

Problems of decision rule elicitation in a classification task

Alexander I. Mechitov

Russian Academy of Sciences, Moscow, Russia

Helena M. Moshkovich

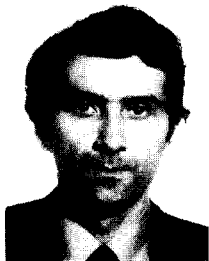
Russian Academy of Sciences, Moscow, Russia

David L. Olson

Texas A&M University, College Station, TX 77843, USA

Intelligent decision support requires knowledge elicitation processes. Two primary approaches for knowledge elicitation in a multiattribute classification task are 1) direct elicitation of decision rules in the form of productions, and 2) classification of multiattribute objects by an expert as a basis for development of the underlying decision rules. This study reports an experiment using a simple classification task, to compare these two forms of knowledge elicitation. Relative consistency and complexity of the resulting rule bases are analyzed. System CLASS was used as a tool for the second approach, as well as a means of analysis for the first approach. It was found that it was easier for subjects to accomplish the task using object classification than it was to formulate production rules directly. High degrees of inconsistency and incomplete rule bases resulted when there was no computer aid for the process of knowledge elicitation.

Keywords: Expert knowledge acquisition; Multiattribute models; Classification tasks

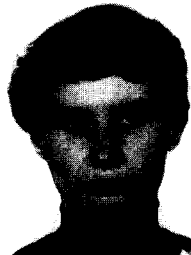


Alexander I. Mechitov received his M.S. in Management Science (1978) from Moscow State University. He holds his Ph.D. in MIS (1987) from the Institute for System Studies of the USSR Academy of Sciences. He is currently Associate Professor in the Decision Sciences Department at that same Institute. Dr. Mechitov's primary research interests include multicriteria decision making, decision support systems and expert systems, and risk analysis. He has been involved in projects including gas pipeline selection, site selection, and diagnostic expert systems in medicine and geology. Publications include over 45 articles in leading Russian journals, proceedings of international conferences, and two books.

Correspondence to: Dr. David L. Olson, Department of Business Analysis & Research, Texas A&M University, College Station, TX 77843, USA.

1. Introduction

Classification is a multiattribute problem type of great importance. A feature of classification tasks is that it is not necessary to rank all alternatives, but only to assign them to a small number of decision groups. Usually these decision groups are ordered, reflecting different degrees of attribute quality. For example, a manager using multicriteria estimations of research and development projects must divide the projects into the two categories of "accepted" and "rejected"; a consumer may desire to divide available products



Helen M. Moshkovich received her M.S. in MIS (1973) from the Moscow University of Management. She received her Ph.D. in Management Science (1984) from the Institute for System Studies of the USSR Academy of Sciences. She is currently Associate Professor in the Decision Sciences Department of that Institute. Dr. Moshkovich's primary research interests include multicriteria decision making, human decision making behavior, and intelligent decision support systems. She has been involved in projects such as R&D planning, oil technology selection, and strategic decision support. She has published over 40 articles in leading Russian journals, as well as in *Organizational Behavior and Human Decision Processes*, *Acta Psychologica*, and two books.



David L. Olson received his Ph.D. in Business Administration (1981) from the University of Nebraska-Lincoln. He is currently Associate Professor in the Department of Business Analysis and Research at Texas A&M University. Dr. Olson's primary research interest is in multicriteria decision making. He has published in journals such as the *European Journal of Operational Research*, *Journal of the Operational Research Society*, *Decision Sciences*, and other journals. He is coau-

thor of one book.

on various levels of quality, perhaps based upon input from a consumer's guide, opinions of others who have used the product, or on past personal experience; economists might use market indicators to classify market trends as increasing, neutral, or decreasing.

To provide effective decision support in such tasks, it is necessary to form a set of classification rules on the bases of decision maker or expert knowledge. There has been active research in knowledge acquisition techniques in general [1], including formal techniques such as machine rule induction. 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. Viewing approaches to elicitation of such rules, we see two main approaches. The first, and most popular, approach is to directly ask the expert to enter his or her implicit rules in the form of productions (see, e.g. [22]). In the second approach experts are asked to enter real examples from their practice, or to directly classify some objects, and then attempt to derive some interpretation of the underlying rules ([2],[7],[13],[21]).

While a great deal of attention has been given to knowledge acquisition ([8],[10],[20],[21]), more effort has been concentrated on development of shell systems, containing easy to use tools for entering information in the form of rules. Shells usually have the ability to explain why conclusions are made. The primary control mechanism to assure consistency is for the system to be tested on a number of examples ([3],[6]). It is assumed that the expert creates a comprehensive system of rules which will be useful to a less qualified specialist, but we can see that this approach is based primarily upon past experiences. The knowledge base may well be incomplete, as an answer for all possible alternatives may not be present. In that case, the knowledge base would be fragmentary. Furthermore, it has been known that a knowledge base can contain different answers for the same set of criteria values [20].

The second approach is implemented in a number of systems (see, e.g. [5],[9],[13]) where experts are asked to classify complex and multiattribute objects. This form of human judgment has been thoroughly investigated by psychologists ([4],[12],[15],[19]). Those studies found that classification accuracy was influenced to a large extent

by task parameters such as the number of criteria, the number of possible values on each criterion scale, and the number of classes into which alternatives are to be categorized. Moreover, it was obvious that when classifying different alternatives, people (even experts) often make mistakes, and need some support in classifying alternatives.

While we consider productions to be the most natural approach for humans to use in verbalizing their rules, we know of no attempts to investigate the cognitive effort required. Therefore, we developed an experimental design to analyze this form of human judgment in a simple classification task and compared the results with those produced with the help of the decision support system CLASS ([11],[13]). CLASS is oriented toward eliciting expert knowledge in the form of multiattribute object classification. Subjects were asked to introduce rules in the form of productions for a classification task with five attributes (three or four possible values each), with four class choices. The same task was solved with the system CLASS, providing a basis for estimation of completeness and contradictions as well as comparison of rule system complexity.

CLASS aids decision makers and/or experts in the task of classification of all possible alternatives, given a set of criteria with measures on discrete ordinal scales. CLASS operates by presenting the decision maker with hypothetical alternatives (combinations of values on criteria scales) which the decision maker is asked to classify into the given set of decision categories. Through this process, possible decision maker inconsistencies are checked for after each decision maker input. If an inconsistency is identified, the decisions leading to the conflict are identified for the decision maker, who is asked to reconcile the inconsistency. A rational interview procedure is used, allowing identification of a complete set of classification rules while presenting the decision maker with only a subset of the possible number of alternative combinations. Decision makers have the ability to review and change the resulting decision rules.

In the next section, the system CLASS is described. The third section presents the experimental design, as well as indices for evaluating the results. The fourth section analyses results, followed by the conclusion.

2. Primary features of system CLASS

2.1. Task formulation

CLASS ([11],[13]) was designed for problems where a decision maker must assign a set of cases (or objects) to N decision classes. These classes are ordered in the sense that each object placed in the first class is preferable to all objects placed in the second class, and so on. Each object can be characterized by values on each of Q criteria. Values on criterion scales are presented to the decision maker in verbal form. The decision maker orders each criterion scale. As there are Q criteria, and each criterion has a given number of discrete values, we are able to form the set of all possible combinations of values in criteria space. A complete classification system is developed when an a priori construction of classification of all possible criteria space combinations is completed. When an experienced decision maker and a real decision context is used, this classification reflects the decision maker's rules, and can be used for categorization of alternatives (objects) for real situations.

Therefore, the initial information necessary to begin to work with the system consists of criteria with scales, and the decision classes. As stated above, all criteria have ordinal scales, and verbal descriptions of quality grades on these ordinal scales. Examples of such criteria and decision classes for the task of assessing job opportunities are given in Appendix 1. All hypothetical combinations of criteria values are formed as a Cartesian product of criteria scales.

2.2. Information points

It is possible to accomplish the task of classification by having the decision maker directly classify all possible vectors of estimates. However, this is impractical even for a relatively small problem, which can involve a large number of possible states. Therefore, a special procedure for elicitation of decision maker classification rules has been developed. Ordinal ranking of attribute scales and decision classes are used in this method, imposing evident constraints on possible alternative classification, because an alternative which dominates another alternative cannot be assigned to a class worse than the alternative it

The following job opportunity is under consideration:

1. Type of the job position is good enough (in field).
2. Job is located far away.
3. The salary is on the average level.
4. There are minimal (almost none) possibilities for training.
5. There are moderate possibilities for promotion.

POSSIBLE ANSWERS:

1. This job is very attractive.
2. This job is acceptable.
3. This job is acceptable if there is nothing better.
4. This job is unacceptable.

YOUR ANSWER:

Fig. 1. Visualization of the situation and menu of possible answers.

dominates, while a dominated alternative cannot be assigned to a better class than its dominating alternative. It is attractive to classify as many alternatives as possible by logical rules inferred from previous classifications given by the decision maker.

The procedure builds the required classification with only a limited number of questions asked of the decision maker. The most "informative" (potentially capable of classifying of the maximum expected number of alternatives) alternatives are presented to the decision maker. More details of the procedure and estimation of the classification rules are given in Larichev and Moshkovich [15]. The system calculates the most informative alternative, and explains it as demonstrated in Figure 1.

2.3. Contradictions

People can make judgmental errors for a variety of reasons. Thus it is necessary to have tools to detect and correct these possible errors. With CLASS, errors can be detected by use of outranking criteria scales and classes. As stated above, a dominated alternative may not be assigned to a better class than an alternative which dominates it. Therefore, if such an alternative is presented to a decision maker, and the decision maker chooses an inappropriate class, we are able to identify a contradiction.

1. Type of the job position is almost ideal.
2. Location of the job is very convenient.
3. The salary is rather high.
4. There are minimal (almost none) possibilities for training.
5. There are almost no possibilities for promotion.

THE SITUATION IS ESTIMATED AS: 1. This job is very attractive.

1. Type of the job position is almost ideal.
2. Location of the job is very convenient.
3. The salary is rather high.
4. There are normal possibilities for training.
5. There are almost no possibilities for promotion.

THE SITUATION IS ESTIMATED AS: 2. This job is acceptable.

The second situation is more preferable to than the first one according to their estimates. It must be put to a not less preferable class than the first situation. Analyze the inconsistency and assess both situations again.

PRESS ANY KEY TO CONTINUE

Fig. 2. Screen presenting contradictory input.

The system checks each decision maker assignment of an alternative to determine if this assignment is compatible with previous information. If the assignment is compatible, then the system spreads the information to other hypothetical alternatives on the basis of dominance relations. A new hypothetical alternative is then presented to the decision maker. This continues until all possible alternatives (as described by combinations of criteria characteristics) have been classified.

If a particular answer contradicts the knowledge previously developed, then the system informs the user, and gives the decision maker the option to either change the current response, or to analyze the contradiction. If the decision maker chooses to analyze the contradiction, the system displays relevant information as in Figure 2.

2.4. Introduction of decision rules

The procedure described above leads to construction of a complete and non-contradictory classification of all possible alternatives in the criteria space. However, it may take a significant amount of time, because the system in a priori unaware of the decision maker's rules. Sometimes, however, a decision maker may explicitly

know of some classification rules (usually the simplest kinds of rules, such as unacceptable performance on one or more criteria automatically result in assignment to the least attractive class). In this case, the system includes the ability to enter such a rule. Usually such rules are expressed in a conjunctive form. For instance, "a job with a poor salary is unacceptable". In this case, it is enough for the decision maker to enter an alternative with all of the best values on all criteria except for the criteria in question (number 3 – salary, in the example), and enter the least acceptable value for that criterion. The system then displays such an alternative, and if a decision maker marks the lowest class for this alternative, the system reflects the knowledge that all alternatives containing such a value on the criterion in question will be classified in the lowest class (in the example, the fourth class – unsuitable job).

2.5. Analysis of decision rules

Once the required classification is built, it can be used to identify the appropriate class for any real alternative. Accomplishment of this task requires that the alternatives (in our example, jobs) be estimated by experts on the set of criteria. Then these alternatives are presented to the system as a set of combinations of values on the set of criteria. The system identifies the appropriate class for each alternative.

As the classification of all possible alternatives (combinations of estimates on criteria) is built, we can identify the boundaries between decision classes. In each class, we can determine the alternatives which do not dominate any other alternative in this class. We also can identify alternatives not dominated by any other alternative in this class. Therefore, all other alternatives in this class are placed between them (these other alternatives dominate the former, and are dominated by the latter). That is why these subsets of alternatives within a class are referred to as "borders". The first subset is the "lower border" of the class, while the second subset is the "upper border" of the same class.

After classification is completed, the system defines the borders in each class (see Figure 3). Upon request, the system can present these borders to the user for information or analysis. If

Vectors defining borders of classes consist of the category for each of the five criteria considered. Each criterion had three categories (with 1 being the best), except for criterion 3, which had 4 categories.

Vector in the upper border of class 1
11111

Vectors in the low border of class 1
22123 32113

Vectors in the upper border of class 2
11211 31121

Vectors in the low border of class 2
22223 22313 32123 32213

Vectors in the upper border of class 3
11321 31221 31311

Vectors in the low border of class 3
22323 32223 32313

Vectors in the upper border of class 4
11131 13111 11411 31321

Vectors in the low border of class 4
33433

Fig. 3. Presentation of border elements.

desired, the decision maker can change some of these borders. In this case, the system inserts this new information into the classification, checks for contradictions, provides the decision maker information to eliminate any such contradiction, and maintains the completeness of classification. The new set of borders can then be presented to the decision maker. The elements in the borders can be viewed as rules which describe the strategy of alternative classification.

3. Experimental design

3.1. Task

The intent of the experiment was to analyze subject direct formulation of rules (in the form of productions) and to compare with the results when using CLASS. The task was evaluation of job opportunities similar to possible offers for the subject group. Subjects were 18 senior undergraduate students, all in the job seeking stage. Five attributes, JOB TYPE, LOCATION, SALARY,

TRAINING, and PROMOTION, were used to characterize each job. Each attribute had a three point ordinal scale except SALARY, which had a four point ordinal scale (see Appendix 1). Students were familiar with the task context, and had dealt with a similar problem context as part of other assignments applying other multicriteria techniques. Therefore, the students could be considered as experts of a sort for this task. The classification task is simple, does not require decomposition of complicated structures, and may be solved by enumeration. The simplicity of the task allowed us to concentrate on the elements of human judgment required.

At the time of the experiment, students were taught the basics of expert systems, and production rules as well as other knowledge base construction approaches were presented to them. Students worked with CLASS, which required them to focus on comparison of job attributes. Four classification categories were used: job is very attractive, job is acceptable, job is unacceptable if there is nothing better, and the job is unacceptable. Note that using the CLASS system results in a comprehensive set of rules, but the subjects have no direct realization that rules are being developed. The subjects were then assigned the task of manually constructing a set of production rules for classifying job offers (instructions provided to students are given in Appendix 2).

3.2. Experimental data processing

To analyze the results of the experiment, it is necessary to develop measures for the constructed sets of rules for both cases. The first logical requirement for constructing a set of rules is completeness (all possible combinations of attributes should be classified by the set of rules). The set of rules should also be noncontradictory (there should be a guarantee that only one class assignment could result for a given set of attribute values). With CLASS, we know that the resulting set of rules is complete and noncontradictory. Direct development of a set of production rules could very easily include gaps and contradictions.

To measure rule sets directly developed by subjects for these characteristics, the following procedure was developed. After subjects introduced a set of production rules for the task, the

authors returned to the CLASS system, introducing the production rules directly given by the subjects. This was possible, as CLASS has a regime for introduction of rules. For example, if the subject gave the rule: "If salary is on the average level and the job type is almost ideal, then class 2". To introduce such a rule within CLASS, we reason as follows: For an alternative to be in the second class, it is sufficient to have an estimate of 1 or 2 on the first criterion (JOB TYPE) and an estimate of 2 for the third criterion (SALARY). In this case, any assigned value on the other criteria will not influence the classification. Therefore, we enter the vector 11211 and assign this to the second class. The system shows this vector to be the best possible alternative influenced by this rule. Next we enter the vector 23233, and assign this vector to class 2. The system displays this as the worst possible alternative influenced by this rule. Therefore, all other alternatives having the required estimates on the first and third criteria will be classified in class 2. Therefore, to introduce this rule we need to classify two specially constructed alternatives. Other rules can require entering more vectors (and in special cases one is sufficient), but it is always possible to enter any production rule into CLASS. The following parameters were estimated for rule bases directly developed by subjects: *the number of alternatives needed to be entered* into the system to complete the classification; *the number of contradictory rules* (if any); *the number of alternatives not classified* (out of the initial 324) by this set of rules.

Each subject's work with CLASS enabled us to identify *the number of alternatives considered* by each subject in building the full classification. In addition, we know *the number of changes* in subject answers (due to contradictory answers encountered in the process of using the system), and *the number of alternatives in the borders*. All of these parameters are fixed by the system. The number of elements in the borders characterize the minimum number of alternatives by which a decision maker is guaranteed of classifying all 324 possible attribute value combinations. Therefore, these parameters measure the effectiveness of CLASS.

In Larichev and Moshkovich [15], it was stated that to characterize human abilities in different cognitive tasks, it is necessary to evaluate the

complexity of the decision strategy (or rules) used by subjects. It is not that people may not have simple rules, but the data must show that people are able to reliably use rather sophisticated rules if they want to. In that same work, it was proposed that this complexity could be measured by *the number of rules used*, and by *the number of attributes used* in each rule. We will try to estimate these parameters. This would be easy to do for a freely built system of rules (as they are formulated as rules on attributes).

For the rule set developed within CLASS, the problem is a bit more complicated. It is sufficient to use the borders of the classes to classify each alternative. Moreover, it is sufficient to use only the upper borders of classes (see Figure 3). If we have an alternative, we first compare it with the elements of the upper border of the fourth class. If this alternative is equal to or dominated by any of the members of this border, then this alternative belongs to the fourth class. If not, we test the border of the third class. If the analyzed alternative is equal to or dominated by any element of the upper border of the third class, then the alternative belongs to the third class. If not, we try the second class, and so on. Therefore, to characterize the number of rules used, we can take *the number of elements in the upper borders* of the classes.

However, as is shown in Figure 3, each element of the upper border contains the same number of attributes (five in our case). Therefore, to characterize the number of attributes used in these rules, we must further analyze these elements. Examine the example of the upper borders given in Appendix 3. Each upper border consists only of one element, and in these elements, all values but one are equal to 1. This means that the subject used only one attribute (salary in this case) to formulate his policy (the value on this attribute defines the class of the alternative). Therefore, we can calculate the number of attributes used in each rule (each element of the border is one rule) by calculating *the number of elements differing from the first (best) values* on corresponding attributes. We are thus able to assume that we have means of measurement that will allow us to characterize the complexity of applied strategies for constructing the rule set in both approaches used in this experiment.

4. Analysis of results

In the experiment, subjects first worked with CLASS, and then directly developed production rules. This may cause initial concern about order effects. However, we argue that any advantage that would exist would be for the second treatment, which in this case was the rule-based approach. This is really immaterial, because subject use of the CLASS system is not at all like developing the rule base directly. CLASS guides the subjects to focus on tradeoffs among alternatives. By working with CLASS first, the subjects were able to gain familiarity with the problem domain prior to development of rules. In Table 1, data for 18 subjects working with CLASS is presented. As we see, the number of alternatives presented to subjects to construct the full classification is rather large (about 100), which is two to three times more than the elements in the borders (which would be the minimal number required for subject consideration). However, we of course do not know what alternatives are going to end up in the borders ahead of time. Post experiment interviews showed that subjects did not use the ability to impose specific rules (probably because they were not familiar with the system). But in

real applications, simple rules can be entered which would reduce the number of alternatives users would be required to consider. We would still expect the number of alternatives considered to be about twice the number of border elements (also found to be typical in practical cases where CLASS was applied [14].

We can also see that there were some errors in subject answers which required correction. But the number of errors is rather low. This, as well as the number of alternatives presented, show that the task of alternative classification was not too cognitively difficult, and produced reliable results.

The analysis of borders show that the number of rules may be rather large (up to 43), and contain rules using quite a few attributes. Although not many subjects used rules reflecting five attributes, almost all used three or four attributes in some of their rules. All subjects had many simple rules, using one or two attributes.

Only one subject (number 11) used a very simple system of four rules, with one attribute. Usually such rules are formulated as follows: "If SALARY is poor, the alternative is to be categorized as belonging to the fourth class". Payne [18] called such rules "non-compensatory", meaning

Table 1
Subject responses using CLASS

	Alternatives presented to DM (324 max)	Changes in answers	Vectors to get full classification	Elements in upper borders	Number of rules with attribute values not equal to 1				
					1	2	3	4	5
1	153	0	57	30	2	17	7	1	2
2	93	2	32	19	3	7	6	1	1
3	111	0	45	27	5	13	5	3	-
4	69	1	17	10	6	1	2	-	-
5	109	1	31	17	4	7	3	2	-
6	142	0	86	43	3	12	11	10	6
7	107	0	13	5	4	1	-	-	-
8	110	1	32	18	6	3	8	-	-
9	114	3	55	28	3	8	7	6	2
10	110	0	23	15	6	3	5	-	-
11	103	0	8	4	4	-	-	-	-
12	110	0	27	14	4	3	2	3	1
13	89	0	28	15	6	4	4	-	-
14	82	0	44	27	6	12	7	1	-
15	101	12	55	27	2	9	8	6	1
16	93	1	20	10	4	3	2	-	-
17	107	2	54	32	6	14	7	4	-
18	157	0	60	33	5	17	8	2	-

Table 2
Measures of subject responses with their own rules

Directly Developed Rules					
N/N	Number of Rules	Alternatives entered to introduce rules	Number of unclassified elements (of 324)	Number of contradictions in rules	Attributes used in rules
1	9	18	256	0	3
2	11	11	17	6	5
3	10	15	0	4	3
4	9	28	0	0	3
5	73	93	0	5	4
6	25	29	99	0	3
7	17	17	54	0	3
8	8	13	168	1	4
9	10	90	0	3	4
10	16	16	278	0	5
11	8	27	84	2	4
12	11	19	58	4	3
13	7	9	0	0	3
14	6	13	174	0	5
15	7	11	54	2	3
16	39	17	12	0	4
17	18	24	59	0	5
18	8	10	187	2	5

that poor quality on one attribute cannot be compensated for by any advantage gained on the other attributes.

Table 3
Complexity of the rule used in subject rule base

N/N	Number of rules by attributes					Type of rules used
	1	2	3	4	5	
1	-	-	9	-	-	production
2	3	1	4	3	-	production
3	7	3	-	-	-	production
4	6	3	-	-	-	production
5	1	-	-	72	-	situations
6	-	-	25	-	-	situations
7	6	6	5	-	-	production
8	3	-	3	2	-	production
9	2	4	2	1	-	production
10	-	-	-	-	16	situations
11	1	-	2	5	-	production
12	2	5	4	-	-	production
13	3	4	-	-	-	production
14	-	-	-	-	6	production
15	2	5	-	-	-	production
16	3	-	-	-	36	situations
17	5	-	-	-	13	situations
18	-	1	5	1	1	production

In Tables 2 and 3, data for the subject built rule base is presented. We note that one column, headed "Type of rules used" was not discussed in the previous section. This is because in the assignment, subjects were shown the example of the rules to appear in their rule base. Nevertheless, there was no special limitation on the form of their formulation. As a result, some of the subjects formulated their rules by enumerating possible attribute values, assigning each to the appropriate class (see the example of the rule base in the form of a production in Appendix 4, and the form of enumeration of possible combinations in Appendix 5).

When subjects tried to enumerate the combinations of attribute values, they used the same type of judgment as when working with CLASS, but without the aid of the system. As these cases demonstrated a different approach in introducing production rules, in the last column of Table 3 the word "productions" is used to indicate the more traditional presentation of rules, and the word "situations" is used for enumeration of possible combinations of attribute values.

We first analyze data for those who used "productions". We can see that the number of rules

formulated for the rule base is less than in the previous case (up to 17 rules). At the same time, only five subjects classified all elements with the help of these rules. All other subjects had rule bases which could categorize only a part of the possible number of alternatives. Only five of the 13 subjects developed a non-contradictory rule base, despite the small numbers of rules used. We can also see that rule complexity is much less than that obtained through CLASS. The majority of subjects used only three or four attributes, and those who avoided contradictions focused on one, or at the most, two attribute rules. We conclude from this that CLASS provides a useful consistency check, and that without such a check, sets of production rules may well include gaps and inconsistencies.

Looking at the five subjects who used “situations” for constructing rule bases, we see that only one had contradictory rules. All others were consistent despite the rather large number of rules used. But none of these five was able to cover all possible combinations of attribute values. We consider these five subjects to have used rather complex rules (three, four, or five attributes considered). Thus, even when subjects adopt the idea of generating and classifying all possible alternatives, they are not always able to accomplish this, even for tasks involving a small number of attributes.

We draw several conclusions. When people are asked to introduce rules, they tend to formulate a small number of these rules, usually covering the most common and simple situations they expect. Even so, they tend to introduce contradictory rules, without noticing. These rules are usually based on a small number of available attributes, not necessarily reflecting attribute significance, but rather trying to avoid cognitive overload [16]. It is difficult for people (and even for analysts) to notice incompleteness and contradictions in a rule base. Moreover, the experimental data shows that although use of productions is a very natural way for people to formulate implicit rules, it still is not an easy task. The data shows that when classifying multiattribute alternatives (“situations”), people tend to make fewer mistakes than when using some sets of simple rules. Therefore, we can assume that that form of rules elicitation through classification of possible combinations of attribute values in some pre-defined

domain (knowing attributes and possible values) is preferable to asking people to directly formulate rules.

5. Conclusion

The problem of knowledge elicitation is very important in decision support. Rule bases are a necessary element in intelligent decision support systems. The results of our experiment show that rule base construction requires a great deal of assistance. It is very important to use means of knowledge elicitation that are comfortable for decision makers, and to check for rule base completeness and contradiction. Assistance provided by a system such as CLASS was shown to be of great value in allowing subjects to develop a valid and diverse set of rules. Working with the system CLASS required classification of a large number of hypothetical alternatives (generally 15 to 30 minutes), but was worth the effort. This matches findings in actual applications with real experts [13],[14]. As a rule, experts are more at ease when they are asked to evaluate (classify) concrete examples than they are in formulating general inference rules, considering all assumptions and exceptions.

This can be explained by different phenomena. Firstly, consideration and classification of definite situations is a more common task for humans than formation of a large set of rules, which are to be mutually dependent and logically consistent. Secondly, it is known that it is difficult for humans to verbalize their knowledge [17].

At first glance, CLASS seems appropriate for small tasks with clearly identifiable order of value. However, it can be used for a broader set of tasks as well. Several medical diagnostic knowledge bases with nonordinal categories have been constructed using CLASS [14].

CLASS does not provide a cure-all for knowledge elicitation, but it can aid some stages of developing a real knowledge base. In practice, the number of attributes and their possible values can be very large. For instance, in medical diagnostic problems, the physician may consider up to forty attributes with three to five value scales on each attribute. Therefore, the total number of states to be classified may be over hundreds of millions, and the problem of complete and non-

contradictory classification becomes intractable. Usually, however, people do not use all of these parameters simultaneously (that would exceed human information processing capability). People tend to initially analyze one group of attributes, and depending on the result of this initial analysis, seek additional information on the next group of attributes.

Therefore, the logical way to deal with large problems is to try to decompose them into smaller problems, solve these smaller problems, and synthesize the results. A system such as CLASS can be very useful in eliciting knowledge for solving these relatively small subproblems, and other elicitation techniques can be used for synthesis of the knowledge base and testing results. We feel that it would be interesting to attempt combination of CLASS and some production shell in order to provide flexible transfer from one elicitation technique to another in constructing knowledge bases.

References

- [1] T.A. Byrd, K.L. Cossick and R.W. Zmud, A synthesis of research on requirements analysis and knowledge acquisition techniques, *MIS Quarterly* 16:1 (1992) 117–138.
- [2] W. Clancey, Heuristic classification, *Artificial Intelligence* 27 (1985) 289–350.
- [3] R. Davis and D.B. Lenat, *Knowledge-Based Systems in Artificial Intelligence*, McGraw-Hill, NY (1982).
- [4] R. Dawes and B. Corrigan, Linear models in decision making, *Psychological Bulletin* 81 (1974) 95–106.
- [5] E.A. Feigenbaum, Expert systems in the 1980s, in *State of the Art Report on Machine Intelligence*, A. Bond, ed., Maidenhead: Pergamon-Infotech (1981).
- [6] W.A. Gale, Knowledge-based knowledge acquisition for statistical consulting system, *International Journal of Man-Machine Studies* 26 (1987) 55–64.
- [7] M.D. Grover, A pragmatic knowledge acquisition methodology, *IJCAI-83* (1983) 436–438.
- [8] A. Hart, Knowledge elicitation: Issues and methods, *Computer-Aided Design* 17:9 (1985) 455–462.
- [9] F. Hayes-Roth, D.A. Waterman and D.B. Lenat, *Building Expert Systems*, Addison-Wesley (1983).
- [10] F. Hickman, Knowledge acquisition: The key to success for commercial expert systems, *Proceedings of the International Conference on Knowledge Based Systems*, London, July (1986), 205–214.
- [11] O.I. Larichev and H.M. Moshkovich, Task of direct classification in decision making, *Doklady Akademii Nauk* 287:6 (1986) 567–570 (in Russian).
- [12] O.I. Larichev and H.M. Moshkovich, Limits to decision making ability in direct multiattribute alternative evaluation, *OBHDP* 42 (1988) 217–233.
- [13] O.I. Larichev and H.M. Moshkovich, Decision support system “CLASS” for R&D planning, in *Proceedings of the First International Conference on Expert Planning Systems*, Brighton, England (1990) 227–232.
- [14] O.I. Larichev, H.M. Moshkovich, E.M. Furems, A.I. Mechitov and V.K. Morgoev, *Knowledge Acquisition for the Construction of Full and Contradiction Free Knowledge Bases*, iec ProGAMMA, Groningen, The Netherlands (1991).
- [15] O.I. Larichev, H.M. Moshkovich and S.B. Rebrik, *Acta Psychologica* 68 (1988) 171–182.
- [16] G. Miller, The magical number seven plus or minus two: Some limits on our capacity for processing information, *Psychological Review* 63 (1956) 87–97.
- [17] R.E. Nisbett and T.G. Wilson, Telling more than we can know: Verbal reports on mental processes, *Psychological Review* 37 (1977) 231–259.
- [18] J.W. Payne, Task complexity and contingent processing in decision making: An information search and protocol analysis, *Organizational Behavior and Human Processes* 16 (1976) 366–387.
- [19] R.H. Phelps and G. Shanteau, Livestock judges: How much information can an expert use? *Organizational Behavior and Human Performance* 21 (1978) 209–219.
- [20] R.G. Vedder, PC-based expert systems shells: Some desirable and less desirable characteristics, *Expert Systems* 6:1 (1989) 28–42.
- [21] G. Wright and P. Ayton, Eliciting and modelling expert knowledge, *Decision Support Systems* 5 (1987) 13–26.
- [22] VP-Expert, *Rule-Based Expert System Development Tool*, Paperback Software International (1987).

Appendix 1. Criteria for evaluation of job opportunities evaluated by subjects

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 very convenient.
2. Location of the job is in some distance.
3. Job is located far away.

Criterion 3. Salary

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

Criterion 4. Possibilities for training

1. There are nice possibilities for training.
2. There are normal possibilities for training.
3. There are minimal (almost none) possibilities for training.

Criterion 5. Possibilities for promotion

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

DECISION CLASSES

1. This job is very attractive.
2. This job is acceptable.
3. This job is rather poor.
4. This job is unacceptable.

Appendix 2. Assigned instructions for directly developed rule base

The purpose of this assignment is to classify a set of jobs into categories by objective. Generate a set of characteristics that can automatically be applied to other alternative choices (either to be encountered in the future, or when there are many such alternatives available).

The decision context we will use is job selection, primarily because it is a topic most of you know a lot about. We will focus on the five criteria we used in prior assignments (pay, location, job type, promotion, and training). Each of these criteria will be represented by three (or in the case of salary, four) categories. You are to formulate a set of rules to categorize every possible combination of criteria classes. This set of rules will be used to categorize jobs into one of the following four classes:

- Class 1: a very attractive job
 Class 2: an acceptable job
 Class 3: a poor job, acceptable in a pinch
 Class 4: an unacceptable job

Production rules of the following type are to be formulated:

- if SALARY is poor, then Class 4;
 if JOB TYPE is inappropriate and LOCATION is far away and SALARY is not high, then CLASS 4;
 if there is not excellent PROMOTION potential, then Class 4;

You may use criterion and value numbers to represent the rules, such as:

- if C3 = 4 then Class 4;
 if C1 = 3 and C2 = 3 and C3 = (2 or 3 or 4) then Class 4;
 if C5 = 2 or C5 = 3 then Class 4;

Appendix 3

Vectors defining borders of classes consist of the category for each of the five criteria considered. Each criterion had three categories (with 1 being the best), except for criterion 3, which had 4 categories.

BORDERS FOR SUBJECT 11

Number of elements in the upper border of class 1 is equal to 1:

11111

Number of elements in the lower border of class 1 is equal to 1:

33133

Number of elements in the upper border of class 2 is equal to 1:

11211

Number of elements in the lower border of class 2 is equal to 1:

33233

Number of elements in the upper border of class 3 is equal to 1:

11311

Number of elements in the lower border of class 3 is equal to 1:

33333

Number of elements in the upper border of class 4 is equal to 1:

11411

Number of elements in the lower border of class 4 is equal to 1:

33433

Appendix 4. Example of a subject's production rules

If $C_3 = (1 \text{ or } 2)$ and $C_1 = (1 \text{ or } 2)$ then Class 1;
 If $C_2 = (1 \text{ or } 2)$ then Class 1;
 If $C_1 = 3$ and $C_2 = 3$ then Class 4;
 If $C_1 = 2$ and $C_3 = 3$ then Class 3;
 If $C_1 = 3$ and $C_3 = 4$ then Class 4;
 If $C_1 = 2$ and $C_2 = 2$ and $C_3 = (2 \text{ or } 3)$ and $C_4 = 2$ and $C_5 = 2$ then Class 3;
 If $C_1 = 1$ and $C_3 = 3$ then Class 2;
 If $C_3 = 4$ then Class 4;

Appendix 5. Example of a subject's situation rule base

If $C_2 = 3$ then Class 4
 If $C_3 = 4$ then Class 4
 If $C_4 = 3$ then Class 4
 If $C_1 = 1$ and $C_2 = 2$ and $C_3 = 3$ and $C_4 = 2$ then Class 3
 If $C_1 = 1$ and $C_2 = 1$ and $C_3 = 3$ and $C_4 = 2$ then Class 3
 If $C_1 = 1$ and $C_2 = 2$ and $C_3 = 3$ and $C_4 = 1$ then Class 2
 If $C_1 = 1$ and $C_2 = 1$ and $C_3 = 3$ and $C_4 = 1$ then Class 2
 If $C_1 = 1$ and $C_2 = 2$ and $C_3 = 2$ and $C_4 = 2$ then Class 2
 If $C_1 = 1$ and $C_2 = 1$ and $C_3 = 2$ and $C_4 = 2$ then Class 2
 If $C_1 = 1$ and $C_2 = 2$ and $C_3 = 2$ and $C_4 = 1$ then Class 2
 If $C_1 = 1$ and $C_2 = 1$ and $C_3 = 2$ and $C_4 = 1$ then Class 2
 If $C_1 = 1$ and $C_2 = 2$ and $C_3 = 1$ and $C_4 = 2$ then Class 1
 If $C_1 = 1$ and $C_2 = 1$ and $C_3 = 1$ and $C_4 = 2$ then Class 1
 If $C_1 = 1$ and $C_2 = 2$ and $C_3 = 1$ and $C_4 = 1$ then Class 1
 If $C_1 = 1$ and $C_2 = 1$ and $C_3 = 1$ and $C_4 = 1$ then Class 1

If $C_1 = 2$ and $C_2 = 2$ and $C_3 = 3$ and $C_4 = 2$ then Class 3
 If $C_1 = 2$ and $C_2 = 1$ and $C_3 = 3$ and $C_4 = 2$ then Class 3
 If $C_1 = 2$ and $C_2 = 2$ and $C_3 = 3$ and $C_4 = 1$ then Class 2
 If $C_1 = 2$ and $C_2 = 1$ and $C_3 = 3$ and $C_4 = 1$ then Class 2
 If $C_1 = 2$ and $C_2 = 2$ and $C_3 = 2$ and $C_4 = 2$ then Class 2
 If $C_1 = 2$ and $C_2 = 1$ and $C_3 = 2$ and $C_4 = 2$ then Class 2
 If $C_1 = 2$ and $C_2 = 2$ and $C_3 = 2$ and $C_4 = 1$ then Class 2
 If $C_1 = 2$ and $C_2 = 1$ and $C_3 = 2$ and $C_4 = 1$ then Class 2
 If $C_1 = 2$ and $C_2 = 2$ and $C_3 = 1$ and $C_4 = 2$ then Class 1
 If $C_1 = 2$ and $C_2 = 1$ and $C_3 = 1$ and $C_4 = 2$ then Class 1
 If $C_1 = 2$ and $C_2 = 2$ and $C_3 = 1$ and $C_4 = 1$ then Class 1
 If $C_1 = 2$ and $C_2 = 1$ and $C_3 = 1$ and $C_4 = 1$ then Class 1
 If $C_1 = 3$ and $C_2 = 2$ and $C_3 = 3$ and $C_4 = 2$ then Class 4
 If $C_1 = 3$ and $C_2 = 1$ and $C_3 = 3$ and $C_4 = 2$ then Class 4
 If $C_1 = 3$ and $C_2 = 2$ and $C_3 = 3$ and $C_4 = 1$ then Class 3
 If $C_1 = 3$ and $C_2 = 1$ and $C_3 = 3$ and $C_4 = 1$ then Class 3
 If $C_1 = 3$ and $C_2 = 2$ and $C_3 = 2$ and $C_4 = 2$ then Class 3
 If $C_1 = 3$ and $C_2 = 1$ and $C_3 = 2$ and $C_4 = 2$ then Class 3
 If $C_1 = 3$ and $C_2 = 2$ and $C_3 = 2$ and $C_4 = 1$ then Class 2
 If $C_1 = 3$ and $C_2 = 1$ and $C_3 = 2$ and $C_4 = 1$ then Class 2
 If $C_1 = 3$ and $C_2 = 2$ and $C_3 = 1$ and $C_4 = 2$ then Class 2
 If $C_1 = 3$ and $C_2 = 1$ and $C_3 = 1$ and $C_4 = 2$ then Class 2
 If $C_1 = 3$ and $C_2 = 2$ and $C_3 = 1$ and $C_4 = 1$ then Class 1
 If $C_1 = 3$ and $C_2 = 1$ and $C_3 = 1$ and $C_4 = 1$ then Class 1