Pergamon

0969–6016(94)E00010–P

An Approach to Ordinal Classification Problems

O. I. LARICHEV and H. M. MOSHKOVICH

Institute for System Analysis (Moscow), Russia

An ordinal classification task is defined. An approach to the construction of full ordinal classification on the basis of a decision maker's knowledge is proposed. It allows elicitation of information (or knowledge) in a natural form for the decision maker (through qualitative attribute scales and verbal descriptions of decision classes). It provides verification of the received judgments for consistency and possibilities for corrections and modifications of the elicited classification rules. Problems with obtaining valid judgments from people in ordinal classification tasks are discussed. The decision support system ORCLASS, developed on the basis of the proposed approach, is described.

Key words: multi-objective decision, man–machine systems, behavioral, expert

INTRODUCTION

Along with multi-criteria choice problems, people may face multi-criteria classification problems. A feature of classification tasks is that it is not necessary to rank the alternatives, but only to assign them to a small number of decision groups. Quite often these classes may be rank-ordered, reflecting different degrees of quality. In this sense the alternatives assigned to the first decision class are better than those assigned to the second class, etc.

Examples of such tasks may be found in different areas of human decision making: an R&D program leader, who decides which projects to incorporate into the program on the basis of their characteristics; a physician, who determines the severity of the disease on the basis of the patient's symptoms; an engineer, who detects the possibility that a definite block in a complicated technical system is the cause of malfunction on the basis of a set of indicators' data; an editor, who decides on the manuscripts according to the referees' evaluations, and others.

Usually one does not need to choose the best variant in them: the task is to categorize each object. So the final goal in such tasks is to distribute alternatives among classes of decision: to include or not this R&D project in a program; to accept, correct or reject the manuscript submitted to a scientific journal; to define the appropriate kind of disease for a patient, and so on.

Tasks with ordered classes were called tasks of ordinal classification by Larichev and Moshkovich (1986).

Though classification problems are quite common in real life, they have been paid much less attention in the theory of multi-criteria decision making than that of ranking or choice problems. Roy (1985) used the term 'segmentation procedures' for procedures suited to assign each action to one of several 'categories'. In recent classification and definition of main types of decision tasks given by Bana e Costa (1993), these problems are referred to as problems in which absolute evaluation of multi-attribute alternatives is needed, but most of the works were connected with nominal classes. De Montgolfier and Bertier (1978) described a procedure of generalizing sets of attribute values into ordinal categories of a more general criterion upon the decision maker's judgment. This task may be viewed as that of ordinal classification. ELECTRE TRI (Yu, 1992) and ROBOT Technique (Bana e Costa, 1992) are the recent examples of an approach to ordinal classification based on the idea to assign multi-attribute alternatives to ordered classes if it is found out that all the alternatives' values (separately) belong to the appropriate class.

The goal of the paper is to present another approach to the solution of ordinal classification problems, characterized by the following main features:

(1) it allows construction of complete classification of all possible objects in a criteria space, via classification of only part of them directly by a decision maker;

---

(2) it provides the possibility to detect and correct errors and inconsistencies in the decision maker's judgments;

(3) it makes possible the constructive analysis of the formed classification, and its modification in case of some changes.

Some mathematical and psychological basis for this approach is given.

THE PROBLEM FORMULATION

The problem under consideration may be presented in the following way (de Montgolfier and Bertier, 1978; Larichev and Moshkovich, 1986). A decision maker has a final set of N decision classes and must assign to them a set of cases (or objects). These classes are ordered for a decision maker (DM) 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 upon criterion scales are presented to the decision maker in verbal form. The decision maker orders each criterion scale from the most to the least preferable one (an example of criteria and classes for the case of manuscripts' evaluation is presented in the Appendix).

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 (Cartesian product of criterion values). 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 real alternatives (objects). For example, the editorial board can construct the complete classification in the formed criteria space (see Appendix), and use the result to categorize submitted manuscripts after their refereeing.

The task can be represented formally in the following way:

Given:

(1) $K = \{q_i\} i = 1,2,\ldots,Q$ – a set of criteria;

(2) ω_q – number of possible values on the scale of the qth criterion $(q \in K)$;

(3) $X_q = \{x_{iq}\}$ – a set of values for the qth criterion (the scale of the qth attribute); $|X_q| = \omega_q (q \in K)$;

(4) $Y = X_1 \times X_2 \times \ldots \times X_Q$ – a set of vectors $y^i \in Y$ of the following type $y^i = (y_1^i, y_2^i, \ldots, y_Q^i)$, where $y_q^i \in X_q$;

(5) $L = |Y| = \prod\limits_{q=1}^{Q} \omega_q$ – capacity of Y;

(6) N – number of ordered decision classes.

Needed: on the basis of a decision maker's preferences (judgments) to build a reflection F: $Y \Rightarrow \{Y_1\}, l = 1,2,\ldots,N$, such that $Y = \bigcup\limits_{l=1}^{N} Y_1$; $Y_l \cap Y_k = \phi$ if $k \neq 1$ (where Y_l – a subset of vectors from Y, assigned to the lth class).

A rather usual task, required from a decision maker in multi-criteria decision problems with verbal values, is that of rank-ordering of possible values for one criterion from the set K. As a result ordinal scales for criteria are formed, in which the first value x_{q1} upon the criterion $q(q \in K)$ is more preferable for a decision maker than the second value x_{q2} upon the same criterion, and so on. If we use natural numbers to denote values in the ordinal scale X_q for the qth criterion, we shall obtain a modified scale $B_q = \{1,2,\ldots,\omega_q\}$, where $b_{iq} < b_{jq}$, if x_{iq} is more preferable for the DM than x_{jq}. So, for each ordinal scale X_q we form the unique ordinal scale B_q, reflecting the DM's preference for values from X_q.

This information from a decision maker defines an irreflexive and transitive binary relation of strict preference (or dominance) P° on the set Y:

$$P^\circ = \{(y^i, y^j) \in Y \times Y \mid \forall q \in K \, b_{iq} \leq b_{jq} \text{ and } \exists q^\circ, \text{ such that } b_{iq^\circ} < b_{jq^\circ}\}.$$

In Table 1 four hypothetical manuscripts estimated upon the criteria from the Appendix are given.

Table 1. Data for four hypothetical alternatives

Alternatives (manuscripts)	Relation to journal's outline	Theoretical value	Practical value	Errors	Quality of text
#1	direct	middle	high	no	good
#2	relative	middle	high	some	good
#3	relative	middle	no	some	good
#4	relative	high	high	no	good

These alternatives may be presented through the following vectors according to the above described principle: vector $b_1 = (1,2,1,1,1)$ for alternative #1, vector $b_2 = (2,2,1,2,1)$ for alternative #2, vector $b_3 = (2,2,3,2,1)$ for alternative #3, and vector $b_4 = (2,1,1,1,1)$ for alternative #4. As we can see, alternative #1 dominates alternative #2 (as it has better values upon criteria one and four), alternative #2 dominates alternative #3 (as it has better values upon criterion three), and alternative #4 dominates alternatives #2 and #3. Alternatives #1 and #4 are uncomparable upon dominance relation.

On the other hand, we know that decision classes are ordered for the DM. This means that any alternative from the first class is more preferable for the DM than any alternative from the second class, and so on. This property may be reflected in the following binary preference relation on the set Y:

$$P^1 = \{(y^i, y^j) \in Y \times Y \mid y^i \in Y_k, y^j \in Y_1, k < 1\}.$$

It is natural to assume that no vector from Y dominating the one under consideration is to be in a less preferable class. For our example, this means that if the alternative #2 from Table 1 belongs to the first class according to the decision maker's opinion, then alternative #1 must also belong to the first class, as it dominates the alternative #2. Formally this may be put down as follows:

if $(y^i, y^j) \in P^o$ and $y^i \in Y_1$, then $y^j \notin Y_k$ if $k < l$.

Let us call the partition of the set Y into classes non-contradictory if this requirement is fulfilled.

The requirement for the partition to be non-contradictory is equal then to the fulfillment of the following:

$$\text{if } (y^i, y^j) \in P^o, \text{ then } (y^j, y^i) \notin P^1. \tag{1}$$

It is possible to accomplish the task of classification by having the decision maker classify directly all possible vectors from Y. However, this is impractical even for a relatively small problem, which can involve a large number of such vectors. Therefore, a special procedure for elicitation of decision maker classification rules has been developed (Moshkovich, 1985).

A RATIONAL PROCEDURE FOR ORGANIZATION OF THE ELICITATION PROCESS

The proposed procedure allows building of the needed classification with the help of a limited number of questions to a decision maker. The idea of the procedure is based on the implementation of the requirement (1). Suppose we have only two decision classes for our example: class 1 means that the manuscript is to be published, and class 2 means that the manuscript is to be rejected. If we ask the decision maker to classify alternative #2 we shall be able to classify alternatives #1 or #3 without presenting it to the decision maker. If the decision maker considers alternative #2 worthy of class 1, then alternative #1 is also to be assigned to class 1 (as it dominates alternative #2). In case the decision maker considers that alternative #2 must be rejected (class 2), we are able to say that alternative #3 must also be rejected (as it is dominated by alternative #2). Thus, any answer of the decision maker for the alternative #2 will determine the appropriate class for one of the other two alternatives from our example. Note, that if we first present the alternative #1 or alternative #3 for classification the consequences may be different. In case we present alternative #1 and the decision maker assigns it to class Y_1 (and this is very probable as it has good attribute values), we shall not be able to make any conclusions about appropriate class for alternative #2 or alternative #3.

Analogous results will be obtained it we present alternative #3 and the decision maker assigns class Y_2 to it.

Therefore, it is attractive to classify as many alternatives as possible by logical rules inferred from previous classifications given by the decision maker. Thus, the choice of a vector from Y for classification by the decision maker may influence the effectiveness of the interview (if effectiveness is evaluated by the number of vectors from Y the decision maker has directly to classify to complete the whole task). In this sense vectors from Y may be differently 'informative' for the construction of a complete classification of vectors from Y, and we can formulate a task of determining the most 'informative' vector at each step of the interview with the decision maker.

Let g_{il} denote the number of vectors definitely classified by assigning class Y_l to vector y^i. Thus, g_{il} characterizes the amount of information gained as a result of such a decision. This amount depends on the class prescribed to the vector y^i. Therefore, we are not able to evaluate it exactly (as we do not know in advance the class the decision maker will assign to the presented vector). This situation leads to the attempt to evaluate the possible amount of information connected with the vector y^i, and the necessity to introduce some index which will characterize the likelihood of class Y_l for vector y^i. We propose the following heuristic approach to this problem.

Let p_{il} denote the index which reflects the likeliness of y^i being assigned to class Y_l. Then the expected amount of information A_i connected with classification of vector y^i may be defined by the formula (2):

$$A_i = \sum_{l=1}^{N} p_{il} \times g_{il}. \tag{2}$$

There may be different heuristics for calculation of p_{il} in this formula. It is clear that the possibility of vector y^i to be assigned class Y_l is connected with some notion of the 'similarity' of y^i and elements of class Y_l. We introduce the formal idea of the center c^l of class Y_l, which is defined according to formula (3). This is an artificial point in the criteria space with averaged values upon all criteria. Though it has no special physical sense it reflects some averaged image of the class representative, and will be used later to evaluate the required 'similarity':

$$c^l = (c_1^l c_2^l, \ldots, c_Q^l), \text{ where } c_q^l = \left(\sum_{y^j \in Y_l} y_q^j \right) / |Y_l|, q = 1, 2, \quad, Q. \tag{3}$$

The index p_{il} then, is based on the measure of the 'distance' between the vector y^i and the center c^l. Smaller distance will reflect a larger possibility for vector y^i to be assigned class Y_l. Therefore, distance d_{il} between vector y^i and the center of lth class will be calculated upon the formula (4):

$$d_{il} = \sum_{q=1}^{Q} |y_q^i - c_q^l|. \tag{4}$$

Let d_{max} denote the maximum possible distance between two vectors from Y:

$$d_{max} = \sum_{q=1}^{Q} (\omega_q - 1).$$

Then p_{il} will be calculated according to the formula (5):

$$p_{il} = (d_{max} - d_{il}) / \left(|G^i| d_{max} - \sum_{j \in G}^{i} d_{ij} \right) \tag{5}$$

where G^i is the set of classes to one of which it is possible to assign vector y^i at the present level of obtained information. At the beginning $G^i = \{1, 2, \ldots, N\}$ for each $y^i \in Y$ as we do not have additional information. As our final goal is to assign each vector from Y one of N classes we need that at the end $|G^i| = 1$, for each $y^i \in Y$.

As we have ordinal criteria scales, formulas (3)–(5) may not be quantitatively meaningful, but as their outputs are used only in a substantive sense (to generate some rough estimation of our expectations), we consider them valid enough and useful in the proposed heuristic procedure.

The rational procedure of interviewing a DM is based on sequential presenting to a DM the most informative vectors, that is y^o for which $A_o = \max_i A_i$.

Let y^o be the most informative vector at the current step. We present it to a decision maker for classification, and receive the answer that $y^o \in Y_l$.

It is natural to assume that for $\forall y^j \in Y$ such that $(y^j, y^o) \in P^o, y^j$ may not belong to the class less preferable than Y_l. This will lead us to $G^j = \{1, 2, \ldots, l\}$. Analogously for $\forall y^k \in Y$ such that $(y^o, y^k) \in P^o, y^k$ may be assigned a class not more preferable than Y_l. This will lead us to $G^k = \{l, l + 1, \ldots, Q\}$. Thus, the information about one vector is able to decrease sets G^i for other vectors, and in some cases, to receive the only appropriate class for vectors not presented to a decison maker.

After the outlined processing of the decision maker's answer, it is necessary to recalculate the informative index for all non-classified vectors from Y, and repeat the procedure. It will end when all vectors from Y are classified ($|G^i| = 1$ for each $y^i \in Y$).

The proposed procedure uses a heuristic approach and therefore needs some evaluation. To evaluate the effectiveness, results of two procedures were used: a proposed one (pr1), and the so-called pattern one (pr2). The second procedure is based on the assumption that the complete classification is given. In this case we know in advance the amount of information gained if we choose vector y^i for classification. It is equal to g_{il} where l is the number of the class to which y^i is assigned. Thus, in this procedure, A_i is calculated not according to the formula (2) but as $A_i = g_{il}$. The result was measured by the number of presented vectors to construct the complete classification.

Classifications were built on the basis of random number generation (Moshkovich, 1985) for different values of the number of criteria (Q), the number of classes (N), and the number of values for each criterion scale (ω).

Data for different values of N, ω, Q (in about 1000 experiments for each case) are given in Table 2.

Table 2. A comparison of a number of vectors presented in two procedures

Q	ω	L	pr1			pr2		
			$N = 2$	$N = 3$	$N = 4$	$N = 2$	$N = 3$	$N = 4$
4	3	81	8	13	17	3	8	11
	4	256	10	14	21	4	9	13
5	3	243	10	18	25	4	9	15
	4	1024	14	24	33	5	12	19

The data show that the proposed procedure provides a rather effective system for interviewing a decision maker in an ordinal classification task.

PROCEDURE FOR LOCATION AND ELIMINATION OF ERRORS IN DM's ANSWERS

People can make judgmental errors for a variety of reasons. Thus, it is necessary to have tools to detect and correct these possible errors in the decision maker's responses.

In our case, the possibility for detection of errors is based on the requirement expressed in relation (1). Violations of this requirement indicate the presence of some inconsistencies in the gained responses, as this relation requires more preferable objects to be assigned to a more preferable class. For example, if alternative #1 (see Table 1) is assigned to the second class, and alternative #2 has been assigned to the first class, we can say that these two responses contradict each other, as alternative #1 has better criterion values. The decision maker must reconsider information and change one (or both) of the responses.

Larichev and Moshkovich (1987) proposed a special approach to location and correction of errors in a *built* classification [with possible violations of the requirement (1)]. An effective procedure allowed detection of responses inconsistent with the largest number of other ones, and, on this basis, to correct all the contradictions at one iteration. On the other hand, our experience shows that correction of errors in the *course* of classification is more convenient, as it requires less effort to find inconsistent responses and provides a 'learning' effect for a decision maker.

Thus, we suggest the following approach to errors' elimination. Let $Y^o \in Y$ be the subset of vectors

being classified up to the moment. Now, the decision maker is presented with vector $y^i \in Y$. G^i is the set of possible classes for y^i according to the information at hand, that is $G^i = \{n,n + 1,\ldots,n + q\}$, where $n \geq 1$ and $n + q \leq Q$. To get a better understanding of the process consider the data in Table 1. If we have three decision classes, and the decision maker has assigned the second class to alternative #2, we are able to conclude that alternative #1 may belong to the second or the first class (as it dominates alternative #1): $G^1 = \{1,2\}$. At the same time alternative #3 may belong to the second or third class, and, as so, $G^3 = \{2,3\}$ (as alternative #2 dominates alternative #3).

Let us suppose that the decision maker assigns class Y_s to y^i, and $s < n$ or $s > n + q$ (e.g. alternative #1 in our example is assigned the third class, or alternative #3 is assigned the first class). This response is inconsistent with the previous one(s), as there exists at least one vector in Y dominating y^i, and assigned to class Y_n (that is why possible classes for y^i in G^i start from n). Analogously, there exists a vector in Y, which is dominated by y^i and is assigned class Y_{n+q}.

Previous considerations may be used for construction of a subset of classified vectors from Y, violating relation (1) paired with $y^i \in Y_s$. Let us denote such a subset as Y_{err}, and define its elements in the following way:

$$\text{if } s < n: \quad Y_{err} = \{y^j \in Y^o \mid (y^j,y^i) \in P^o; \; y^j \in Y_z z > s\} \tag{6}$$

$$\text{if } s > n + q: \quad Y_{err} = \{y^j \in Y^o \mid (y^i,y^j) \in P^o; \; y^j \in Y_z z < s\}. \tag{7}$$

For our example Y_{err} will contain only one vector, that is alternative #2. In the general case this set may be large enough.

Then the decision maker is consequently presented with pairs of vectors (y^i,y^j), where $y^j \in Y_{err}$, and classes assigned to them. The decision maker is to analyze the contradiction and to change one or both of these responses to eliminate it.

After that, the correction of the appropriate classes in accordance with new assignments is carried out, and y^j is eliminated from Y_{err}. When $Y_{err} = \phi$, the procedure stops.

It is necessary to note that when we have more than two classes of decision, the newly given responses may be different. That is why the elimination of all elements from Y_{err} does not guarantee the absence of contradictions in the whole classification of the set Y^o (that is: new classes for some vectors may be inconsistent with classes of other vectors from Y^o previously assigned to them). Let us illustrate it by the example for data from Table 1.

Let us consider that we know classes of alternatives #2, #3 and #4. Alternatives #2 and #4 belong to the second class, and alternative #3 is assigned to the third class. This gives us $G^1 = \{1,2\}$ for alternative #1. The decision maker gives this alternative the third class. We have a contradiction, and $Y_{err} = \{\text{alternative } #2\}$. We present the decision maker with alternatives #1 and #2, and request the necessary change of the received responses. The decision maker analyzes the situation, and decides that alternative #1 and alternative #2 must both belong to the first class. In this case, we eliminate the previous contradiction ($Y_{err} = \phi$), but there appears to be a new one: alternative #4 dominates alternative #2 and belongs to the second class, while alternative #2 is now assigned to the first class.

Thus, only if the decision maker changes the class for y^i, are we able to continue our procedure. Otherwise, it is necessary to check requirement (1) for all $y^j \in Y_{err}$ with changed classes. If new contradictions appear, the procedure is to be repeated.

To make the task less complex we suggest in this case to build a matrix of contradictions A. Let $I = |Y^o|$. Then $A = \|a_{ij}\|$:

$$a_{ij} = \begin{cases} 1, & \text{if } (y^i,y^j) \in P^o, i,j = 1,2,\ldots,L, i \neq j; \\ 0, & \text{otherwise} \end{cases}$$

(such a matrix for our example is presented in Fig. 1).

Let us underline a_{ij} equal to 1, for which $y^i \in Y_s$, $y^j \in Y_t$ and $s < t$ (the matrix with underlined elements for our example is presented in Fig. 2). It is evident that these underlined elements reflect contradictions in the decision maker's responses. If there are no marked elements, we are able to continue the interview to build the classification. Otherwise, it is necessary to present the decision maker with a pair of vectors corresponding to the underlined element, correct the class, and correct

Vectors of Y^0		#1	#4	#2	#3
	Number of the class	3	2	2	3
#1	3		0	1	1
#4	2			1	1
#2	2				1
#3	3				

Fig. 1. Matrix A for data in Table 1.

Vectors of Y^0		#1	#4	#2	#3
	Number of the class	3	2	2	3
#1	3		0	<u>1</u>	1
#4	2			1	1
#2	2				1
#3	3				

Fig. 2. Matrix A with underlined elements.

Vectors of Y^0		#1	#4	#2	#3
	Number of the class	1	2	1	3
#1	1		0	1	1
#4	2			<u>1</u>	1
#2	1				1
#3	3				

Fig. 3. Matrix A after reclassification of alternatives #1 and #2.

the corresponding information in the matrix A. In Fig. 3 you can see the modified matrix A after assigning the first class to alternatives #1 and #2.

Thus, the correction process is fulfilled, only if there are no underlined elements in the matrix A. Otherwise, the underlined elements are being processed in the same manner.

The convergency of such a sequential correction procedure is guaranteed by the fact that at each iteration the decision maker is to decrease the class number for vectors from Y_{err} (in case $s < n$), and consequently to increase them (in case $s > n + q$). In these circumstances, finite, and not large numbers of possible classes limit the possibility of appearance of new contradictions, providing a high speed of convergency.

After the elimination of all contradictions in the decision maker's responses, it is necessary to simulate the procedure of processing these responses for vectors from Y^o, as if newly obtained from a decision maker. This will allow correction of sets G^k for all non-classified vectors from Y.

Thus, the suggested approach allows us to carry out an effective procedure of 'on-line' correction of possible errors in the decision maker's responses.

HUMAN BEING'S CAPACITIES IN ORDINAL CLASSIFICATION TASK

The proposed approach is based on the direct classification by a decision maker of complicated multi-criteria objects (alternatives). This may cause rather a large load on a human short-term memory. That is why a series of investigations of the human capacities in ordinal classification task was carried out (Larichev et al., 1988).

The complexity of the classification problem was assumed to depend on the problem variables, i.e. the number of attributes (Q), number of values on their scales (ω), and number of decision classes

(N). A hypothesis was that human behavior could vary following a certain change in some problem variable.

In the course of the experiments, subjects were requested to classify all possible objects (vectors from Y), using the prescribed decision classes. Four main measures of human performance were used:

(a) Number of inconsistencies (errors). It was viewed as a violation of the requirement (1) as in the previous section.

(b) Evaluation strategy. Special procedure (Larichev and Moshkovich, 1987) was used to modify each subject's responses in such a way that the resulting classification has no contradictions. The procedure searches for a minimal number of changes in classification of vectors from Y necessary to make the whole classification consistent. On the basis of the constructed consistent classification, sets of productive rules the subject might have used in assigning options to classes, were elaborated (details of elaborating these strategies may be found in Larichev and Moshkovich, 1988).

(c) Complexity. The number and type of rules were taken as an indication of the complexity of a subject's strategy. The easy way to obtain consistent classification is to use a set of simple conjunctive rules (Payne, 1976). Complex strategies must result in compensatory rules.

(d) Solubility. According to the type of errors and complexity of the used strategies the fulfillment of the task by individual subjects was characterized as successful or unsuccessful.

The experimental results confirmed the hypothesis that there are certain 'limits' to the subject capacities in multi-attribute ordinal classification problems (Larichev et al., 1988). They are summarized in Table 3.

Table 3. The marginal number of attributes under which the subjects still manage solution of new multi-attribute ordinal classification problems

Number of values on ordinal scales	Number of decision classes			
	2	3	4	5
2	7–8	6–7	4–5	3
3	5–6	3–4	2–3	2
4	2–3			

In other ranges of task parameters, the number of inconsistencies sharply increased. Subjects failed in managing the problem, and their responses did not allow elaboration of meaningful decision rules of classification. What is behind the phenomenon?

The analysis showed that when the subjects managed the task the number of rules they used did not exceed eight. In cases when they failed, a formal analysis revealed a much larger number of rules.

The most suitable explanation for the above data is probably as follows. In assigning an alternative to some class or other, the subject has to keep all the rules in the short-term memory, constituting structural units of information (chunks) he (or she) operates. As is known, the volume of short-term memory is limited. Different scientists (Miller, 1956; Simon, 1974; 1981) indicate that it does not exceed five–nine structural units of information (chunks), and they may differ in size.

When the subject employed nine or fewer rules for classification, they managed the task. If more, then a part of the rules was abundant for the operating in short-term memory which sharply increased the number of errors and inconsistencies.

DECISION SUPPORT SYSTEM

On the basis of the described approach a decision support system ORCLASS for ordinal classification tasks was developed.

The initial information necessary to start the work with the system consists of criteria with their scales, and lists with classes of decision (see the Appendix for an example). As it has been stated above, all criteria have ordinal scales and verbal descriptions of quality grades. Decision classes are also ranked from the best to the worst and also have verbal descriptions. All hypothetical combinations of criterion values are formed by a Cartesian product of scales. ORCLASS calculates

The following situation with the manuscript is under consideration:

1. The article has correspondence to the journal's outline.
2. The results have some theoretical value in the field.
3. The results have practical value.
4. Too many errors and inconsistencies in the article.
5. The article needs some refinement in style.

POSSIBLE ANSWERS:

1. Article may be directly published.
2. Article has to be revised and then published.
3. Article has to be revised and then reviewed.
4. Article has to be rejected.

YOUR ANSWER:

Fig. 4. Visualization of the situation and menu of possible answers.

1. The article directly corresponds to the journal's outline.
2. The results have theoretical value in the field.
3. The results are of high practical value.
4. Too many errors and inconsistencies in the article.
5. The article is awfully written.

THE SITUATION IS ESTIMATED AS:
    2. The article has to be revised and then published.

1. The article directly corresponds to the journal's outline.
2. The results have theoretical value in the field.
3. The results are of high practical value.
4. Too many errors and inconsistencies in the article.
5. The article needs some refinement in style.

THE SITUATION IS ESTIMATED AS:
    3. The article has to be revised and then reviewed.

The second situation is more preferable than the first one
according to the criteria values. It must be put to a not less
preferable class than the first situation. Analyze the
inconsistency and assign both situations once again.

PRESS ANY KEY TO CONTINUE

Fig. 5. Display of inconsistent responses with explanations.

the most 'informative' vector from Y and displays it. An example of such a presentation is given in Fig. 4.

If the response of the decision maker (class for the presented alternative), contradicts the previous ones, then ORCLASS informs the user about this fact and suggests to alter this response or to analyze the situation. If the user prefers to analyze the contradiction, then the system displays relevant information as in Fig. 5.

The interview is continued up to the moment when the classification is built.

After that, the system provides the user with the possibilities to analyze the classification rules, used by him (or her) in this task. This is done through presenting the user with boundary elements (the most and the least preferable vectors of each class according to dominance relation), and corresponding explanations. An example of such boundary elements is given in Fig. 6.

If the user disagrees with some elements, he (or she) is able to change them. The system will help to eliminate contradictions and form modified classifications.

If the user is satisfied, the system may be used for classification of real objects. In this case, the user is to enter into the system the appropriate alternatives with corresponding values upon the formed set

CLASS 1

The most preferable vectors:
11111

The least preferable vectors:
23212   23311

CLASS 2

The most preferable vectors:
11121   11113   11312

The least preferable vectors:
23323

CLASS 3

The most preferable vectors:
11131

The least preferable vectors:
13333   22333   23233   23332

CLASS 4

The most preferable vectors:
31111   23333

The least preferable vectors:
33333

Fig. 6. Presentation of the most and least preferable vectors in each decision class.

of criteria. The alternatives are presented in the system ORCLASS as a corresponding vector from Y. The system ORCLASS finds the assigned class for a vector from Y, which describes the considered alternative and presents it to the user.

CONCLUSION

We have proposed an approach to the solution of ordinal classification tasks, based on the possibility of a decision maker to classify separate alternatives. To our mind these tasks are wide spread in practice, but not very popular with specialists in decision-making. System ORCLASS allows one to elicit information (or knowledge) in a traditional form (through qualitative criteria scales and verbal descriptions of decision classes). It provides the possibility to receive reliable information as it tests this information for consistency. ORCLASS generates a complete classification rule, which makes it possible to find the decision class for any combination of criteria values, reducing the number of questions by choosing the most informative ones.

This system has been used in a wide range of practical tasks, from R&D planning context (Larichev and Moshkovich, 1987) to medical diagnostics (Larichev et al., 1986).

REFERENCES

Bana e Costa, C. A. (1992). Absolute and relative evaluation problematics. The concept of neutral level and the MCDA Robot technique. In Cerny, D. Gluckaufova and D. Loula (Eds) Proceedings of the International Workshop on Multicriteria Decision Making, Methods – Algorithms – Applications Liblice, 18–22 March, 1991, Prague.
Bana e Costa, C. A. (1993). Les problematiques dans le cardre de l'activite d'aide a la decision. In Document du LAMSADE, p. 80. Paris: Universite de Paris-Dauphine.

Larichev, O. I. & Moshkovich, H. M. (1986). Direct classification problems in decision making. *Doklady Akademii Nauk*, Vol. 287, pp. 567–570.

Larichev, O. I. & Moshkovich, H. M. (1987). Direct classification method and problems of eliciting reliable expert information. *Izvestia Akademii nauk SSSR. Technicheskaya Kibernetica*, Vol. 1, pp. 151–161.

Larichev, O. I. & Moshkovich, H. M. (1988). Limits to decision making ability in direct multiattribute alternative evaluation. *Organizational Behavior and Human Decision Processes*, Vol. 42, pp. 217–233.

Larichev, O. I., Moshkovich, H. M. & Furems, E. M. (1986). Decision support system 'CLASS'. In B. Brehmer, H. Jungerman, P. Lourens and G. Seton (Eds) *New Directions on Research in Decision Making*. Amsterdam: North Holland.

Larichev, O. I., Moshkovich, H. M. & Rebrik, S. B. (1988). Systematic research into human behavior in multiattribute object classification problems. *Acta Psychologica*, Vol. 68, pp. 171–182.

Miller, G. (1956). The magical number seven plus or minus two: some limits on our capacity for processing information. *Psychological Bulletin*, Vol. 63, pp. 81–97.

de Montgolfier, J. & Bertier, P. (1978). *Approche Multicritere des Problems de Decision*. Paris: Edicions Hommes et Techniques.

Moshkovich, H. M. (1985). *A Rational Procedure of Decision Maker Interview for Assigning Alternatives to Decision Classes*. Moscow: VNIISI Press.

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

Roy, B. (1985). *Methodologie Multicritere d'Aide a la Decision*. Economica.

Simon, H. A. (1974). How big is a chunk? *Science*, Vol. 183, pp. 482–488.

Simon, H. A. (1981). *The Science of Artificial*. (London: MIT).

Yu, W. (1992). ELECTRE TRI, aspects methodologiques et manuel d'utilisation. In *Document du LAMSADE*, 74. Paris: Universite de Paris-Dauphine.

APPENDIX

Criteria for manuscripts submitted to the journal:

Criterion 1. Correspondence to the journal's outline:

(1) The article is directly related to the journal's outline.
(2) The article has correspondence to the journal's outline.
(3) The article has rather low correspondence to the journal's outline.

Criterion 2. Theoretical value of the results:

(1) The results are of sound theoretical value in the field.
(2) The results have some theoretical value in the field.
(3) The results have no theoretical value in the field.

Criterion 3. Practical value of the results:

(1) The results are of high practical value.
(2) The results have practical value.
(3) The results have no practical value.

Criterion 4. Errors:

(1) There are no errors and inconsistencies in the article.
(2) There are some errors and inconsistencies in the article.
(3) Too many errors and inconsistencies in the article.

Criterion 5. Quality of the text:

(1) The article is written in a good language.
(2) The article needs some refinement in style.
(3) The article is awfully written.

Decision classes:

Class 1. Article may be directly published.
Class 2. Article has to be revised and then published.
Class 3. Article has to be revised and then reviewed.
Class 4. Article has to be rejected.