts.

[18] Pregibon, D. (1985) A do-it-yourself guide to statistical strategy. In *Artificial Intelligence and Statistics,* ed. W. Gale, Addison-Wesley, Massachusetts.

[19] Pregibon, D. and Gale, W. (1984) REX: an expert system for regression analysis. *COMPSTAT-84,* Prague, Czechoslovakia.

[20] Wetherill, G.B., Daffin, C., and Duncombe, P. (1985) A user-friendly survey analysis program. *Bulletin of the International Statistical Institute,* 45th Session, Vol. 3, Amsterdam, August, p20.4-1 to 20.4-14.

Choice

[21] Hakong L., and Hickman, F.R. (1985) Expert system techniques: an application in statistics. In *Expert Systems 85,* ed. Merry M., Cambridge University Press, Cambridge, p43-63.

[22] Hand, D.J. (1985) Choice of statistical techniques. *Bulletin of the International Statistical Institute,* Proceedings of the 45th Session, Vol. 3, Amsterdam, August, p21.1-1 to 21.1-16.

[23] O'Keefe, R. (1982) An expert system for statistics. Paper presented at the Technical Conference on Theory and Practice of Knowledge Based Systems, Brunel University.

Knowledge representation

[24] Ellman, T. (1986) Representation of statistical computations: towards expert systems with a deeper understanding of statistics. In *Artificial Intelligence and Statistics,* ed. W. Gale, Addison-Wesley, Massachusetts.

[25] Thisted, R.A. (1985) Representing statistical knowledge and search strategies for expert data analysis systems. In *Artificial Intelligence and Statistics,* ed. W. Gale, Addison-Wesley, Massachusetts.

Interfaces

[26] Finn, J.D. and Bock, R.D. (1981) *MULTISTAT/MULTIVARIANCE manual,* National Educational Resources Ince., Chicago.

[27] Smith, A.M.R., Lee, L.S., and Hand, D.J. (1983) Interactive user-friendly interfaces to statistical packages. *The Computer Journal, 26,* p199-204.

[28] Sundgren, B. (1985) How to satisfy a statistical agency's need for general survey processing programs. *Bulletin of the International Statistical Institute,* Proceedings of the 45th Session, Vol. 3, Amsterdam, August, p20.1-1 to 20.1-10.