Developments In Business Simulation & Experiential Exercises, Volume 17, 1990

STATUTOR:
An Expert System for Selecting Analytical Techniques for Analyzing Marketing Research Data

Ronald S. Rubin, University of Central Florida

ABSTRACT

STATUTOR is an expert system designed to assist marketing research students select an appropriate statistical technique for a particular research problem. This paper describes an instructional technique that enables students to use a knowledge-base as a decision aid. It seeks to automate the symbolic reasoning and the heuristic processes employed in the selection of such analytical techniques.

INTRODUCTION

In recent years, the use of computer-assisted instruction (CAI) techniques has proliferated in the marketing curriculum. In 1987, the Journal of the the Academy of Marketing Science developed a special issue on the application of microcomputers In marketing. Several articles focussed on such topics as: the current status of micro-computers in marketing programs of AACSB-accredited colleges and universities (Kurtz and Boone, 1987); the application of micros in sales manage- merit (Collins, 1987); the application of expert systems in sales management (Steinbert and Plank, 1987); the integration of micros in marketing research (Malhotra, Tashchian. and Mahmound, 1987); and a modular approach for the integration of micro applications over the entire marketing curriculum (Dyer. 1987).

Many publishers have also incorporated computer-assisted instruction programs with their texts, for example, programs accompanying basic marketing texts include: Computer-Aided Problems to Accompany Basic Marketing (McCarthy and Perreault. 1987); Boone and Kurtz Marketing Disk (Boone and Kurtz. 1989); Computer-Aided Problems (Berkowitz. Kerin, Rudelius. 1989): Computer Exercise Diskette (Evans and Berman, 1987). Texts in other marketing functional areas have also incorporated the use of the microcomputer.

Despite computer-aided programs increased usage over recent years. Kearsley, Hunter, and Seidel (1982) argue that little progress has been made toward new quality instructional strategies. It has been realized that current CAL applications of computers are often too simplistic to result in any substantive effects. At the 19B8 International Conference on Education and Technology, it was noted that the current generation of educational technology offers a great potential to improve learning, but that there is a "desperate need" for more definitive research in order to substantiate its effectiveness (Goodspeed. 1988). At present there is a great deal of interest in an area of Artificial Intelligence (Al) commonly called expert systems (ES). These are problem-solving programs that model the reasoning and decision-making of experts. According to Kearsley (19871 and Good (19871 expert systems not only hold great potential for education. but will have a larger impact on education and training than the traditional models of computer-based instruction.

On the basis of the present review of the literature. 1 decided to initiate the development of an expert system for both reaching and reinforcing subject matter in my marketing research course which is difficult to convey to students by traditional means: selecting the appropriate statistical technique for analyzing marketing data. Converting book knowledge into problem-solving skill is a major stumbling block for many students, especially in choosing an appropriate analytical technique that relies heavily on analysis, synthesis and evaluation. My experience in teaching this course has been that a number of students are either not sufficiently interested in the subject or have a great difficulty with it as stated above. These students registered for the course only because it was required. Many faculty members believe that microcomputers can be used in several ways to lessen the apprehension students have regarding data analysis. Some instructors use the microcomputer to do all the difficult calculations. Another way is to use it as a tutor, providing the student with a mode' of the decision processes involved in the selection of an appropriate statistical technique. It is this second use that I discuss in this paper.

Selecting Appropriate Analytical Methods

Some marketing research studies go no farther than summarizing data and cross tabulations; however, many studies employ more sophisticated methods. Therefore, an important area of learning for students is to answer the question: "Which statistical technique should I use to analyze my data?" The students are exposed to a bewildering variety of statistical tools for hypothesis testing, estimation, examination of the strength of relationships, and multivariate analysis. The guidelines of when to use these techniques involve number of criteria such as the scale by which variables are measured, number of variables to be analyzed, whether the observa

tions are independent, etc. Mv goal has been that the student's know when to use the techniques and also how to interpret their results. But examinations and project report results have indicated, for the most part, that students typically have difficulty synthesizing the variety of statistical tools and factors determining their appropriateness. In discussions with the students about their difficulty in learning this material two responses emerged: 1) they believe they really don't have to know the material because they would refer to their old marketing research text, or 2) they would consult with an expert. In the business community, the necessity to constantly refer to the above approaches as solutions indicates a students lack of understanding of the problem situation at hand, and a weakness in the ability to render an executive decision. It, was the second response that led me to develop an expert system that would help facilitate the learning process of synthesizing the variety of statistical techniques and the criteria for their selection. The system is based on the following questions:

1. What is the analytical technique intended to show? The possible answers are summarizing the data, determine observed differences and investigation of associations.
2. What is the scale by which the variables are measured? The possible answers are nominal, ordinal, interval, and ratio.
3. What are the number of variables to be analyzed? The possible answers are one or two or more.
4. Are some of the variables dependent on others? The possible answers are dependence analysis, and interdependence analysis.
5. How many samples are involved? The possible answers are one and two or more.
6. What is the relationship of the samples? The possible answers are independent and related.

The expert system developed. STATUTOR asks the above questions in a decision tree form and then indicates the correct answer to the various combinations of answers.

Building the Expert System STATUTOR Using ESIE (Expert System Inference Engine)

An overview of how expert systems are developed should provide a better understanding of the potential pedagogical benefits. Expert systems are typically constructed in one of several programming languages, for example. LISP, PROLOG and even PASCAL and BASIC. However, one does not have to be proficient in these languages to construct an expert system. Instead, they can be built using an expert system "shell.' A Shell" is the portion of an expert system that remains after all application-specific rules and user interfaces have been removed. The shell is a computer program, which acts as a translator between the English language and the programming language. In general, these shells are much easier t_0 learn than the programming languages; and therefore, the construction of an expert system is much more efficient. Also, most expert systems are able to keep you informed as to what it is looking for and what information it has learned. This request is readily available since the relationships are contained in the knowledge base and held as a stream of knowledge that can be readily accessed (Bauer 1987).

Several expert systems have been developed which would answer the question: "Which statistical technique should I use on my data?" However, each had limitations, which made them inappropriate for classroom use in marketing research. For example. Which Statistic? (Robertson, Jr. 19B71, developed for use in a political science class employs a classification scheme for the level of measurement that differs from S. S. Stevens (1960) which has been used in most, if not all, introductory texts in statistics and marketing research. I found the proposed classification scheme by Robertson, Jr. to be confusing and inappropriate for the task at hand. The Statistical Consultant (Sechrist 1988) uses a criteria scheme that is confusing to beginning marketing research students and details techniques that our students will never come into contact with in their careers. Therefore. I set out to develop an expert system in which the combination of expert knowledge and the interaction which takes place between the Student and the system makes the expert system a viable educational alternative. Students learn by following the reasoning of an expert and the interaction available through the system.

ESIE (pronounced easy) is available as Public Domain Disk #398. This permits many instructors with limited budgets to acquire a quality. stripped-down expert system shell which offers a coherent set of tools with which to construct many appropriate knowledge bases for educational use.

Unlike expert system shells that use an artificial intelligence language such as LISP or PRLOG. ESIE'S knowledge base consists of a goal declaration, rules (without headings or end designations), questions that elicit facts from the user, the format for the answer (outcome) of the search for the goal, a list of answers permitted for the user (if restricted), and, if desired, any beginning and ending text material. The syntax is about the simplest possible. Figure 1 shows an edited version of STATUTOR (one branch of the decision tree) to illustrate what a knowledge base in ESIE is like and how it operates. It is a three-rule knowledge base that determines the appropriate measure of central tendency.

Many text editors can be used for preparing and debugging the knowledge base; however, to run in ESIE they must produce "flat" ASCII files (no control characters). I used Displaywrite 4 and saved the text as an ASCII file with no problems.

Figures 2. 3, 4. 5 and 6 show a typical consultation dialog with ESIE. The program can be run in two modes, the consultation or GO mode and the testing or TRACE mode.

Figure 2 illustrates a typical "opening" screen used to introduce the expert system to the user. Figure 3 illustrates the results of the consultation.

In the TRACE mode, the inner workings of the inference engine are made visible as it chains through the knowledge base to select various rules. This mode is used to debug the knowledge base. The screen is divided into two window. The top window shows the user dialog, the one at the bottom indicates the rule generating the question. Figure 4 shows the internal status of the TRACE command of the expert system during a questioning session, while Figure 6 shows a typical TRACE display at the conclusion of a consultation.

Conclusion

The intention of educational software is to make difficult-to-understand material more accessible to students. Expert systems offer a relatively new way to organize data and retrieve information by allowing a computer program to handle complex reasoning. Expert system technology provides the opportunity for sharing the knowledge and decision-making skills of recognized experts in a filed with novices, such as our students. As instructional strategy, STATUTOR is a practical application of this new technology to increase fundamental understanding of difficult material and allows a new approach to teaching this material.

I believe that the construction of knowledge bases such as STATUTOR is a novel instructional process, as well as a tool. As such it holds important promises to all constituents in the learning process. For the student it affords the concrete opportunity to solve problems in an interesting and relevant way. Students in my marketing research classes were given a copy of STATUTOR to use and arrive at the appropriate techniques. Since the students used this program outside the class period, all of them did not use it as a study aid. But the students who did avail themselves and used the program seemed to have much less trouble selecting an appropriate analytical technique than those who did not use it.

For the instructor, the construction of the expert system acts as a tool to explore the acquisition, organization and utilization of knowledge. It should be pointed out that only a limited number of analytical techniques can be included in any introductory marketing research class because students can absorb only so much in a given rime frame. Texts vary considerably as to the number of techniques presented. Therefore the selection of a proper technique is dependent to a great degree upon the techniques included in a selected text plus those added by the instructor. However, if an instructor desired to lecture on several analytical techniques, which are not covered in the expert systems knowledge base, it would be relatively easy to modify and customize the expert system to include such techniques. Expert system programming techniques exploit object oriented and symbolic programming methods; that is, separating the domain knowledge from the rules of thinking about the domain. The power of this approach is that it is relatively easy to alter or extend a system; the program can include a new technique by simply inserting it into the knowledge base. So, if we want a system to keep current with changing knowledge, it is relatively easy to update the knowledge base where this information is needed. Moreover, once inferencing rules exist, minimal subsequent programming is required since the existing rules are the starting point. The instructor's expertise lies not in computer programming but in knowledge engineering; that is not in how to extract Information needed from experts and describe that knowledge formally.

STATUTOR has not solved all the problems of teaching the selection of an appropriate analytical technique, but it has enhanced current available methods to a significant degree.

REFERENCE

Bauer, Richard J. (1987), "The Why of ExpertSystems,' Business Quarterly. (Winter), 86- 88.

Berkowitz, Eric, Roger Kerin, and William Rudelius, (1989). Computer-Aided Problems, Homewood, Illinois: Irwin.

Boone, Louis and David Kurtz (1989), Boone and Kurtz Marketing Disk, New York, New York: The Dryden Press,

Collins, Robert H, (1987). 'Sales Force Support Systems: Potential Applications to Increase Productivity, Journal of the Academy of Marketing Science, Vol. 15, No. 2, 49-54.

Dyer, Robert F. (1987), 'An Integrated Design for Personal Computers in the Marketing Curricu-

lum," Journal of the Academy of Marketing Science, Vol. 15. No. 2, 16-24

ESIE (Expert System Inference Engine) (1986). PC-SIG Library Disk #398, Lightwave, P.O. Box 16858, Tampa, FL.

Evans, Joel R. and Barry Berman (1987), Computer Exercise Diskette, New York, New York, Macmillan.

Good. R. (1987). "Artificial Intelligence and Science Education,' Journal of Research in Science Teaching. 24 (4), 325-342.

Goodspeed, J. (May/June. 1988, "World Leaders Call for Better Research in Education Technology,' Electronic Learning. 7 (8). 24-26.

Kearsley, G. Hunter, B. and R. Seidel (1982). "Two Decades of Computer Based Instruction Projects: What Have We Learned?" The Journal, 10 (3 & 4)). 90-96.

Kearsley, G. (19871. Artificial Intelligence & Instruction: Applications and Methods. New York, New York: Adison-Wesley.

Kurtz, David L. and Louis E. Boone (1987), "The Current Status of Microcomputer Usage in the Marketing Programs of AACSB-Accredited Colleges and Universities,' Journal of the Academy of Marketing Science, Vol. 15, No. 2, 10-15.

Malhotra, Naresh K., Tashchian Armen, and Mahound Eassom (1987). "The integration of Microcomputers in Marketing Research and Decision Making," Journal of the Academy of Marketing Science, Vol. 15. No. 2, 6982.

McCarthy, E. Jerome and William D. Perreault, Jr. (1987). Computer-Aided Problems to Accompany Basic Marketing Homewood. Illinois: Irwin.

Robertson. O. Zeller Jr. (1987), "Which Statistic?" Social Science Microcomputer Review, Vol. 5, No. 3, 350-353.

Sechrist, Richard (1988), The Statistical Consultant, PC-SIG Library, Disk #949, Dept. of Geography & Regional Planning, Indiana University of Pennsylvania, Indiana, Pa.

Steinberg, Margery and Richard E. Plank (1987), "Expert Systems: The Integrative Sales Management Tool of the Future." Journal of the Academy of Marketing Science, Vol. 15, No. 2, 55-62.

Stevens, S. S. (1960), On the Theory of Scales of Measurement." In Arthur Danto and Sidney Morgenbesser (Eds.). Philosophy of Science New York: the World Publishing Company, 141- 149.

FIGURE 1. EXAMPLE OF STATUTOR SYNTAX IN ESIE

```
Goal is stat
[relevancexxx are y n *
if descriptive is y
and nominal is y
then stat is Measure.of.Central.Tendency.is...Mode
question descriptive is"
Is the method intended to summarize the data (that is, descriptive
statistics)?
question nominal is"
Is 'are' the variable(s) NOMINAL?
if descriptive is y
and ordinal is y
then stat is Measure.of.Central.Tendency.is...Median
question ordinal is" Is (are) the variable(s) ORDINAL?
if descriptive is y
and interval is y
then stat is Measure.of.Central.Tendency.is...Mean
question interval is" Is (are) the variable(s) INTERVAL?
answer is"
The analytical technique you should consider is:' stat
incorrect is"
    Because of the large number of statistical techniques
available, the researcher may have difficulty in selecting the
technique most appropriate for the situation.
    STATUTOR is an expert system designed to assist you in
selecting the appropriate statistical technique for your problem.
    The system will ask you a series of questions, starting with,
what is the method intended to show? Responses to the questions
lead to the identification of a particular statistical) technique.
    The questions are phrased for yes and no responses. In these
cases one need only type Y or N."
terminate is"
    If you would like another analysis selection type GO at the *>
symbol.
    If you wish to end your session and return to DOS the type
EXIT at the *> symbol"
```

FIGURE 2. OPENING SCREEN OF CONSULTATION

```
**>go

                              STATUTOR:
A SYSTEM FOR SELECTING ANALYTICAL TECHNIQUES FOR ANALYZING MARKETING
                          RESEARCH DATA

Because of the large number of statistical techniques available, the
researcher may have difficulty in selecting the technique most
appropriate for the situation.
    STATUTOR is an expert system designed to assist you in selecting
the appropriate statistical technique for your problem.
    The system will ask you a series of questions, starting with,
what is the method intended to show? Responses to the questions lead
to the identification of a particular statistical technique.
    The questions are phrased for Yes or No responses. In these cases
one need only type Y or N.
    Please begin by answering the question below.

Is the method intended to summarize the data (that is, descriptive
statistics)?
```

FIGURE 3. TYPICAL CONSULTATION SESSION SCREEN

```
Is the method intended to summarize the data (that is, descriptive
statistics)?
n

Is the method intended to determine observed differences in the data?
n

Is the method intended to investigate associations?
y

Are some of the variables DEPENDENT ON OTHERS?
y

Are the number of dependent variables equal to ONE?
y

Is the scale of the dependent variable NOMINAL?
n

Is the scale of the dependent variable ORDINAL?
n

Is the scale of the dependent variable INTERVAL?
```

FIGURE 4. TYPICAL CONCLUSION SCREEN

```
Is the scale of the independent variable INTERVAL?
y

Is the number of independent variables = ONE?
n

Is the number of the independent variables MORE THAN ONE?
y

Technique you should consider:MULTIPLE.REGRESSION...AND...CORRELATION.ANALYSIS

    If you would like another analysis selection type GO at the *>
symbol.

    If you wish to end your session and return to DOS then type EXIT
at the *> symbol.

I have completed this analysis.

**>
```

FIGURE 5. CONSULTATION SESSION IN TRACE MODE

```
Thinking: is ASSOCIATION  Y?

Is the method intended to investigate associations?
y
It has been learned that ASSOCIATION
    is Y.
Thinking: is STAT  STAT?
Thinking: is DEPENDENCE  Y?

Are some of the variables DEPENDENT on others?
y
It has been learned that DEPENDENCE
    is Y.
Thinking: is STAT  STAT?
Thinking: is CRITERION.ONE  Y?

Are the number of dependent variables equal to ONE?
y
It has been learned that CRITERION.ONE
    is Y.
Thinking: is STAT  STAT?
Thinking: is CRITERION.NOMINAL  Y?

Is the scale of the dependent variable NOMINAL?
```

FIGURE 6. CONCLUSIONS IN TRACE MODE

```
Thinking: is STAT  STAT?
Thinking: is PREDICTOR.TWO  Y?

Is the number of the independent variables MORE THAN ONE?
y
It has been learned that PREDICTOR.TWO
    is Y.
Thinking: is STAT  STAT?
It has been learned that STAT
    is MULTIPLE.REGRESSION..AND..CORRELATION.ANALYSIS.

Technique you should consider:MULTIPLE.REGRESSION. and .CORRELATION ANALysis

     If you would like another analysis selection type GO at the **
prompt.

     If you wish to end your session and return to DOS then type Exit
at the ** symbol.

I have completed this analysis.

**>
```