Pergamon

Expert Systems With Applications, Vol. 8, No. 1, pp. 203–212, 1995

0957-4174/95 $9.50 + .00

0957-4174(94)E0010-R

The Role of Rules and Examples in the Process of Knowledge Acquisition in Direct Classification Tasks

D. L. OLSON

Department of Business Analysis & Research, Texas A & M University, College Station, TX

A. I. MECHITOV AND H. M. MOSHKOVICH

Department of Decision Sciences, Institute for Systems Analysis of the Russian Academy of Sciences, Moscow, Russia

Abstract—*Shells provide a means for experts to easily develop expert systems for their area of expertise. However, rule bases need to be complete and free of contradictions. A set of* 30 *subjects, unfamiliar with shells except for initial orientation and training, were asked to develop a system for their personal preferences for a decision problem. The results of these systems were analyzed, leading to a number of conclusions. First, three types of rules used by the subjects were identified. Cutoff rules reflect preemptive treatment of decision rules. Examples reflect an attempt to enumerate all combinations of decision factors. Compensatory rules reflect attempts to balance trade-offs among the relative performance of decision cases. The implications of using these three types of rules are evaluated. Subjects validated their systems on a test bank of* 18 *cases. Subject responses to the impact of these test cases were evaluated, revealing that they thought that the test cases yielded more complete systems. Posttest evaluation of their systems for completeness and consistency also revealed that the systems still included significant gaps in rules. We conclude that computer aids to assist experts need to include means to assure consistency and completeness of knowledge bases. Further, at least some compensatory rules should be included for those cases that involve trade-offs.*

1. INTRODUCTION

CLASSIFICATION TASKS are very important in their own right, but especially in the context of expert systems. Classification tasks require assignment of all possible cases to a small number of decision groups. Usually these decision groups are ordered, reflecting different degrees of attribute quality. For example, research and development projects may need to be categorized into the classes of "accepted" and "rejected"; a consumer may want to divide available products on various levels of quality, perhaps based on input from a consumer's guide, the opinions of others who have used these products, or on past personal experience; economists may want to use market indicators to classify market trends as increasing, neutral, or decreasing. In Boose's (1989) taxonomy of knowledge-based applications, classification tasks are a major subheading under analysis problems. Classification tasks are the most common application of expert system shells.

Requests for reprints should be sent to D. L. Olson, Department of Business Analysis & Research, Texas A & M University, College Station, TX 77843-4217, USA. E-mail: DOLSON@TAMVM1.TAMU.EDU

To provide effective expert systems capable of dealing with such tasks, it is necessary to form a set of classification rules on the bases of expert knowledge. There has been quite a bit of research in knowledge acquisition techniques in general. Gaines (1987) provided a very comprehensive early overview of the field. That research includes structured elicitation methodology given by Garg-Janardan and Salvendy (1988) and machine rule induction (see Byrd, Cossick, & Zmud, 1992). Here we are dealing with a specific kind of knowledge to be acquired. The expert has developed the ability to classify cases into some finite set of categories. The most popular approach is to directly ask the expert to enter his or her implicit rules in the form of productions (as in any shell system). Another approach is to ask experts to enter real examples from their practice, or to directly classify some objects, and then attempt to derive some set of rules that fits these classifications (as in Clancey, 1985; Grover, 1983; Larichev & Moshkovich, 1990; Wright & Ayton, 1987). In Mechitov, Moshkovich, & Olson (in press), two forms of knowledge elicitation were investigated in a small classification task of job selection by students anticipating graduation. Subjects were asked to form a full and consistent set of rules to classify one of four possible classes of job offers described by three or four possible values on each of five attributes. This yielded a $3 \times 3 \times 4 \times 3 \times 3$ design, representing the first approach of direct entry of production rules.

Elicitation of rules is a vital and necessary element in the knowledge acquisition process. Edmonds, O'Brien, & Bayley (1993) cited unsatisfactory capture and transfer of domain expertise as a major hindrance to more widespread use of expert systems. It is necessary to accurately identify knowledge, and to assure the quality of that knowledge (O'Keefe, Balci, & Smith, 1987), two tasks called crucial and difficult by Preece (1993). Computer aids that provide more complete and consistent rule bases are needed.

We report an experiment where we used a well-known system, VP-EXPERT, as a direct elicitation tool for the conventional rule-based approach. The experimental results are not limited by the use of VP-EXPERT, as many other good systems incorporate the same type of rule-based structure. VP-EXPERT is easy to use and has the properties of most production shells, making it easy to introduce production rules into the rule base; it can also be used to test the rule base on a set of example cases. Validation of knowledge bases is the traditional approach for knowledge acquisition. Nazareth (1989) reviewed some techniques for system knowledge verification.

Section 2 of this paper will describe the experimental design of our study. Section 3 will present analysis of the results, followed by discussion in Section 4. The conclusions are presented in section 5.

2. EXPERIMENTAL DESIGN

The intent of the experiment was to analyze subject direct formulation of rules (in the form of productions) with the help of VP-EXPERT. As is true with all shell systems we are aware of, VP-EXPERT allows users to enter rules that could be inconsistent. The focus of the comparison is on completeness and consistency of the resulting rule bases.

2.1. Task

The 30 subjects were senior level undergraduate students (from a business college). The task was evaluation of job opportunities similar to those possible for the subject group. The rule base was to reflect each subject's preference structure for the task. The students were on the job market. Therefore, subjects were considered experts for this task. Jobs were characterized by five attributes, JOB TYPE, LOCATION, SALARY, TRAINING, and PROMOTION. Four of the attributes had three-point ordinal scales (attractive, medium, and unattractive, with appropriate labels—see Appendix 1). SALARY was given four categories. Subjects were asked to categorize job opportunities into one of four ordinal result categories: attractive, acceptable, poor, and unacceptable job opportunities. The classification task is simple, not requiring decomposition of complicated structures. There were 324 possible combinations of attribute categories, which could feasibly be enumerated, but this number was large enough that enumeration was unlikely. The simplicity of the task allowed us to concentrate on the elements of human judgment required.

At the time of the study, subjects were taught the basics of expert systems, and production rules as well as other knowledge base construction approaches were presented to them. Students were taught how to use VP-EXPERT. The subjects were assigned the task of constructing a set of production rules to classify job opportunities using VP-EXPERT.

Because the most common approach espoused for knowledge acquisition is based on experts entering sets of rules into a rule base and then testing them on a set of test cases reflecting the expected real environment, subjects were instructed to test their rule bases on 18 test cases. The test cases (Appendix 2) were selected on the basis of involving tough trade-offs among the five criteria, with no one test case standing out as obviously superior. Subjects were allowed to change their rule bases with the intent of improving them, based on their test results. The final rule bases were analyzed in this study. Subjects were also asked to answer two questionnaires. The first questionnaire sought to measure subject overall attitude toward the system and its results (see Appendix 3). The second questionnaire examined the process subjects used to construct their rule bases (see Appendix 4).

2.2. Experimental Data Processing

Analysis of the results focuses on three areas. First, we examine how subjects worked with the system, their attitude toward the system, its results, and the overall process. Second, we analyze the rule bases constructed by subjects. Third, we test some hypotheses about the relationships of different features involved in the process of rule construction. Subject attitudes toward the process were estimated on the basis of their responses to the two questionnaires.

The first area studied involved an overall estimation of subject attitude toward the process and their effort in using the system. The first questionnaire measured:

- ease of working with the system;
- satisfaction with the results given by the system;
- understandability of the system;
- subjective evaluations of the "speed" of the process;
- usefulness of the results;
- how reasonable actual use of this system would be.

The second questionnaire measured the time and labor required by the process.

- time to construct, test, and change the rule base (in hours);
- ease of constructing the rule base, on a three-point scale;
- ease of rule base testing, on a three-point scale;
- the degree of rule base modification after testing, on a three-point scale.

The second area investigated looked at the rule bases from the following perspectives:

- number of rules in the rule base;
- type of rules used in the rule base;
- rule base completeness;
- rule base consistency.

Because these measures are central to our analysis, we will discuss the reason for including each.

2.2.1. *Number of Rules in the Rule Base.* This measure provides an indication of rule base complexity. It is easily measured by counting the number of VP-EXPERT rules. It is a rough measure, because it is possible to have one rule serve the same function as a set of a dozen rules in some cases.

2.2.2. *Type of rules used in the rule base.* Most of the rules used by subjects gave a set of attribute values connected by logical operators AND and OR, as well as possible values for one attribute connected by the logical operator OR, or by arithmetic operators =, <, and >. Rules came in three basic types. *Cutoff rules, examples,* and *compensatory rules.*

A cutoff rule is based on the psychological differentiation between more simple and more complicated decision rule types (Payne, 1976). People often use very simple rules to eliminate unsatisfactory alternatives. They usually focus on one or at most two attributes to formulate rules. For example, "IF SALARY is poor, the job is unacceptable" or IF SALARY is high and JOB IS LOCATED IN AN ATTRACTIVE PLACE, the job is attractive." Cutoff rules are very easy to formulate, to analyze, and to understand. Another peculiarity of cutoff rules is their ability to classify large numbers of possible variants of the entire set of 324 combinations. However, they are noncompensatory. Implementation of such rules simplifies rule base construction, but they usually give an insufficient reflection of real decision-maker preference.

Examples are rules that describe only one of the possible set of combinations (or in the test case used, full description of one of the 324 possible variants). This means that a rule would have all attributes present, and only one value included for each attribute. A complete rule base for a case involving 324 possible variants would have 324 rules, a complete enumeration of all possible cases.

We label as compensatory rules all rules that are not cutoff rules or examples. Compensatory rules try to balance trade-offs in relative performances on different attributes. Compensatory rules can combine different numbers of attributes and their values in one rule. An example of such a rule is:

IF POSITION = inappropriate AND
LOCATION = poor OR LOCATION = average AND
SALARY = high OR SALARY = average AND
PROMOTION = good OR PROMOTION = moderate
THEN class = acceptable.

We believe, with support in Mechitov et al. (in press), that the type of rules used correlates with other characteristics of the rule base, which is why we pay so much attention to this feature.

2.3. Rule Base Completeness

Rule base completeness is a major concern in knowledge elicitation. Although most real expert systems do not claim to be fully complete, relying upon expansion during implementation, they should be as complete as possible. It is always desirable to construct a knowledge base that covers all cases that could arise. Testing by using examples is designed to spot gaps in the knowledge base, as well as to check for rule base correctness.

Our subjects were asked to construct a set of rules for classifying all possible variations of jobs. The task was relatively small, so therefore it was natural to expect complete rule bases. Each rule introduced into the rule base determines an appropriate decision class for a set of the 324 possible combinations. If the rule is formulated on a small number of attributes, all of the possible values for those attributes not addressed by the rule are considered the same. For example, if we have the rule

RULE A
IF SALARY = high OR SALARY = average AND
LOCATION = great OR LOCATION = average
THEN class = attractive

this rule covers all possible values for the attributes POSITION, TRAINING, and PROMOTION. We can easily calculate how many variants will be defined as belonging to the attractive class by this rule: $3 \times 2 \times 2 \times 3 \times 3 = 108$. This method was used to evaluate the number of variants classified by the rules. The union of all variants, covered by different rules, will form the set of classified variants.

2.4. Rule Base Consistency

Almost all expert system specialists agree that rule base consistency is needed. However, the usual approach to assure consistency is to let the expert check the rule base in some way, usually by testing some example cases. If inconsistencies are found, the rule base is changed. To check the consistency of rule bases in this study, special efforts were made to identify contradictory rules.

Let us return to the last rule presented, which covered 108 combinations of the possible 324. Let us also consider the following rule:

RULE B
IF TRAINING = poor AND
PROMOTION = none
THEN class = inappropriate.

What does this rule mean? It means that all variants covered by rule A that have the worst value on the fourth and fifth attributes belong to the fourth class of outcome. We can interpret the range of variants of rule B by a pair of variants. The most preferable of this set, expressed in vector form (with class 1 for each criterion representing the best case), is 11133 and the least preferable is 33433. However, the range of rule A assigned variant 11133 to the first class. Therefore, we have a case of inconsistency, or contradiction. Two rules are considered contradictory if they assign the same alternative to different decision classes. The number of contradictory rules is calculated by counting the number of contradictory rules in the rule base.

The third area analyzed is our hypotheses about the relationship of the different types of rules in the rule bases with the number of contradictory rules. We shall test five hypotheses:

1. The number of contradictory rules is dependent upon the number of rules in the rule base.
2. The number of contradictory rules is dependent upon the number of cutoff rules.
3. The number of contradictory rules is dependent upon the number of compensatory rules.
4. The number of contradictory rules is dependent upon the number of example rules.
5. The number of contradictions is dependent upon efforts to modify the rule base after testing.

3. ANALYSIS OF RESULTS

Table 1 presents data for the 30 subjects, giving their responses to the questionnaires (all 30 completed the first questionnaire, 26 of these 30 completed the second questionnaire). We see that subjects considered VP-EXPERT moderately easy to work with, but required a great deal of time. Most subjects understood how the system obtained results (question 3), but at the same time, they were not as thoroughly satisfied with the system results (question 2), even though the rule bases were what they chose. Subjects had no definite opinion concerning the usefulness of such a system (questions 5 and 6). In general, subject attitude toward the shell system (VP-EXPERT) was positive.

The process of constructing and testing the rule base did not seem difficult to most of the subjects. However, most of the subjects found this process time consuming. Most of the subjects made slight modifications to the rule base after their testing (22 out of 27). Most subjects commented that they did not change the rule base, but rather added to it to cover combinations of job opportunities that they had inadvertently overlooked. Overall, subjects were satisfied with their rule bases, and most of them were confident that they had complete and consistent rule bases.

In Table 2 we present the results of the objective analysis of rule bases. We can see that the number of rules varied from 4 to 30, averaging about 14 rules. Only one of the subjects was able to cover all possible variants, which he did using only 6 rules. He used 4 simple cutoff rules, covering most of the variants, and the other two rules were fairly simple compensation rules. Thorough analysis of his rule base led us to conclude that he used only 4 attributes and 3 decision classes. So his task was actually of smaller dimensionality than the others (108 variants in 3 classes).

None of the other subjects was able to develop a complete rule base. The average number of classified variants was around two-thirds of the initial set. Six subjects out of 30 were able to avoid contradictory rules in their rule bases. We expected that such variants would be related to the small number of rules used. But those subjects who avoided contradictions had quite large rule bases (18 to 25 rules).

We examined the structure of rule types. Three subjects included a large number of rules in their rule bases (subjects 6, 13, and 21). Almost all of the rules used by these subjects were examples. This supports our previous findings (Mechitov et al., in press) that it is much easier for people to be consistent when using examples rather than other types of rules. This leads us to the conclusion that rule bases developed from expert description of tasks need to be verified by testing. Two other subjects (10 and 14) had no contradictions because they had rule bases with a very small number of simple rules.

We are able to conclude that simple tasks involving simple cutoff rules are the most likely to avoid inconsistency. In practice, people tend to make a large proportion of mistakes (4 out of 14 rules, on average), and need some help to cope with this problem. It is obvious that use of testing examples gave additional rules, but was not able to identify and overcome internal rule base inconsistency. This is supported by the results of testing the different parameters for their correlation with the number of contradictions through ANOVA. The results are presented in Table 3. Hypotheses 1 and 5 are accepted with a very high level of probability. Hypothesis 3 may be accepted with a 10% confidence. Hypotheses 2 and 4 must be rejected.

TABLE 1
General Data on Subjects' Answers to Questionnaire

Subject	CONV	RESU	UNDE	SPEE	USEF	EASE	Time (hr)	East of Construct	Ease of Testing	Change to Rulebase
1	3	2	1	2	2	2	2.0	2	1	1
2	3	2	1	4	2	2	1.5	2	1	2
3	2	2	1	1	2	2	2.0	1	1	1
4	3	2	1	2	1	2	2.0	2	1	2
5	2	3	2	4	3	2	3.5	1	1	1
6	4	1	1	2	1	2	1.0	2	1	2
7	3	2	1	3	2	2	2.5	1	1	1
8	3	2	1	1	2	2	0.5	2	1	0
9	2	2	1	2	1	1	1.0	2	1	1
10	4	4	2	4	3	2	1.0	3	3	0
11	2	2	1	3	1	2	1.75	2	1	2
12	2	2	1	1	2	2	1.0	3	3	2
13	2	1	1	3	3	2	2.0	1	1	0
14	3	1	1	3	1	1	3.0	2	1	0
15	3	1	1	3	3	3	3.0	1	2	1
16	1	1	2	3	2	1	1.0	1	1	0
17	3	2	1	3	2	2	1.5	2	2	2
18	2	2	1	2	1	2	2.0	1	1	2
19	4	2	2	1	3	3	2.0	1	1	2
20	2	3	1	2	1	2	1.0	1	1	1
21	2	2	1	1	2	1	2.0	2	1	2
22	3	4	1	2	3	3	1.0	1	2	2
23	2	2	1	2	2	2	1.0	1	2	1
24	4	1	1	2	2	2	3.0	2	1	1
25	4	2	2	3	2	2	1.5	3	3	0
26	2	3	2	3	2	2	2.5	3	2	1
27	3	3	1	1	2	2				
28	3	2	3	2	3	3				
29	2	2	2	2	2	2				
30	1	2	2	1	2	1				
Average	2.6	2.1	1.3	2.3	2.0	2.0	1.8	1.7	1.4	1.2

CONV—convenience of the system.
RESU—satisfaction with the result.
UNDE—understanding of the system process.
SPEE—speed of the system (subjective estimate of time required).
USEF—usefulness of the system output.
EASE—ease of using the system.

4. DISCUSSION

The data we present allows us to arrive at several conclusions concerning human ability to construct rule bases in the form of production rules. First, we are able to discriminate between at least three types of rules used by subjects, each requiring different cognitive efforts and leading to different results relative to completeness and consistency. These three rule types are cutoff rules, examples, and compensatory rules. Each type of rule demonstrated a different impact on results.

Cutoff rules are simple but effective. The subject can use cutoff rules to eliminate a set of unsatisfactory variants (IF SALARY = poor THEN job = inappropriate) or to select a set of preferred variants for the most preferred class (IF SALARY = high THEN job = attractive). Cutoff rules are easily formulated (and therefore easy to understand), cover a lot of variants efficiently (thus and seldom result in contradictory rule bases (hypothesis 2 in Table 3, contending that the number of contradictions would be related to the number of cutoff rules, was rejected). Almost all subjects used this type of rule. Despite all of these good features of cutoff rules, it must be understood that in real tasks there is no possibility to cover all necessary variants with cutoff rules. We would expect the need to apply compensatory rules to accurately represent the complications of real decisions (see Larichev, 1992).

Examples represent rules defining the appropriate class for one variant out of the initial set. It is very close to the concept of an expert who records those cases that they have experienced in the past. The primarily peculiarity of such a rule is that it only represents one variant, and thus is not generalized knowledge. Examples, like cutoff rules, are very simple for users

TABLE 2
Analysis of Subjects' Rule Bases

Subject	Number of Rules	Types of Rules Used: Cut-offs	Compen	Examples	Classification Variants	Number of Contradictions
1	12	0	11	1	236	5
2	10	0	8	2	173	3
3	11	5	5	0	243	3
4	12	3	9	0	307	2
5	11	4	7	0	242	5
6	18	0	0	18	100	0
7	25	2	0	23	137	2
8	8	4	4	0	318	3
9	6	3	2	1	268	8
10	4	4	0	0	127	0
11	10	0	8	2	110	5
12	12	9	3	0	238	9
13	18	0	0	18	82	0
14	6	4	2	0	324	0
15	9	1	7	1	183	6
16	6	4	2	0	288	2
17	21	1	20	0	150	7
18	12	4	8	0	184	8
19	10	1	8	1	216	6
20	20	5	10	5	306	5
21	25	3	4	18	218	0
22	16	6	10	0	117	10
23	24	0	3	21	137	2
24	29	1	27	1	296	8
25	7	2	4	1	250	4
26	12	3	6	3	228	7
27	30	22	8	0	304	11
28	10	7	3	0	238	2
29	6	3	3	0	216	0
30	17	4	13	0	291	3
Average	13.9	3.5	6.5	3.9	217.6	4.2

Number of rules—number of rules in the rule base.
Cut-offs—number of cut-off rules in the rule base.
Compen—number of compensatory rules in the rule base.
Examples—number of examples in the rule base.
Classification variants—number of the 324 possible variants classified.
Contradictions—number of contradictory rules in the rule base.

to implement. Hypothesis 4 proposed that the number of examples in the rule base would be correlated with the number of contradictory rules. This hypothesis was rejected (see Table 3). The primary disadvantage of examples is that they cover a small number of variants, and complete rule bases using examples would be very inefficient.

Compensatory rules are what is traditionally meant when discussing production rules. Compensation rules combine a rather large number of possible variants

TABLE 3
Results of Hypotheses Testing

Hypothesis	F-test	P-value	
1 Number of contradictory rules depends on the number of rules	44.79**	<.0001	Supported
2 Number of contradictory rules depends on the number of cut-off rules	0.57	0.47	Rejected
3 Number of contradictory rules depends on the number of compensatory rules	3.5*	0.07	Supported
4 Number of contradictory rules depends on the number of example rules	0.05	0.82	Rejected
5 Number of contradictory rules depends on the number of rule base modifications after testing	24.4**	<.0001	Supported

through small ranges of possible values upon several attributes. They play an intermediate role in providing complete rule bases, because they cover broad sets of variants as cutoff rules do, but require far fewer rules to cover the same variants as examples would. Compensatory rules resulted in the greatest number of contradictions. Compensatory rules are relatively more difficult to formulate. Hypothesis 5 proposed that the number of rule base modifications would be related to the number of contradictions. This hypothesis was not rejected (see column 11 in Table 1), which indirectly states that the more subjects tried to modify their rule base, the more contradictions they created.

We have data for 9 subjects on their classifications of test examples prior to testing. Analysis of this data showed large differences in the pretest and posttest classifications (see Table 4).

We see that only one subject had identical pretest and posttest classifications. Subjects 17 and 22 rated some test cases as attractive while their rule bases classified these job opportunities as inappropriate. This is in spite of the fact that both subjects claimed that they were sure that they had made all necessary modifications to make their rule bases complete and noncontradictory.

We note that the subjects were students completing a class assignment, and therefore their motivation is less than desired. Yet we consider them experts in the task they were asked to perform, and they were under the impression that complete and noncontradictory rule bases would be to their benefit. The results of this study reflect that correction of your own rule base is a complex and difficult task. Otherwise, subjects would have managed to avoid incomplete and inconsistent rule bases.

Why is this task so difficult? We think that the main source of contradictions is related to the features of compensatory rules. Subjects select several important attributes and formulate appropriate decision classes. But as we showed in section 2, this action by the subject allows **any** possible value upon other attributes not used in the rule. In many cases it is appropriate that if these other attributes have attractive or unattractive enough values, the expert may wish to modify the original classification based on the salient attribute. What we observe is that these combinations of other attributes may be addressed by later compensatory rules, but the resulting complexity of the overall rule base structure covers the contradiction. In order to identify such contradictions, humans need assistance.

How can this difficulty be overcome? One possibility is to provide the system with the ability to expand the entered rule to consider values on all attributes. This simple modification would simplify rule testing for users, as they would focus on the rule they were formulating. System CLASS (Larichev & Moshkovich, 1990) provides a means to check for both consistency and completeness. CLASS is oriented on complete classification of all possible variants over all attributes, operating by presenting subjects with hypothetical examples from this set for classification. CLASS uses dominance relationships to gain some efficiency, meaning that not all combinations of vectors need be considered by the subject. But completeness is assured by mapping all responses and resulting inferences. Contradictions are also easily identified. In previous work we have done (Mechitov et al., in press), subjects averaged more than 100 classification examples for a problem of similar size.

Formulation of more generalized rules, covering many more variants from the set of all possible combinations of attribute values, is a much quicker way to guarantee full classification. That is why CLASS includes the ability to enter some rules as examples. This could be done in VP-EXPERT terms:

IF SALARY = high OR SALARY = average AND
LOCATION = great OR LOCATION = average
THEN class = attractive.

As we have rank-ordered decision classes and rank-ordered values on attribute scales, we are able to enter boundaries on the extreme classes (enter the worst vector that is to be classified in the best class, and enter the best vector that is to be classified in the worst class). We calculated the number of example rules that would be needed to cover all rules, presented in Table 5. These numbers are much less than those presented in Mechitov et al. (in press) where the average number of examples that subjects were asked to classify by using CLASS was greater than 100.

TABLE 4
Analysis of Testing Examples

	Subject								
	1	5	7	17	18	19	22	24	30
Number of inconsistent classifications	8	1	9	1	5	8	10	0	4
Number of differences in 2 classes	0	0	0	0	0	0	0	0	1
Number of differences in 3 classes	1	0	0	0	0	0	2	0	0

TABLE 5
Number of Examples Used in CLASS to Represent Rules

Subject	Number of Example Rules Required
1	28
2	14
3	14
4	21
5	22
6	18
7	30
8	22
9	14
10	9
11	15
12	13
13	18
14	7
15	21
16	11
17	26
18	18
19	22
20	23
21	35
22	23
23	28
24	49
25	19
26	26
27	32
28	19
29	11
30	17
Average	20.87

5. CONCLUSIONS

People tend to use a variety of rule types in using expert system shells. Compensatory rules are most appropriate relative to decision-making theory (see Keeney & Raiffa, 1976; Larichev, 1992), because they involve focused consideration of tradeoffs. On the other hand, cut-off rules are much more efficient, but in effect discard from consideration those attributes not included in the rule. Example rules have the positive feature that they obviously will assure consistency, but are usually impractical due to the number of attribute combinations possible.

Our study found no relationship between the number of cut-off rules and contradictions in the rule base, nor was any relationship found between the number of example rules and contradictions in the rule base. In fact, only compensatory rules had a significant relationship with contradictions. We infer that cut-off rules and examples are efficient ways to avoid contradictions, but are dangerous in that they demonstrate a cursory consideration of the decision problem on the part of the expert. Compensatory rules reflect greater expert focus on the decision, but create additional need for experts to be concerned about contradictions in their rule bases. In like manner, more complex rule bases (as measured by the number of rules) also lead to more contradictions, as would be expected.

The conventional technique to minimize rule base errors is to test the system on a set of cases. Our study found that testing identified a number of contradictions, and subjects thought that they had corrected their rule bases. But the rule bases still included contradictions. Again, this indicates to us that a computerized aid to detect inconsistencies would be useful.

We conclude that people must use all three types of rules when constructing knowledge bases. Usually each process of rule base construction starts with simple and indisputable cut-off rules. Then the expert starts to formulate more complicated compensatory rules to cover more complex cases, often based on examples gained from past experience. At this moment, a powerful aid to assist in interpreting and to verify such rule bases is needed. Larichev, Moshkovich, Furem's, Mechitov, & Morgoev (1991) demonstrated with simulation studies that in this process, experts classified almost 80% of possible variants. After this level of classification, it was found that the expert needed help to finish the job, identifying gaps in the knowledge base and presenting them to the decision maker for classification.

Our study focuses on the application of expert systems to classification tasks, and therefore is not generalizable to other expert system applications. An additional limitation is that any research involving human subjects requires replication. In this study, student subjects were used. The task environment was one that they understood very well. But grade-driven motivation is always a shortcoming. Future research comparing results using system CLASS as well as a shell system is underway. Future work is planned using students as well as other subjects.

REFERENCES

Boose, J.H. (1989). A survey of knowledge acquisition techniques and tools. *Knowledge Acquisition*, **1**, 3–37.

Byrd, T.A., Cossick, K.L., & Zmud, R.W. (1992). A synthesis of research on requirements analysis and knowledge acquisition techniques. *MIS Quarterly*, **16**(1), 117–138.

Clancey, W. (1985). Heuristic classification. *Artificial Intelligence*, **27**, 289–350.

Edmonds, E.A., O'Brien, S.M., & Bayley, T. (1993). Constructing end-user knowledge manipulation systems. *International Journal of Man–Machine Studies*, **38**, 51–70.

Gaines, B.R. (1987). An overview of knowledge acquisition and transfer. *International Journal of Man–Machine Studies*, **26**, 453–472.

Garg-Janardan, C., & Salvendy, G. (1988). A structured knowledge elicitation methodology for building expert systems. *International Journal of Man–Machine Studies*, **29**, 377–406.

Grover, M.D. (1983). A pragmatic knowledge acquisition methodology. *IJCAI-83*, 436–438.

Keeney, R.L., & Raiffa, H. (1976). *Decisions with multiple objectives: Preferences and value tradeoffs*. New York: Wiley.

Larichev, O.I. (1992). Cognitive validity in design of decision aiding techniques. *Journal of Multicriteria Decision Analysis*, **1**, 127–138.

Larichev, O.I., & Moshkovich, H.M. (1990). Decision support system "CLASS" for R&D planning. In *Proceedings of the First Inter-*

national Conference on Expert Planning Systems (pp. 227–232). Brighton, England.

Larichev, O.I., Moshkovich, H.M., Furems, E.M., Mechitov, A.I., & Morgoev, V.K. (1991). *Knowledge acquisition for the construction of full and contradiction free knowledge bases*. iec ProGAMMA, Groningen, The Netherlands.

Mechitov, A.I., Moshkovich, H.M., & Olson, D.L. (in press). Problems of decision rule elicitation in a classification task. *Decision Support Systems*.

Nazareth, D.L. (1989). Issues in the verification of knowledge in rule-based systems. *International Journal of Man-Machine Studies*, **30**, 255–271.

O'Keefe, R.M., Balci, O., & Smith, E.P. (1987). Validating expert system performance. *IEEE Expert*, **2**(6), 81–90.

Payne, J.W. (1976). Task complexity and contingent processing in decision making: An information search and protocol analysis. *Organizational Behavior and Human Processes* **16**, 366–387.

Preece, A.D. (1993). A new approach to detecting missing knowledge in expert system rule bases. *International Journal of Man-Machine Studies*, **38**, 661–688.

Wright, G., & Ayton, P. (1987). Eliciting and modelling expert knowledge. *Decision Support Systems* **5**, 13–26.

APPENDIX 1: CRITERIA AND CATEGORIES FOR JOB CLASSIFICATION

Criterion 1. Type of the job position

1. Type of the job position is almost ideal.
2. Type of the job position is good enough (in field).
3. Type of the job position is not appropriate.

Criterion 2. Job location

1. Location of the job is where you want to be.
2. Location of the job is in some distance from where you want.
3. Job is located far away from where you want.

Criterion 3. Salary

1. The salary is high.
2. The salary is on the average level.
3. The salary is a bit lower than the average level.
4. The salary is poor.

Criterion 4. Possibilities for training

1. There are good possibilities for training.
2. There are average possibilities for training.
3. There are minimal (almost no) possibilities for training.

Criterion 5. Possibilities for promotion

1. There are good possibilities for promotion.
2. There are moderate possibilities for promotion.
3. There are almost no possibilities for promotion.

JOB CLASSES

1. This job is very interesting — *highly attractive*
2. This job is acceptable — *acceptable*
3. This job is poor, would only take in a pinch — *poor*
4. You would not take this job — *unacceptable*

APPENDIX 2: TEST CASES

Check Your Set of Rules with the Following 18 Cases:

Case	Position	Location	Salary	Training	Promotion
1	Almost ideal	Average	Poor	Good	Moderate
2	Almost ideal	Average	Below average	Normal	Moderate
3	Almost ideal	Poor	Below average	Normal	Good
4	Almost ideal	Poor	Below average	Poor	Moderate
5	Almost ideal	Average	Average	Normal	None
6	Almost ideal	Average	Average	Poor	Moderate
7	Good enough	Great	Below average	Normal	Moderate
8	Good enough	Average	Below average	Good	Moderate
9	Good enough	Poor	Below average	Good	Good
10	Good enough	Great	Average	Normal	Poor
11	Good enough	Average	Average	Poor	Good
12	Good enough	Poor	Average	Good	Moderate
13	Inappropriate	Great	Average	Good	Poor
14	Inappropriate	Average	Average	Normal	Good
15	Inappropriate	Poor	Average	Good	Good
16	Inappropriate	Great	High	Normal	Poor
17	Inappropriate	Average	High	Poor	Good
18	Inappropriate	Average	High	Normal	Moderate

APPENDIX 3: SUBJECTIVE QUESTIONNAIRE

Question 1. Was it convenient to work with the system?

1. It was easy and comfortable.
2. There were some difficulties while working with the system, but on the whole it was convenient.
3. There were difficulties, but the system may be considered convenient enough.
4. There were essential difficulties. I would say it was not very convenient to work with the system.
5. It was inconvenient to work with the system.

Question 2. To what extent are you satisfied with the result?

1. I fully agree with the result obtained from the system.
2. I almost fully agree with the result obtained from the system.
3. I agree only to a small extent with the result obtained from the system.

4. I am doubtful about the result obtained by the system.
5. I am not satisfied with the result obtained from the system.

Question 3. To what extent do you understand the obtained result?

1. The result is easily understandable (according to the information I gave to the system).
2. I am able to see some correspondence between the result and the information I gave to the system.
3. It is difficult for me to understand the result on the basis of the information I gave to the system.

Question 4. How quick did you obtain the result?

1. The result was obtained quickly.
2. The result was obtained quickly enough.
3. The result was obtained not quickly but in reasonable a time.
4. It took a long time to obtain the result.
5. It took a very long time to receive the result.

Question 5. Was the system useful to you?

1. Work with the system helped me more clearly see what I want (or prefer).
2. After working with the system I began to better understand what I want (or prefer).
3. Work with the system did not help me in understanding my preferences.

Question 6. Would you like to use the system for real choice?

1. I think that such a system should be used before making a real choice.
2. I admit that such a system may be useful in making a real choice.
3. I don't think such a system would be useful in making a real choice.

APPENDIX 4: QUESTIONNAIRE ON SUBJECT PROCESS

Please answer the following questions:

1. What was your strategy in constructing your rule base with VP-EXPERT?
2. How much time did the VP-EXPERT rule base development take?
3. Did you modify your rule base after checking it with the 18 cases above?
   If yes, were there quite a few changes, or just minor changes?
4. Was the construction of the rule base an easy thing to do?
5. Was it difficult to verify the rule base?