Incomplete classifications of a finite set of objects using Monotone Systems*

Yu. M. Zaks and I. B. Muchnik UDC 519.237.8

The problem of incomplete classification is solved using a monotone system of a special kind. A classification method based on identification of the minimal cores of the monotone system is proposed. Existence conditions of a complete classification are given. The method is compared with previously published methods.

1. Introduction

Automatic classification methods, in addition to looking for complete classifications (i.e., partitioning of the initial set of objects), when classes are nonintersecting and form a covering of the entire set, also consider classifications with intersecting classes, with fuzzy classes, with macrostructure, etc. [1-3]. An independent group comprises incomplete classification methods, which identify a special class of "atypical" (background, special, or intermediate) objects and then assign the rest of the objects to nonintersecting classes [4-6]. Incomplete classifications are constructed when it is desirable to form classes comprising "strongly separated" subsets of objects, with all the background objects collected in a single class. This is archived either by two-step procedures, in which the first stage involving identification of the special objects is independent of the second stage involving classification proper, or by single-stage processing in which the classification functional is defined on the set of two-level classifications, thus complicating the discrete optimization problem. Moreover, both cases require specifying in advance the number of classes and the cardinality of the set of special objects, which leads to multi-alternative computations. Finally, most of the known algorithms are crudely approximate.

In this paper, the sought incomplete classification is implicitly described by a separate estimate for each identified class of nonspecial objects, and this estimate should be extremal and equal on all classes of the sought classification. The proposed approach requires minimum prior information: we only need to know the measure of association of one object with a subset of objects. The number of classes and the number of objects identified, as special, is not fixed in advance. The proposed algorithm guarantees exact solution of the corresponding extremal problem.

---

* Moscow. Translated from Avtomatica i Telemekhanika, No. 4, pp. 155 – 164, April, 1989. Original article submitted October 16, 1987.

2. Statement of the problem

Consider a finite set W of objects ($|W| = N$) and a function $\pi(i,H)$, which evaluates the association of the element $i \in W$ with the subset $H \subseteq W$ and satisfies the monotonicity condition

$$\pi(i,H) \le \pi(i,H \setminus k) \;\; \forall i \in W,\; \forall k \in H \subseteq W,\; i \ne k. \tag{1}$$

Introduce two set functions

$$F(H) = \min_{i \in W \setminus H} \pi(i,H) \tag{2}$$

and

$$F'(H) = \max_{i \in H} \pi(i,H). \tag{3}$$

It is easy to show [7] that these functions satisfy the following conditions:

a) quasiconcavity: $\forall H_1, H_2 \subset W$, such that $H_1 \cap H_2 \ne \varnothing$, we have

$$F(H_1 \cap H_2) \ge \min\left(F(H_1), F(H_2)\right); \tag{4}$$

b) quasiconvexity: $\forall H_1, H_2 \subset W$, $H_1 \ne \varnothing$, $H_2 \ne \varnothing$, we have

$$F'(H_1 \cup H_2) \le \max\left(F'(H_1), F'(H_2)\right). \tag{5}$$

Using these conditions, we can easily show that the family of proper subsets H of the set W (augmented with the empty set, if necessary) satisfying the condition $F(H) \ge u$ [1], is closed under the intersection (forms a lower semilattice of sets [8]), and the family of subsets satisfying the condition $F'(H) \ge u$, is closed under union (forms upper semilattice). Similarly for the function $\pi(i,H)$ such that

$$\pi(i,H) \ge \pi(i,H \setminus k) \;\; \forall i \in W,\; \forall k \in H \subseteq W,\; i \ne k, \tag{6}$$

we obtain that the functions

$$F(H) = \max_{i \in W \setminus H} \pi(i,H) \tag{7}$$

and

$$F'(H) = \min_{i \in H} \pi(i,H) \tag{8}$$

---

[1] Notification changed from the original $F(H) \le C$ to $F(H) \le u$. Note by J.M.

generate, respectively, a lower for $F(H) \le u$ and an upper $F'(H) \ge u$ semilattice, because in this case we obviously have $F(H_1 \cap H_2) \le max\left(F(H_1), F(H_2)\right)$, $\forall H_1, H_2 \subset W$ and $H_1 \cap H_2 \ne \varnothing$, and $F'(H_1 \cup H_2) \ge min\left(F'(H_1), F'(H_2)\right)$, $\forall H_1, H_2 \subset W$ and $H_1 \ne \varnothing$, $H_2 \ne \varnothing$.

Given the nature of the semilattices generated by $F(H)$ and $F'(H)$, we call the triple $\langle W, \pi, F \rangle$ an I-system and the triple $\langle W, \pi, F' \rangle$ a S-system, regardless of whether the function $\pi(i, H)$ satisfies condition (1) or the reverse condition.

S-systems [2] were proposed by Mullat in [9] and have since been used in applications, which require identifying the subset of objects that have the strongest association in some sense [10]. S-systems were applied in [11] to construct a complete classification by multiply solving the problem.

Unlike previous studies, our paper uses I-systems. [3] Similarly to the description of S-system [9], we introduce the following definition.

Definition. The cores [4] of the I-system $\langle W, \pi, F \rangle$ are the proper subsets of W on which $F(H)$ attains a maximum: H^* is a core if

$$H^* = arg \max_{\substack{H \subset W \\ H \ne \varnothing}} F(H). \tag{9}$$

We have noted above that the family of proper subsets of W satisfying the conditions $F(H) \ge u$, augmented with the empty set is necessary, is a lower semilattice. We see from the definition that a core is a particular case of such a family when $u = F(H^*)$. This means that the family of all cores is closed under intersection if the intersection of all the elements of the family is nonempty. If it is empty, then we augment it with the empty set, setting $F(\varnothing) = F(H^*)$, where H^* is a core.

2 Such systems originally were called "monotonic" (not "monotone"). To call such systems "monotonic" is more suitable, because as it turns out later the term "Monotone System" is common notification for at least four different subjects under investigation in the scientific community. Note by J.M:

3 We consider I-systems that satisfy condition (1).

4 These cores in [9] vocabulary are kernels contrary to widely used core solution in N -persons coalitions games. Note by J.M.

Thus, the family T_0 of cores of $\langle W,\pi,F\rangle$ is a lower semilattice. Therefore, the set of minimal (by inclusion) nonempty cores of $\langle W,\pi,F\rangle$ generates a classification of the elements of W into nonintersecting classes (in general, this classification is incomplete). The substantive interpretation of the function $\pi(i,H)$ (the distance of the element $i \in W \setminus H$ from the set H) and $F(H)$ (the measure of isolation of H) suggests that the classes of this classification are in fact the sought "strongly separated" subsets of objects, while the elements that are not assigned to any of these subsets are the "atypical" objects that occupy an intermediate position between the separated classes.

Our discussion suggests the following problem: design an algorithm to find the set of all minimal (by inclusion) cores of the system $\langle W,\pi,F\rangle$. The solution of the problem is the subject of our paper.

3. Construction of the determining [5] order of a monotone system

Let P be the set of all proper subsets of W.

Definition. Quasicores of 1-st level are the elements H^* of the set $\mathrm{P} \setminus \mathrm{T}_0$ such that $\tilde{H} = arg \max_{H \in \mathrm{P} \setminus \mathrm{T}_0} F(H)$ (we denote this set by T_1); quasicores of k-th levels are obtained from the condition $\tilde{H} = arg \max_{H \in \mathrm{P} \setminus \bigcup_{i=0}^{k-1} \mathrm{T}_i} F(H)$, where T_i is the set of quasicores of i-th level, $0 \le i \le k-1$, and quasicores of 0-th zero level are cores.

Let m be the highest level of the quasicores of the system $\langle W,\pi,F\rangle$. Then $\mathrm{P} = \bigcup_{i=0}^{m} \mathrm{T}_i$. Let $\mathrm{L}_s \underset{def}{=} \bigcup_{j=0}^{s} \mathrm{T}_j$ $(s = \overline{0,m})$, and let u_s be the value of F on the quasicores of the s-th level. Clearly, $u_0 \ge u_1 \ge \ldots \ge u_m$.

The set L_0, as we have noted above, may be regarded as a lower semilattice. In order to ensure that this property holds for any L_s, we must define $F(H)$ on the empty set, since starting with some $\hat{k}$ all the set L_s $(s \ge \hat{k})$ contain pairwise nonintersecting elements. We have the following theorem.

[5] Originally called "defining sequence", might be a translation circumstance. Note by J.M.

Theorem 1. If we set $F(\varnothing) = u$, then for any s $(0 \le s \le m)$, L_s is a lower semilattice.

The proofs of all the theorems are collected in the Appendix.

In what follows we always assume that $F(H)$ is defined on the empty set according to Theorem 1.

Let K_s be the zero of the semilattice L_s (its least element: $H \in L_0 \Rightarrow H \supseteq K_s$). Note that $F(K_s) \ge u_s$.

Theorem 2. The system of zeros of the semilattices L_s $(s = 0,...,m)$, is a chain, i.e., $K_0 \supseteq K_1 \supseteq ... \supseteq K_m$.

Eliminating repeating elements from this chain, we obtain a maximum-length chain of strictly included elements $K_{j_0} \supset K_{j_1} \supset ... \supset K_{j_t}$. Clearly $K_{j_0} = K_0$ and $K_{j_t} = \varnothing$. The maximum chain has the following properties:

$$F(K_{j_0}) > F(K_{j_1}) > ... > F(K_{j_t}); \tag{10}$$

$$\forall H : K_{j_{d-1}} \supset H \supseteq K_{j_d} \Rightarrow F(H) \le F(K_{j_d}),\ d = \overline{1,t}; \tag{11}$$

$$\forall H : H \supset K_{j_0} \Rightarrow F(H) \le F(K_{j_0}). \tag{12}$$

Consider a fixed element $i \in W$. Let $P^i = \{H \in P : i \in H\}$. Restrict the domain of definition of $F(H)$ to P^i. Let $\langle W, \pi, F, i \rangle$ denote the resulting monotone system. If $T_0^i,...,T_r^i$ is the system of its quasicores, then we can easily show that $L_s^i \underset{def}{=} \bigcup_{j=0}^{s} T_j^i$ $(s = \overline{0,r})$ are lower semilattices. Identify their zeros $K_0^i \supset ... \supset K_r^i$ with the corresponding values of $u_0^i,...,u_r^i$ and construct the maximum chain $K_{j_0}^i \supset ... \supset K_{j_h}^i$. If $i \in K_s$ $(s = \overline{0,m})$, then obviously $T_d^i = T_d$, $u_d^i = u_d$, $d = \overline{0,s}$.

Let us now proceed to construct a determining order. Let J^i be some order on W such that i is the first element in this order: $J^i = \langle j_1 = i, j_2,..., j_N \rangle$. It induces a sequence of sets from $P^i : \overline{H}^i = \langle H_1^i,...,H_N^i \rangle$, where $H_d^i = \{j_1, j_2,..., j_d\}$, $d = \overline{1,N}$.

Definition. The order J^i is called a determining order if the sequence of sets $\overline{H}^i = \left\langle H_1^i,...,H_N^i \right\rangle$ contains a subsequence $\overline{\Gamma}^i = \left\langle \Gamma_1^i,...,\Gamma_{\mathrm{p}}^i \right\rangle$, $\Gamma_1^i = \{i\}$, $\Gamma_1^i \subset ... \subset \Gamma_{\mathrm{p}}^i$, such that

$$F(\Gamma_1^i) < ... < F(\Gamma_{\mathrm{p}}^i); \tag{13}$$

$$\pi(j_{\mathrm{k}+1}, H_{\mathrm{k}}^i) \le F(\Gamma_{\mathrm{d}}^i) \ \ \forall j_{\mathrm{k}+1} \in \Gamma_{\mathrm{d}+1}^i \setminus \Gamma_{\mathrm{d}}^{\mathrm{i}},\ \mathrm{d} = \overline{1,\mathrm{p}-1}; \tag{14}$$

$$\pi(j_{\mathrm{k}+1}, H_{\mathrm{k}}^i) \le F(\Gamma_{\mathrm{p}}^i) \ \ \forall j_{\mathrm{k}+1} \in W \setminus \Gamma_{\mathrm{p}}^{\mathrm{i}}. \tag{15}$$

Note that these inequalities are dual analogs of the inequalities for the S-system [12].

Algorithm 1 to Construct the Determining Order J^i.

Step 1. $j_1 = i$, $H_1^i = \{j_1\}$, $j_2 = \underset{j \in W \setminus H_1^i}{arg\ min}\, \pi(j, H_1^i)$. Set the threshold $u_1 = \pi(j_2, H_1^i)$.

Step k. The sequence $\langle j_1,..., j_{\mathrm{k}} \rangle$ and the set $H_{\mathrm{k}}^i = \{j_1,..., j_{\mathrm{k}}\}$ have already been constructed. Find $j_{\mathrm{k}+1} = \underset{j \in W \setminus H_{\mathrm{k}}^i}{arg\ min}\, \pi(j, H_{\mathrm{k}}^i)$. If $\pi(j_{\mathrm{k}+1}, H_{\mathrm{k}}^i) \le u_{\mathrm{k}-1}$, then $u_{\mathrm{k}} = u_{\mathrm{k}-1}$, else $u_{\mathrm{k}} = \pi(j_{\mathrm{k}+1}, H_{\mathrm{k}}^i)$. For $\mathrm{k} < N-1$ go to step $\mathrm{k}+1$, else end.

Theorem 3. The sequence $\langle j_1,..., j_N \rangle$ constructed by Algorithm 1 defines the determining order J^i, and the subsequence $\left\langle j_{\mathrm{k}_1} = j_2, j_{\mathrm{k}_2},..., j_{\mathrm{k}_p} \right\rangle$, on which the value of the threshold u changes determines $\overline{\Gamma}^i$, where $\Gamma_{\mathrm{d}}^{\mathrm{i}} = \{j_1, j_2,..., j_{\mathrm{k_h}-1}\}$, $\mathrm{d} = \overline{1,\mathrm{p}}$.

A characteristic feature of this order is that it satisfies the condition

$$F(H_{\mathrm{k}}^i) = \pi(j_{\mathrm{k}+1}, H_{\mathrm{k}}^i),\ \mathrm{k} = \overline{1, N-1}. \tag{16}$$

Theorem 4. The sequence $\overline{\Gamma}^i$ of any determining order coincides with the maximum chain of zeros $K_{j_0}^i \supset ... \supset K_{j_{\mathrm{h}}}^i$ of semilattices $\mathrm{L}_{j_0}^i,...,\mathrm{L}_{j_{\mathrm{h}}}^i$ respectively, i.e., $K_{j_0}^i = \Gamma_{\mathrm{p}}^i,...,K_{j_{\mathrm{h}}}^i = \Gamma_1^i$ $(\mathrm{h} = \mathrm{p}-1)$.

The theorem asserts, in particular, that the order of the indices of the zeros of the corresponding semilattices is the inverse of the order of the indices of the subsets Γ_{k}^i.

Corollary 1. If J_1^i and J_2^i are two different determining orders, then the corresponding subsequences $\overline{\Gamma}_1^i$ and $\overline{\Gamma}_2^i$ coincide.

Corollary 2. $\overline{\Gamma}_{\mathrm{p}}^{i}$ is the only minimal core of the system $\langle W,\pi,F,i\rangle$.

It easy to see that the sequence constructed by Algorithm 1 coincides with the sequence constructed by the well-known Spectrum algorithm [4] if $\pi(i,H)$ is used as the measure of association of the element $i\in W\setminus H$ with the set H. Thus, Theorem 3 highlights certain extremal properties of sets in the sequence constructed by the Spectrum algorithm. Other extremal properties are noted in [13], where the Spectrum algorithm is used to find a classification corresponding to the local optimum of the weighted-average variance functional.

In the next section, the construction of this sequence is the basis for the general algorithm to find all the minimal cores of an I-system.

4. Finding the minimal cores

Let $\Omega=\{J\}$ be the set of all orders J on W, and let $\Pi(J,H)=\pi(i_J,H)$, where $H\in \mathrm{P}$, i_J is the first element of the set $W\setminus H$ on J. Then the following theorem holds.

Theorem 5. If the order J_*^i for all $i\in\{\overline{1,N}\}$ satisfies condition (16), then H_{k}^i ($\mathrm{k}\in\overline{1,N-1}$) taken from this order is a core of the system $\langle W,\pi,F,i\rangle$ if and only if for all $J\in\Omega$ and all $H\in \mathrm{P}^i$ we have the inequalities

$$\Pi(J,H_{\mathrm{k}}^i)\geq\Pi(J_*^i,H_{\mathrm{k}}^i)\geq\Pi(J_*^i,H). \qquad (17)$$

<u>Remarks</u>.

1. If condition (17) holds for any $H\in \mathrm{P}$, then H_{k}^i is a core of the system $\langle W,\pi,F\rangle$ (this is clear from the proof of the theorem).
2. If H^* is the minimal core of the monotone system $\langle W,\pi,F\rangle$, then for any $i\in H^*$ Algorithm 1 constructs an order J_*^i that satisfies (16) and such that for some $\mathrm{k}^*\in\{1,\ldots,N-1\}$ we have $H^*=H_{\mathrm{k}^*}$.

The algorithm to find all the minimal cores of $\langle W,\pi,F\rangle$ will be constructed by the following scheme: first for each of the systems $\langle W,\pi,F,i\rangle$, $i=1,\ldots,N$, find a minimal core H_i^*

using Algorithm 1. Then in this set of cores identify the sought set of minimal cores of the system $\langle W,\pi,F\rangle$. WE now give a formal description of the algorithm.

Algorithm 2.

Step 1. For each $i=1,...,N$, use Algorithm 1 to find $H_i^*=\Gamma_p^i$ and the number $F(H_i^*)$.

Step 2. From the family $\mathrm{H}=\{H_1^*,...,H_N^*\}$ separate the subfamily $\mathrm{H}'=\{H_{d_1}^*,...,H_{d_p}^*\}$ of subsets such that on each subset F attains its maximum value over the elements of H.

Step 3. Minimize H' by inclusion and retain only different elements, $\mathrm{H}^*=\{G_1^*,...,G_s^*\}$.

Theorem 6. The set of cores H^*, generated by Algorithm 2 contains all the minimal cores of the monotone system $\langle W,\pi,F\rangle$.

Algorithm 2, in view of the similarity between Algorithm 1 and the Spectrum algorithm notes in Sec. 2, suggests the following stronger form of the algorithm to construct a complete classification minimizing the weighted-average variance [13]. For each $i=1,...,N$, constructs the sequence J^i until the first class is obtained, as described in [13]. From these N alternatives select the one for which the identified class has the least variance. Remove the chosen class from W. Repeat the procedure until W has been exhausted. The algorithm is clearly independent of the choice of the first element i_1 (unlike the algorithm of [13]).

5. Quasiflow monotone systems

Consider a special class of monotone systems whose representatives are both I-systems and S-systems at the same time, i.e. the semilattices generated by their functions $F(H)$ are also lattices. This class was introduced in [7], where it is shown that all such systems are representable in the form [6]

$$\pi(i,H)=\min_{j\in H} a_{ij}, \tag{18}$$

$$F(H)=\min_{i\in W\setminus H}\ \min_{j\in H} a_{ij}, \tag{19}$$

where a_{ij} is the measure of association defined on the couples $(i,j): i,j\in W$.

---

[6] Comparison of (1)-(2) with (18)-(19) shows that the definition is that of an I-system. The fact that this is an S-system follows from the representation $\pi'(j,H)=\min_{i\in W\setminus H} a_{ij}$, $F(H)=\min_{j\in H}\pi'(j,H)$ and its comparison with (6) and (8).

The interpretation of $F(H)$ as the magnitude of the cut $(H, W \setminus H)$ on the graph defined on the vertex set W with the matrix of arc length $A = \|a_{ij}\|$ accounts for the term "quasi-flow" applied to these systems in [7]. Further analysis is restricted to a particular type of such systems, symmetric monotone quasiflow systems, when A is a symmetric matrix.

Thus, let $G = [W, E]$ be a graph on W without loops and parallel edges, with a system of lengths defined on its set of edges E by a symmetric matrix A; $F(H)$ is the lengths of the minimal edge joining H with $W \setminus H$. In this case, we obviously have $F(H) = F(W \setminus H)$. Recall that the minimal tree of a graph is its spanning subtree with a minimal sum of edge lengths.

Theorem 7. The values of $F(H)$ for all $H \subset W$ are equal to the length of some edge of the minimal tree of the graph G.

Note that the definition of a symmetric monotone quasiflow system does not require that the G be a complete graph. It suffices to have a connected graph, setting $a_{ij} = \infty$ for $(i, j) \notin E$.

An edge of maximum length will be called a maximal edge. We thus obtain:

Corollary. The value of $F(H)$ on a core of a monotone system is equal to the length of maximal edge in the minimal tree of the graph.

Let T be the minimal tree of the graph G and ℓ_0 the length of its maximal edges, ℓ_1 the length of the maximal among the remaining edges, [7] and so on, ℓ_m the length of the minimal edge of the tree T. Then by Theorem 7 the function $F(H)$ may take only $(m+1)$ different values $\ell_0, \ell_1, \ldots, \ell_m$ and the sets L_k $(k = 0, \ldots, m)$ are determined from the relationship

$$L_k = \{H \subset W \mid F(H) \geq \ell_k\}.$$

Clearly the sets L_k are Boolean lattices [8], because $F(H) = F(W \setminus H)$, i.e., each element of the lattice L_k has a complement. We know [8] that any element of a Boolean lattice

[7] For any two minimal trees of a graph, there exists an isomorphism that preserves the lengths of the edges [14].

has a unique irreducible representation in terms of $\vee$-irreducible elements [8] and the set of $\vee$-irreducible elements coincides with the set of atoms of the Boolean lattice – nonempty elements that are minimal by inclusion. Thus, the minimal cores of a symmetric monotone quasiflow system play the role of generators, i.e., any core is the union of minimal cores.

Theorem 8. The family of all atoms of the lattice L_k of a symmetric monotone quasiflow system coincides with the set of components of an arbitrary minimal tree of the graph G obtained when this tree is cut by all its edges, which are not shorter than ℓ_k.

Corollary. The set of minimal cores of a symmetric monotone quasiflow system forms a complete partitioning of the set W.

In order to find the family of atoms of the lattice L_k we need to construct the minimal tree on the graph G and to cut it by all the edges, which are not shorter than ℓ_k. This procedure is identical to the well-known procedure of partitioning into components the α-similarity graph [9] [4] with $\alpha = \ell_k$ or to the correlation Pleiades method [4]. It is also identical to the Wroclaw taxonomy method [3] with a given number of classes $m = m_k$, where $(m_k - 1)$ is the number of minimal tree edges not shorter than ℓ_k. More precisely, let R_k be the set of all possible partitions of W into m_k nonempty classes. Let W_S^ℓ be the ℓ-th class of some partition $R_s \in \mathsf{R}_k$, $\mathcal{L}(W_S^\ell)$ the length of the minimal tree of the subgraph G_S^ℓ defined by the set of vertices W_S^ℓ. On R_k consider the functional

$$J(R_s) = \sum_{\ell=1}^{m_k} \mathcal{L}(W_S^\ell). \tag{20}$$

Theorem 9. The functional (20) attains its minimum on the family of atoms of the lattice L_k.

The family of atoms of the lattice L_k coincides also with the set of classes obtained by the “nearest neighbor” method [5] on the step $|W| - m_k$ with a given number of classes m_k, because the classes corresponding to the nearest neighbor method are obtained by adding edges from the minimal tree in the order of increasing lengths, i.e., the minimal-length tree completely characterizes the agglomeration of classes by the nearest neighbor method.

---

[8] The element H ($H \neq \varnothing$) is called $\vee$-irreducible if $H_1 \cup H_2 = H$ implies $H_1 = H$ or $H_2 = H$.

[9] An α-similarity graph is a graph without edges that are not shorter that α.

APPENDIX

Let us prove the quasiconcavity property (4). Indeed, $F(H_1 \cap H_2) = \pi(i^*, H_1 \cap H_2)$, where $i^* \in W \setminus (H_1 \cap H_2)$. If $i^* \in W \setminus H_1$, then $\pi(i^*, H_1 \cap H_2) \geq \pi(i^*, H_1) \geq F(H_1)$ [from (1)]. Similarly, if $i^* \in W \setminus H_2$, then $\pi(i^*, H_1 \cap H_2) \geq F(H_2)$. We have proved (4). The quasiconvexity property (5) is proved similarly.

We will show that the family of subsets $\mathsf{H} = \{H \subset W : F(H) \geq u\}$, augmented with the empty set if necessary, is the lower semilattice. Indeed, let H_1, H_2 be such that $F(H_1) \geq u$, $F(H_2) \geq u$. If $H_1 \cap H_2 \neq \varnothing$, then $F(H_1 \cap H_2) \geq u$ [from (4)], i.e., $H_1 \cap H_2 \in \mathsf{H}$. If $H_1 \cap H_2 = \varnothing$, then let $F(\varnothing) \geq u$. Thus H is a lower semilattice.

<u>Proof of Theorem 1</u>. Since $\hat{\mathrm{k}}$ is the minimal index for which $\mathsf{L}_{\hat{\mathrm{k}}}$ contains nonintersecting sets, then for $\mathrm{i} < \hat{\mathrm{k}}$, L_i is a lower semilattice: the elements of this family satisfy the condition $F(H) \geq u_\mathrm{i}$, $\{\varnothing\} \notin \mathsf{L}_\mathrm{i}$. If $\mathrm{i} \geq \hat{\mathrm{k}}$, then $u_\mathrm{i} \leq u_{\hat{\mathrm{k}}}$ and so $F(\varnothing) \geq u_\mathrm{i}$ by construction; for $H \neq \varnothing$, $F(H) \geq u_\mathrm{i}$ from the definition of L_i. ■

<u>Proof of Theorem 2</u>. This follows from the inclusion of the semilattices $\mathsf{L}_\mathrm{s} \subseteq \mathsf{L}_{\mathrm{s}+1}$ ($\mathrm{s} = \overline{0, \mathrm{m}-1}$).

Let us prove property (11) [properties (10) and (12) are obvious. Assume that there exists H^* such that $K_{\mathrm{j_{d-1}}} \supset H^* \supset K_{\mathrm{j_d}}$, but $F(H^*) > F(K_{\mathrm{j_d}})$. Let i be the index such that $F(H^*) = u_\mathrm{i}$, then $u_{\mathrm{j_d}} < u_\mathrm{i}$. On the other hand, $u_\mathrm{i} < u_{\mathrm{j_{d-1}}}$, because $H^* \subset K_{\mathrm{j_{d-1}}}$. Thus $\mathrm{j_{d-1}} > \mathrm{i} > \mathrm{j_d}$. Then there exists K_i such that $K \supset K_{\mathrm{j_{d-1}}}$ and $K_\mathrm{i} \supset K_{\mathrm{j_d}}$, because $F(K_\mathrm{i}) \geq u_\mathrm{i} > F(K_{\mathrm{j_d}})$. A contradiction with the maximality of the chain. ■

<u>Proof of Theorem 3</u>. This is obvious.

<u>Proof of the Theorem 4</u>. It suffices to show that any element of the sequence $\overline{\Gamma}^i$ is a zero of some semilattice L_j^i and any zero K_j^i of the semilattice L_j^i coincides with some element from $\overline{\Gamma}^i$.

1. We will show that Γ_d^i ($\mathrm{d} \in \{\overline{1, \mathrm{p}}\}$) is a zero of the semilattice L_j^i defined by the condition $F(H) \geq u_j^i$, where $u_j^i = F(\Gamma_\mathrm{d}^i)$, $j \in \{\overline{0, \mathrm{r}}\}$. Consider $H \in \mathsf{P}^\mathrm{i}$. Assume that

$H \not\supseteq \Gamma_{\mathrm{d}}^{i}$, i.e., $\Gamma_{\mathrm{d}}^{i} \setminus H \neq \varnothing$, and let j_{s} be the first element from $\Gamma_{\mathrm{d}}^{i} \setminus H$ in the determining order J^{i}, i.e., $H_{\mathrm{s}-1}^{i} \subseteq \Gamma_{\mathrm{d}}^{i} \cap H$, $H_{\mathrm{s}}^{i} \not\subseteq \Gamma_{\mathrm{d}}^{i} \cap H$. Then

$$F(H) \overset{(a)}{\leq} \pi(j_{\mathrm{s}}, H) \overset{(b)}{\leq} \pi(j_{\mathrm{s}}, H_{\mathrm{s}-1}^{i}) \overset{(c)}{\leq} F(\Gamma_{\mathrm{d}-1}^{i}) \overset{(d)}{\leq} F(\Gamma_{\mathrm{d}}^{i}).$$

In this chain of inequalities, (a) holds by the definition of $F(H)$; (b) follows from the monotonicity of $\pi(j, H)$; (c) and (d) follow, respectively, from (14) and (13). Thus, if H satisfies the condition $F(H) \geq u_{j}^{i}$ (i.e., $H \in \mathrm{L}_{\mathrm{j}}^{\mathrm{i}}$), then $H \supseteq \Gamma_{\mathrm{d}}^{i}$. On the other hand, $\Gamma_{\mathrm{d}}^{i} \in \mathrm{L}_{j}^{\mathrm{i}}$, and so Γ_{d}^{i} is a zero of the semilattice $\mathrm{L}_{j}^{\mathrm{i}}$.

2. We will show that $u_{j}^{i} \leq F(\Gamma_{\mathrm{p}}^{i})$ for any $j = 0, \ldots, \mathrm{r}$. Assume that this is not so. Then there is H such that $F(H) > F(\Gamma_{\mathrm{p}}^{i})$. If $H \not\supseteq \Gamma_{\mathrm{p}}^{i}$, then from the proof of part 1 it follows that $F(H) < F(\Gamma_{\mathrm{p}}^{i})$. Let $H \supseteq \Gamma_{\mathrm{p}}^{i}$ and let j_{s} be the first element of the set $W \setminus H$ in the determining order J^{i}. Then

$$F(H) \leq \pi(j_{\mathrm{s}}, H) \leq \pi(j_{\mathrm{s}}, H_{\mathrm{s}-1}^{i}) \leq F(\Gamma_{\mathrm{p}}^{i}).$$

This chain is similar to the chain in the proof of part 1, except the last inequality, which follows from (15). Thus, for any $H \subset W$, $F(H) \leq F(\Gamma_{\mathrm{p}}^{i})$. Moreover, $\forall H \not\supseteq \Gamma_{\mathrm{p}}^{i} : F(H) < F(\Gamma_{\mathrm{p}}^{i})$ and $\forall H \supseteq \Gamma_{\mathrm{p}}^{i} : F(H) \leq F(\Gamma_{\mathrm{p}}^{i})$, i.e., $\Gamma_{\mathrm{p}}^{i} = K_{0}^{i}$.

Consider K_{j}^{i} ($j \in \{\overline{0, \mathrm{r}}\}$) and let $F(\Gamma_{\mathrm{d}-1}^{i}) < u_{j}^{i} \leq F(\Gamma_{\mathrm{d}}^{i})$, ($\mathrm{d} \in \overline{1, \mathrm{p}}$). We will show that in this case $K_{j}^{i} \supset \Gamma_{\mathrm{d}}^{i}$ and it therefore coincides with Γ_{d}^{i}, because $\Gamma_{\mathrm{d}}^{i} \in \mathrm{L}_{j}^{i}$. Assume that this is not so, i.e., $\Gamma_{\mathrm{d}}^{i} \setminus K_{j}^{i} \neq \varnothing$. Then from the proof of part 1 it follows that $F(K_{j}^{i}) \leq F(\Gamma_{\mathrm{d}-1}^{i}) < u_{j}^{i}$, a contradiction. ■

Corollary 1 is obvious, Corollary 2 follows from the fact that K_{0}^{i} is the unique minimal core of the system $\langle W, \pi, F, i \rangle$.

<u>Proof of Theorem 5</u>. *Sufficiency*. Let condition (17) hold, then $F(H) = \min_{j \notin H} \pi(j, H) \leq$ $\leq \Pi(J_{*}^{i}, H) \geq \Pi(J_{*}^{i}, H_{\mathrm{k}}^{i})$. On the other hand, $\Pi(J_{*}^{i}, H_{\mathrm{k}}^{i}) = \min_{J \in \Omega} \Pi(J, H_{\mathrm{k}}^{i}) = F(H_{\mathrm{k}}^{i})$. Thus, $F(H_{\mathrm{k}}^{i}) \geq F(H)$ $\forall H \in \mathrm{P}^{\mathrm{i}}$, i.e., H_{k}^{i} is a core of the system $\langle W, \pi, F, i \rangle$.

Necessity. Consider the couple $\left(J_*^i, H_k^i\right)$. By (16) it follows that $\Pi(J_*^i, H_k^i) = F(H_k^i) = = \min_{J\in\Omega} \Pi(J, H_k^i)$. Thus $\Pi(J, H_k^i) \geq \Pi(J_*^i, H_k^i)$.

Consider $H \in \mathsf{P}^i$, and let j_{d+1} be the first element of the set $W \setminus H$ in the order J_*^i. Then $H \supseteq H_d^i$ and

$$\Pi(J_*^i, H) = \pi(j_{d+1}, H) \overset{(a)}{\leq} \pi(j_{d+1}, H_d^i) = F(H_d^i) \overset{(b)}{\leq} F(H_k^i) = \Pi(J_*^i, H_k^i).$$

In this chain, (a) follows from (1) and (b) from (9). ■

Proof of Theorem 6. Let $H = \{i_1, i_2, \ldots, i_t\}$ be a minimal core, then it coincides with the minimal core of the system $\langle W, \pi, F, i_1 \rangle$ and it is therefore generated in the first step of Algorithm 2. Then clearly H' is the set of cores of the system $\langle W, \pi, F \rangle$ containing all the minimal cores, and these will be obtained in the third step.

Proof of Theorem 7. Let T be a minimal tree, $F(H) = a_{i^* j^*}$, $e = \left(i^*, j^*\right)$, but $e \notin T$. Since H and $W \setminus H$ form a partition of the vertex set of the graph G, then there is an edge $e' \in T$, connecting H and $W \setminus H$. From the definition of $F(H)$ it follows that $\ell(e) \leq \ell(e')$, where $\ell(e)$ is the length of the edge e. The edge e together with the tree T form a cycle that contains e'. Thus $T \setminus e' \cup e$ is also a tree whose length is $\ell(T) + \ell(e) - \ell(e')$ where $\ell(T)$ is the length of the tree T. From the minimality of T it follows that $\ell(e) = \ell(e')$.

Proof of Theorem 8. We will first show that the assertion of the theorem is well-defined, i.e., it is independent of the particular choice of minimal tree.

LEMMA. Let T_1 and T_2 be two minimal trees and $\mathsf{R}_1 = \{R_1^1, \ldots, R_p^1\}$, $\mathsf{R}_2 = \{R_1^2, \ldots, R_s^2\}$ two partitions obtained when trees T_1 and T_2 are cut, respectively, by all the edges that are not shorter than ℓ_k. Then R_1 and R_2 coincide.

Proof of Lemma. Assume that this is not so. Then there is a component of the tree T_1 (or T_2) that intersects with two components of the tree T_2 (T_1) so that these two components are connected by the edge e of the tree T_1 (T_2), and $\ell(e) < \ell_k$. For definiteness, let $e \in T_1$. Adjoining e to T_2, we obtain a cycle that contains at least one edge $e' \in T_2$ joining the components of T_2, and therefore its length is $\ell(e') \geq \ell_k$. Then we can reduce the length of the tree T_2, by replacing e' by e. But this is a contradiction with the minimality of T_2.

Let us now prove the theorem. Let T contain $(t-1)$ edges of length $\ell \geq \ell_k$ and let $\mathsf{R} = \{R_1, \ldots, R_t\}$ be the partition obtained when T is cut by these edges. From the proof of Theorem 7 it follows that $F(R_i) \geq \ell_k$ $(i = 1, \ldots, t)$, i.e., R_i is an element of the lattice L_k. We will show that R_i is an atom of the lattice L_k. Assume that contrary, i.e., there exists an element $G \in \mathsf{L}_k$, $G \neq \varnothing$, such that $G \subset R_i$. But $F(G)$ is not greater than the length of the edge of the tree T joining G with $R_i \setminus G$, i.e., $F(G) < \ell_k$, a contradiction. Thus, R_i is an atom of the lattice L_k.

Other atoms of the lattice L_k that are not components of the partition obviously do not exist. This follows directly from the fact that the components R_i $(i = 1, \ldots, t)$ form a complete partition of the set W. ■

Proof of Theorem 9. From the graph G and the partition $R \in \mathsf{P}_k$ construct the graph $G_R = [R, E']$ whose vertices are the partition classes and the edges are the shortest edges of the graph G joining a given pair of classes. The minimal tree T_R of the graph G_R is termed the minimal tree of the partition R, $\ell(T_R)$ is the length of the tree T_R. Let T_s^ℓ be the minimal tree of the subgraph G_s^ℓ of the graph G. Then $\bigcup_{\ell=1}^{m_k} T_s^\ell \cup T_{R_s}$ is a tree of the graph G and therefore $J(R_s) + \ell(T_{R_s}) \geq \ell(T)$, where $\ell(T)$ is the length of a minimal tree of the graph G. Thus, the minimum of the criterion $J(R)$ is attained for maximum $\ell(T_R)$. We will show that the maximum value of $\ell(T_R)$ is attained on the family of atoms of the lattice L_k. To this end, we need the following lemma.

LEMMA. For any minimal tree T_R of the partition R there exists a minimal tree T of the graph G such that all the edges of T_R are edges of T.

Proof of Lemma. Assume the contrary: there exists a partition R and the tree T_R such that for any minimal tree T of the graph G there is an edge of T_R that does not belong to T. Then consider the tree T that has the maximal intersections with T_R. Let $e \in T_R$, $e \notin T$, $e = (i^*, j^*)$. Then there is a path in G connecting i^* and j^*, and all the edges of this path are not longer than $\ell(e)$ by minimality of T. On the other hand, there exists an edge $e' \in T$

of this path which joins two components of the tree T_R obtained by cutting the edge e, and $e' \notin T_R$. By minimality of T_R, $\ell(e') \geq \ell(e)$ and therefore $\ell(e') = \ell(e)$. Then replacing the edge e' with e in T, we obtain the minimal tree T' of the graph G, which has the largest number of identical edges with T_R. A contradiction. ■

The lemma implies that the maximum value of $\ell(T_R)$, $(R \in \mathrm{R}_k)$ is equal to the sum of the $m_k - 1$ longest edges of the minimal tree of the graph G, i.e., the sum of lengths of all the edges of the minimal tree which are not shorter than ℓ_k. Thus, the maximum of $\ell(T_R)$ and the minimum of $J(R)$ are attained on the family of atoms of the lattice L_k. ■

LITERATURE CITED

1. E. V. Bauman, A. A. Dorofeyek, and A. L. Chernyavskii, “Methods of structural analysis of empirical data,” *Izmer., Kontrol’, Avtomatizatsiya, No. 12*, 57-69 (1985).
2. J. C. Besdek, Pattern Recognition with Fuzzy Objective Functions Algorithms, *Plenum Press*, New York (1980).
3. B. G. Mirkin, Analysis of Qualitative Attributes and Structures [in Russian], *Statistika*, Moscow (1980).
4. É. M. Braverman and I. B. Muchnik, Structural Methods of Data Analysis [in Russian], *Nauka*, Moscow (1983).
5. T. Éshankulov, “Identification of special elements in the solution of taxonomic problems,” *Izv. Akad. Nauk UzSSR, No. 2*, 57-60 (1983).
6. M. M. Kamilov and T. Éshankulov, “A classification method with identification of special elements, “ 9th *All-Union Conf. on Control Problems*, abstracts of papers [in Russian], Erevan (1983), pp.141-142.
7. A. V. Malishevskii, “On the properties of ordinal set functions,” *All-Union Seminar on Optimization and Its Applications*, abstracts of papers [in Russian], Dushanbe (1986), English translation http://www.datalaundering.com/download/order.pdf .
8. G. Graetser, General Lattice Theory [Russian translation], *Mir*, Moscow (1982).
9. J. E. Mullat, “Extremal subsystems of monotone systems. I,” *Avtomat, Telemekh., No. 5*, 130-139 (1976), http://www.datalaundering.com/download/extrem01.pdf .
10. E. N. Kuznetsov and I. B. Muchnik, “Analysis of the distribution of functions in an organization,” *Avtomat. Telemekh., No. 10*, 119-127 (1982), http://www.datalaundering.com/download/organiza.pdf .
11. J. E. Mullat and L. K. Vyhandu, “Monotonic systems in scene analysis,” *Symp. Math. Processing of Cartographic Data*, Tallinn (1979), pp. 63-66.
12. E. N. Kuznetsov, I. B. Muchnik, and L. V. Shvartser, “Monotone systems and their properties,” in: Analysis of Nonnumerical Information in Social-Science Research [Russian translation], *Mir*, Moscow (1985), pp. 29-57.
13. B. G. Mirkin, “Additive clustering and qualitative factor analysis methods for similarity matrices,” *J. Classification, No. 4*, 5-29 (1987).
14. M. Swami and K. Thulasiraman, Graphs, Networks, and Algorithms [Russian translation], *Mir*, Moscow (1984).
15. J. Raisin (ed.), Classification and Clustering [Russian translation], *Mir*, Moscow (1980).