INFORMATICA, 2001, Vol. 12, No. 1, 109–118

Method for Approximation of Diverse Individual Sorting Rules

Alexey PETROVSKY
Institute for Systems Analysis, Russian Academy of Sciences
Prospect 60 Let Octyabrya 9, 117312 Moscow, Russia
e-mail: pab@isa.ru

Received: January 2001

Abstract. This paper considers the technique to construct the general decision rule for the contradictory expert classification of objects which are described with many qualitative attributes. This approach is based on the theory of multiset metric spaces, and allows to classify a collection of multi-attribute objects and define the classification rule which approximates the set of individual sorting rules.

Key words: multi-attribute objects, classification, sorting rules, multiset metric spaces.

1. Introduction

One of the wide-spread problem in the decision making area is the decomposition of objects' collection into several classes, that is based on the set of individual sorting rules. These individual rules, which are assigned each object into the specific class, may be similar or discordant. So often it is desirable to construct a general rule for the object classification that would take into account object characteristics and maximally coincide with the individual rules.

The difficulties of object classification increase when the objects are described with many qualitative attributes and can exist in several copies with various values of attributes. The problems of these kinds are, for example, the competitive selection of projects estimated by several experts with many qualitative criteria, the recognition of graphic symbols, the problem-oriented sorting of textual documents. In all these cases the object (project, symbol, document) can be presented as the set of repeating attributes.

This paper describes an approach to generate the classes of such objects and define the boundaries between classes. The technique for the multi-attribute objects' classification and the construction of general decision rule is based on searching for the best decomposition of multisets in the metric spaces.

2. Competitive Selection of Projects

One of the practical case, where we had faced the problem how to find the general decision rule for the contradictory expert classification of multi-attribute objects, is related to the preparation of the State scientific-technological Program on hightemperature-superconductivity in the former USSR (Larichev *et al.*, 1989). In order to select R&D projects for the Program the special competition was performed. The Scientific Board, which was responsible for the Program elaboration on the whole, organized 5 Competitive Commissions on the subprograms, which were to choose the projects correspondent to the Program goals. The Competitive Commissions worked together with the groups of experts who evaluated applications submitted for the Program.

Consider the expert procedure of project evaluation in more detail. Suppose that $A_1, \ldots, A_k$ are the projects presented at the competition. Each project is evaluated by n experts with m qualitative criteria $Q_1, Q_2, \ldots, Q_m$. In our case the questionnaire includes the following criteria for the project estimation: Q_1 – "The project contribution to the Program goals"; Q_2 – "A long-range value of the project"; Q_3 – "A novelty of the approach to solve the task"; Q_4 – "A qualification level of the team"; Q_5 – "Resources available for the project realization"; Q_6 – "A character of the project results".

Each criterion has a nominative or ordered scale of verbal estimates. For instance, the scale of the criterion Q_5. "Qualification level of the team" looks like this:

q_4^1 – the team is one of the best by the experience and qualification level;
q_4^2 – the team has the experience and qualification level sufficient for the project realization;
q_4^3 – the team has the experience and qualification level insufficient for the project realization;
q_4^4 – an experience and qualification level of the team are unknown.

An expert evaluates the project with all criteria and makes one of the following recommendations:
r_1 – to approve the project;
r_2 – to reject the project;
r_3 – to consider the project later after improving.

Note that one and the same project can be evaluated by different experts with identical or diverse criteria estimates. The expert recommendations on the project approval (sorting rules) also may coincide or not.

More than 250 applications were estimated by experts. Taking into account expert recommendations Competitive Commissions selected the projects for each subprogram. The decisions of Competitive Commissions were based on their own preferences, which were different and usually not expressed on the language of criteria estimates. Finally, the Scientific Board considered the Commissions' decisions, approved about 170 projects and included them in the Program.

In this case, we have an hierarchy of decision rules that is the individual sorting rules of many experts, the selection rules of several Competitive Commissions, and the final decision rule of the Scientific Board. So, members of the Competitive Commissions and Scientific Board were in need of simple general decision rules for the project selection both for the subprograms, and for the Program on the whole, which could approximate the variety of individual expert rules. Thus the decision makers would obtain an opportunity to use these approximating rules as the final decision rules or work out another decision rules.

There are various ways how to find decision rules for the classification of multi-attribute objects (Larichev and Moshkovich, 1997; Larichev *et al.*, 1989; Pawlak and Slowinski, 1994). But in our case it was necessary to provide a compromise between plural individual sorting rules and the final solution. In the considered situation, the final decision rule was found with a help of the specific DSS, which was constructed for this Program. The DSS consisted of the bases of project data, expert estimates, and special multiple criteria decision making methods. The satisfactory decision rule was worked out in the process of multi-stage interactive dialogue between the Scientific Board members and DSS. Various complex heuristic strategies were used to search for a concordance with the decisions of the Competitive Commissions and Scientific Board.

The following general decision rule was suggested for the project selection (Larichev *et al.*, 1989): "The project is to be very important or important for the achievement of the major program goals (the estimates q_1^1 or q_1^2 on the criterion Q_1), the team would have the experience, qualification level, and resources sufficient for the project realization (the estimates q_4^1 or q_4^2 on the criterion Q_4, and q_5^1 or q_5^2 on the criterion Q_5)". Almost all of the approved projects met these requirements. But the problem how a set of diverse individual sorting rules can be approximated with the simple general rule for the project classification stood over in that time.

3. Problem of Object Classification

The classification deals with combining the initial collection of objects into several groups or sorting them out of the predefined categories. Information about the object properties can be presented with the set of attributes whose values are numerical and/or verbal. From the formal logic point of view, a procedure of object classification can be written as a sequence of the following decision rules:

$$\text{IF } \langle\text{conditions}\rangle, \quad \text{THEN } \langle\text{decision}\rangle. \tag{1}$$

There are direct and indirect classifications. The direct classification is an enumeration of the objects within the class. So in this case, the term ⟨conditions⟩ includes the names of objects or the list of attribute values that describe the objects. The indirect classification is based on the properties common for the class. And the term ⟨conditions⟩ expresses the relations between different attributes and/or their values. The term ⟨decision⟩ marks that the object belongs to the specific class.

In the case of the project selection, the description of competitive project A_i $(i = 1, \ldots, k)$ consists of attributes that can be combined into several groups $G = \{Q_1, \ldots, Q_m, R\}$. Each group of criteria estimates $Q_s = \{q_s^e s\}$ $(s = 1, \ldots, m;$ $e_s = 1, \ldots, h_s)$ is the attribute family, which expresses the project property. The expert recommendations to approve, reject or correct the project (individual sorting rules) form the group of attributes $R = \{r_t\}$, which characterizes the project assignment into the class X_t ($t = 1, \dots, f$). Since each project is evaluated by several experts, some of the attributes may occur more than ones. So, the project can be described with the following set of repeating attributes:

$$A_i = \{(n_i(g_j) \bullet g_j)\} = \{(n_i(q_s^e s) \bullet q_s^e s), (n_i(r_t) \bullet r_t)\}. \tag{2}$$

Here g_j is an attribute from the set $G = \{g_j\}$ ($j = 1, \dots, h$, $h = h_1 + \dots + h_m + f$); $n_i(g_j)$ is a number of attribute g_j, which is equal to a number of experts who has estimated the project A_i with the attribute g_j; the sign $\bullet$ denotes that there are $n_i(g_j)$ copies of attribute g_j in the description of project A_i. The arguments in the formula (2) are associated with the decision rule (1) as follows: the various combinations of criteria estimates $q_s^e s$ correspond to the term $\langle$conditions$\rangle$; belonging the object A_i to the class X_t reflects the term $\langle$decision$\rangle$. The object A_i is said to be a member of class X_t if $n_i(r_t) > \sum_{p \neq t} n_i(r_p) = 0$.

When the objects are sorted by many experts, there is a family of decision rules which may be similar, diverse, and contradictory. Individual sorting rules are coincident or similar when the objects with the identical or resemble values of attributes are included in the same class. Contradictory rules assign the weakly discernible objects into diverse classes. The inconsistencies of individual rules may be caused, for instance, by errors in the expert classification of objects, the incoherence between expert' estimates of objects and decision classes, the intransitivity of expert judgements, and by other reasons. Note that knowledge bases of expert systems are built in the same manner.

If the number of objects and attributes is rather small, then decision rules are reviewed and utilized relatively easily. The more the family of decision rules, the more difficult analysis of these rules. In this case, the problem arises: how to generate the simple approximating rule(s) for the indirect classification of objects, which would maximally coincide with the set of contradictory sorting rules. The final decision rule would include a minimal number of attributes and assign the objects into the given classes with the admitted accuracy. A construction of the generalized decision rule allows us also to discover divergences in the initial direct classification and correct some of the expert sorting rules if necessary.

The classification of multi-attribute qualitative objects has some additional peculiarities. First, the amount, complexity and peculiarities of information necessary to specify qualitative objects are essentially larger and more varied then that for the quantitative objects. Second, a multiplicity and redundancy of factors, that express a substance of the problem considered, are possible. Third, the multi-attribute space and indexes of similarity/difference between objects are to be chosen corresponding to the qualitative nature of object properties. And finally, in order to classify qualitative objects, a lot of verbal and numerical data are to be taken into consideration simultaneously and processed without unfounded transformations (like "averaging", "mixing", "weighting" attributes, and so on). So the special procedures to collect and process these types of data are needed (Larichev and Moshkovich, 1997; Petrovsky, 1994; Petrovsky, 1995).

4. Multiset Model of Multi-Attribute Objects

The multiset or set with repeating elements (also called the bag) is a very convenient mathematical model in order to present and analyze a collection of objects that are described with many qualitative attributes and can exist in several copies with various values of attributes. Give the brief review of the multiset theory (Knuth, 1969; Yager, 1986; Petrovsky, 1994; Petrovsky, 1995).

Let $G = \{g_1, g_2, \dots, g_j, \dots\}$ be a crisp set, where all elements g_j are different. A is called a multiset over the domain G if A can be presented by the set of pairs as follows

$$A = \{(n_A(g) \bullet g)\},$$

where $n_A(g)$ is called a counting function of multiset A. This function defines the number of occurrences of the element g in the multiset A, and $n_A\colon\ G \to \mathbf{N}^+$. Unlike the set, each element may occur in the multiset more than once. The element g is said to be a member of the multiset A ($g \in A$), and there are k copies of g in A, if $n_A(g) = k > 0$. If $n_A(g) = 0$, then $g \notin A$. When $n_A(g)$ is equal to $\chi_A(g) = \{0, 1\}$, the multiset A becomes an ordinary set. The set $G = \{g_j\}$ is said to be a generic domain for the collection of multisets X, if all multisets from X are composed from the elements of G. The multiset is called: the empty multiset $\varnothing$, if $n_\varnothing(g) = 0$ for $\forall g \in G$; the maximal multiset Z, if $n_Z(g) = \max n_A(g)$ for all multisets $A \subseteq Z$. The cardinality of multiset A is a total number of all elements $|A| = \sum_{g \in G} n_A(g)$; and the dimensionality of multiset A is a number of different elements $[A] = \sum_{g \in G} \chi_A(g)$.

The following operations under multisets can be defined:

a union of multisets	$A \cup B = \{g \mid n_{A\cup B}(g) = \max(n_A(g), n_B(g))\}$;
an intersection of multisets	$A \cap B = \{g \mid n_{A\cap B}(g) = \min(n_A(g), n_B(g))\}$;
an addition of multisets	$A + B = \{g \mid n_{A+B}(g) = n_A(g) + n_B(g)\}$;
a difference of multisets	$A - B = \{g \mid n_{A-B}(g) = n_A(g) - n_{A\cap B}(g)\}$;
a symmetrical difference of multisets	$A \Delta B = \{g \mid n_{A\Delta B}(g) = \lvert n_A(g) - n_B(g)\rvert\}$;
a multiplication on scalar k	$k \cdot A = \{g \mid n_{k\cdot A}(g) = k n_A(g),\ \ k > 0\}$;
a multiplication of multisets	$A \cdot B = \{g \mid n_{A\cdot B}(g) = n_A(g) \cdot n_B(g)\}$;
the complement of multiset A	$\overline{A} = \{g \mid n_{\overline{A}}(g) = n_Z(g) - n_A(g)\}$.

A family of multisets closed under operations of the union, intersection and complement is said to be an algebra $L(Z)$ of multisets where the maximal element Z is the unit of the algebra. A real-valued function $m(A)$ defined on the algebra $L(Z)$ is called a measure of multiset A if $m(A) \geqslant 0$, $m(\varnothing) = 0$, $m(A) + m(B) = m(A \cup B) + m(A \cap B) = m(A + B)$. The measure of multiset A can be determined in the various ways, for instance, as a linear combination of counting functions: $m(A) = \sum_j w_j n_A(g_j)$, $w_j > 0$.

The metric spaces of multisets were introduced in (Petrovsky, 1994; Petrovsky, 1995). Different metric spaces (X, d) can be determined for the same collection of objects by introducing the various types of distances $d(A, B)$. The following metrics can exist in multiset spaces:

$$d_0(A, B) = m(A\Delta B);$$
$$d_1(A, B) = m(A\Delta B)/m(Z);$$
$$d_2(A, B) = m(A\Delta B)/m(A \cup B).$$

Functions $d_1(A, B)$ and $d_2(A, B)$ satisfy the normalization condition $0 \leqslant d(A, B) \leqslant 1$. Note, that due to the continuity of the multiset measure, the distance $d_2(A, B)$ is undefined for $A = B = \varnothing$. So $d_2(\varnothing, \varnothing) = 0$ by the definition. Nonmetric measures of multiset similarities are connected with the distances:

$$s_1(A, B) = 1 - m(A\Delta B)/m(Z);$$
$$s_2(A, B) = m(A \cap B)/m(A \cup B);$$
$$s_3(A, B) = m(A \cap B)/m(Z).$$

Let $X = \{A_1, \ldots, A_k\}$ be a collection of k objects, that are described with many qualitative attributes $G = \{g_1, \ldots, g_h\}$, and presented in the form (2). In other words, the expression (2) is the multiset $A_i \in X$ drawn from the set $G = \{g_j\}$. Thus the problem of multi-attribute object classification can be considered now as the problem of similarity/difference between multisets, that present objects in the metric space of multisets.

In order to simplify the problem, assume that the collection of objects $X = \{A_1, \ldots, A_k\}$ is to be sorted into only two classes X_a and X_b. In this case the objects' collection X can be presented as the following decompositions of multisets:

$$X = \sum_{t=a,b} X_t = \sum_{t=a,b} \Big(\sum_{s=1}^{m} Q_{st} + R_t\Big),$$
$$Q_{st} = \sum_{e_s=1}^{h_s} Q_{st}^{e_s}, \quad Q_{st}^{e_s} = \sum_{i \in I_{st}^{e_s}} A_i, \quad R_t = \sum_{i \in I_{rt}} A_i, \tag{3}$$

where $I_{st}^{e}s = I_{s}^{e}s \cap I_t$; I_t is the subset of indexes i for $A_i \in X$ with $n_i(r_t) > \sum_{p \neq t} n_i(r_p) = 0$; $I_s^e s$ is the subset of indexes i for $A_i \in X$ with $n_i(q_s^e s) \neq 0$, $n_i(q_v^e v) = 0$, $v \neq s$, $n_i(r_t) = 0$; $I_{rt} = I_r \cap I_t$; I_r is the subset of indexes i for $A_i \in X$ with $n_i(r_t) \neq 0$, $n_i(q_s^e s) = 0$.

The relations between the collection of objects $X = \{A_i\}$ and the set of their attributes $G = \{g_j\}$ can be expressed with the matrix $C = \|n_i(g_j)\|$. The matrix C is used often in data analysis, pattern recognition and called the "object-attribute" table, information table or decision table (Pawlak and Slowinski, 1994; Petrovsky, 1994; Petrovsky, 1995). In our case, information on properties of the multi-attribute objects A_i and information that the object A_i belongs to a certain decision class can be presented as the decision table C, that has a dimension $k \times h$, and consists of $2(m + 1)$ boxes which correspond to multisets Q_{sa}, Q_{sb} and R_a, R_b (k is a number of objects, h is a number of object attributes, $m+1$ is a number of attributes groups). The reduced decision table $C' = \|n'_i(g_j)\|$ has a dimension $2 \times h$, and consists of two rows $n'_A(g_j)$ and $n'_b(g_j)$ which correspond to the classes X_a and X_b.

The demand to sort objects into two classes is not the principle restriction for us. Whenever objects are to be classified into more than two classes, it is possible to divide the collection X into two groups, then into subgroups, and so on. For instance, the competitive projects can be classified into the projects approved and not approved, then the not approved projects can be divided into the projects rejected and considered later, and so on.

5. Approximation of Individual Sorting Rules

The main idea of approximating a large family of sorting rules with a compact decision algorithm or simple decision rule can be formulate as follows. In the metric space of multisets (X, d), the pairs of new multisets for every group of attributes $Q_1, \ldots, Q_m$, R would be generated. The multisets within each pair are to be spaced at the maximal distance d, and be the mostly coincident with the initial expert sorting of the objects into the classes X_a and X_b. Combinations of the attributes, that define the pairs of the generated multisets, produce the generalized decision rule for the object classification.

Obviously, the decomposition $R = \{R_a, R_b\}$ is the best partition of the object collection $X = \{A_i\}$ ($i = 1, \ldots, k$) into the classes X_a and X_b. The distance between the multisets R_a, R_b in the metric space (X, d) is maximal and equal to

$$d(R_a, R_b) = \max d(R_a, R_b) = d^*. \tag{4}$$

In the case of the ideal classification without inconsistencies of the individual sorting rules, the maximal distance in the metric space (X, d) is equal correspondingly to $d_0^* = kn$, $d_1^* = 1/h$, $d_2^* = 1$ (k is a number of the classified objects, n is a number of the individual sorting rules per an object, which is equal to a number of experts who evaluated the object, h is a total number of the object attributes).

The problem of how to approximate rules for sorting a collection of multi-attribute objects is transformed into the problem of how to find the best binary decompositions $Q_s = \{Q_{sa}, Q_{sb}\}$, where the multisets Q_{sa}, Q_{sb} ($s = 1, \ldots, m$) are maximally far from each other in the metric space (X, d). In other words, the following m optimization problems should be solved:

$$d(Q_{sa}, Q_{sb}) \to \max d(Q_{sa}, Q_{sb}) = d(Q_{sa}^*, Q_{sb}^*). \tag{5}$$

The solution of each problem (5) is the best binary decomposition of the multiset $Q_s^* = \{Q_{sa}^*, Q_{sb}^*\}$, where the multisets Q_{st}^* is a sum of submultisets $Q_{st}^{*1} + Q_{st}^{*2}$ ($t = a, b$). The attribute q_s^*, which belongs to the submultiset Q_{st}^{*1}, is called the approximating attribute.

Combinations of the approximating attributes q_s^* for various numbers s of criteria groups define the conditions for assigning the object $A_i \in X$ into a certain class X_t.

The approximating attributes q_s^* for different criteria groups s can be ordered according to the values of distances $d(Q_{sa}^*, Q_{sb}^*)$. Then attributes q_s^*, which occupy the first places in this ranking, are to be included in the generalized decision rule. The nearer the distances $d(Q_{sa}^*, Q_{sb}^*)$ to the maximal distance d^*, the more accurate the approximation of individual sorting rules. The rate of the sorting rules approximation for the attribute q_s^* can be estimated by the expression

$$\rho_s = d\left(Q_{sa}^*, Q_{sb}^*\right)/d^*, \tag{6}$$

where the distance d^* is determined by the formula (4). The approximating attribute q_s^* with the approximation rate $\rho_s \geqslant \rho_0$ (ρ_0 indicates the demanded rate of approximation) is to be included in the generalized decision rule (1) for the objects classification. Note that the value of approximation rate ρ_s characterizes the comparative importance of the criterion Q_s with respect to the final classification of multi-attribute objects.

The procedure for a generation of the generalized decision rule can be summarized as follows.

1. Compute the decision table $C = \|n_i(g_j)\|$ of dimension $k \times h$, that presents the collection of multi-attribute objects $X = \{A_i\}$ and consists of $2(m+1)$ boxes which correspond to multisets Q_{sa}, Q_{sb} and R_a, R_b.
2. Combine the objects A_i, which is related to the given classes X_a, X_b, by using the formula (3). Obtain the reduced decision table $C' = \|n'_i(g_j)\|$ of dimension $2 \times h$, that corresponds to the specified classes X_a and X_b, and consists of two rows $n'_a(g_j)$, $n'_b(g_j)$.
3. Solve the optimization problem (5) for the every binary decomposition Q_s and find the approximating attributes q_s^* in the every s-th box of the reduced matrix C'.
4. Range the approximating attributes q_s^* according to the values of distances $d(Q_{sa}^*, Q_{sb}^*)$.
5. Select the attribute q_s^* that provides the demanded approximation rate ρ_s (6). The set of these attributes $\{q_s^*\}$ forms the generalized decision rule for sorting the objects.

6. Case Study

Let us illustrate the proposed method for the approximation of contradictory sorting rules on the base of expert decisions, which are related to the Program considered above (Larichev *et al.*, 1989). The followings are a part of decision table C, the reduced decision table C', the distances between the multisets $d(R_a, R_b)$, $d(Q_{sa}^*, Q_{sb}^*)$ in the metric space (X, d_0), and the rates ρ_s of the sorting rules approximation for the attribute q_s^*, that are calculated in accordance with the given algorithm:

Objects	Attributes																					
	q_1^1	q_1^2	q_1^3	q_2^1	q_2^2	q_2^3	q_3^1	q_3^2	q_3^3	q_4^1	q_4^2	q_4^3	q_4^4	q_5^1	q_5^2	q_5^3	q_5^4	q_6^1	q_6^2	q_6^3	r_a	r_b
A_1	1	2	0	2	1	0	3	0	0	2	1	0	0	0	2	1	0	2	1	0	3	0
...																						
A_i	1	1	1	0	2	1	1	2	0	0	2	1	0	0	1	2	0	0	0	3	2	1
A_{i+1}	1	1	1	0	2	1	1	2	0	0	2	1	0	0	1	2	0	0	0	3	1	2
...																						
A_k	0	2	1	0	1	2	0	3	0	0	1	1	1	0	0	2	1	0	3	0	0	3

Classes of objects	Attributes																					
	q_1^1	q_1^2	q_1^3	q_2^1	q_2^2	q_2^3	q_3^1	q_3^2	q_3^3	q_4^1	q_4^2	q_4^3	q_4^4	q_5^1	q_5^2	q_5^3	q_5^4	q_6^1	q_6^2	q_6^3	r_a	r_b
X_a	144	360	21	81	324	120	99	336	90	219	297	9	0	72	435	18	0	126	300	99	510	15
X_b	45	156	51	27	93	132	36	111	105	51	132	63	6	60	147	30	15	45	135	72	78	174
d	333			297			303			393				327				273			591	
ρ_s	0.563			0.503			0.517			0.665				0.553				0.462				

The approved projects $A_1 - A_i$ belongs to the class X_a, and the not approved projects $A_{i+1} - A_k$ belongs to the class X_b. Observe that the projects A_i and A_{i+1} have the same collection of criteria estimates $\{q_s\}$ but their individual sorting rules do not coincide, that is $A_i \in X_a$, $A_{i+1} \in X_b$. The set of the approximating attributes q_s^*, which is ordered by the distances $d(Q_{sa}^*, Q_{sb}^*)$, consists of the following attributes:

$$\{q_s^*\} = \{q_4^1, q_4^2; q_1^1, q_1^2; q_5^1, q_5^2; q_3^1, q_3^2; q_2^1, q_2^2\}\,.$$

Note that there is no optimal solution of the problem (5) for the criterion Q_6. Thus, the all attributes q_6^p are not approximating. By choosing the demanded rate of approximation ρ_0, one can find the following general decision rules to select the competitive projects:

"The team must be one of the best or have the experience and qualification level sufficient for the project realization" (the estimates q_4^1 or q_4^2; the approximation rate $\rho_s \geqslant 0.65$);

"The project is to be very important or important for the achievement of the major program goals, the team must be one of the best or have the experience, qualification level, and resources sufficient for the project realization" (the estimates q_1^1 or q_1^2, and q_4^1 or q_4^2, and q_5^1 or q_5^2; the approximation rate $\rho_s \geqslant 0.55$).

The last rule is the same as mentioned above. One can also find the various general decision rules for the separate subprograms and discover the inconsistencies between the individual sorting rules and the general decision rules.

7. Conclusion

The quality of decisions depends on the quality and completeness of the problem investigation. However, there is a kind of problems where plurality and redundancy of data characterizing objects, alternatives, situations, and their properties are essential. In this paper we have suggested the tools for classifying a collection of objects represented by many qualitative attributes, when a lot of copies of objects or values of attributes describing them can exist. These techniques can be applied to analyze decisions in the wide-range situations, where the relations between objects and their attributes are presented as the "object-attribute" matrix, information or decision table.

The proposed tools are based on the theory of multiset metric spaces. The multiset approach allows us to discover, present and utilize the available information, which is contained in the object descriptions and could not be processed formerly by other methods. This approach provides the technique to classify multi-attribute objects and interpret the peculiarities of classification, especially in the cases of a large number of the objects considered and a variety of inconsistencies between the object properties and sorting rules.

This work is supported by the Russian Foundation for Basic Research (grant N 99-01-00476).

References

Knuth, D.E. (1969). *The Art of Computer Programming. Vol.2. Seminumerical Algorithms.* Addison-Wesley, Reading, MA.

Larichev, O.I., E.M. Moshkovich (1997). *Verbal Decision Analysis for Unstructured Problems.* Kluwer Academic Publishers, Boston.

Larichev, O.I., A.S. Prokhorov, A.B. Petrovsky, M.Yu. Sternin, G.I. Shepelev (1989). The experience of planning of the basic research on the competitive base. *Vestnik of the USSR Academy of Sciences*, **7**, 51–61 (in Russian).

Pawlak, Z., R. Slowinski (1994). Rougf set approach to multi-attribute decision analysis. *European Journal of Operational Research*, **72**, 443–459.

Petrovsky, A.B. (1994). An axiomatic approach to metrization of multiset space. In G.H. Tzeng, H.F. Wang, U.P. Wen, P.L. Yu (Eds), *Multiple Criteria Decision Making.* Springer-Verlag, New York. pp. 129–140.

Petrovsky, A.B. (1995). Metric spaces of multisets. *Doklady of the Academy of Sciences*, **344**(2), 175–177 (in Russian).

Yager, R.R. (1986). On the theory of bags. *International Journal of General Systems*, **13**, 23–37.

A. Petrovsky has graduated from Moscow State University in 1967, and received Ph.D degree (Theoretical and Mathematical Physics, 1970) from V.A. Steklov Mathematical Institute of the USSR Academy of Sciences. He is a Doctor of Technical Sciences (Systems Analysis and Automated Control, 1994) and a Head of the Laboratory for Decision Support Systems at the Institute for Systems Analysis of Russian Academy of Sciences. His research interests include theory of multisets, decision analysis, information technologies, R&D forecasting, planning and management.

Skirtingų individualių rūšiavimo taisyklių aproksimavimo metodas

Alexey PETROVSKY

Straipsnyje nagrinėjamas metodas bendrai sprendimo taisyklei sudaryti, kai ekspertai skirtingai klasifikuoja kokybiškai aprašytus objektus. Metodas remiasi aibių metrinėse erdvėse teorija ir leidžia klasifikuoti daugiaatributinius objektus, apibrėžiant taisyklę, aproksimuojančią individualių rūšiavimo taisyklių aibę.