SIMULATION OF BEHAVIOR AND INTELLIGENCE

Russian version: http://www.datalaundering.com/download/extrem01-ru.pdf

Extremal Subsystems of Monotonic Systems I. [i]

J. E. Mullat

Department of Economics, Tallinn Technical University (1973 – 1980)

Translated from Avtomatica i Telemekhanika, No. 5, pp. 130 – 139, May, 1976.
Original article submitted August 13, 1975.

Correspondence Address: Byvej 269
2650 Hvidovre
Copenhagen Denmark
Phone (45) 36 49 33 05
E-mail: jm@mail-telia.dk

Abstract

A general theoretical method is described which is intended for the initial analysis of systems of interrelated elements. Within the framework of the model, a specially postulated monotonicity property for systems guarantees the existence of a special kind of subsystems called kernels. A number of extremal properties and the structure of the kernels are found. The language of description of monotonic systems of interrelated elements is described in general set-theoretic terms and leads to a constructive system of notions in the case of systems with finite number of elements. A series of properties of special finite sequences of elements are studied whereby kernels in monotonic systems are classified.

Keywords: monotonic, system, matrix, graph, cluster

1. Introduction

For the study of a complex system, it is often necessary to encounter the problem of analyzing concrete numerical data about the system functioning. Sometimes based on similar data it is required to show whether in the system there exist special elements or subsystems, reacting in one way to some “actions” as well as “relations” between one-type subsystems. Information on the existence of the indicated peculiarities or on the “structure” of the system under study is necessary, for example, before carrying out extensive or expensive statistical investigation.

Concerning wide application of computational techniques, at the present time, to initial detection of the structure of a system an approach based on various kind of heuristic models is planned [1-4]. For constructing models, many authors start with intuitive formulations of the problem and also with the form of presentation of the initial data [5,6].

A natural form of presentation the data for the purpose of studying complex systems is that of a graph [7]. A matrix, for example, a data matrix [8] also serves as a widely spread carrier of information. Matrices and graphs easily admit isolation of two minimal structural units of the system: “elements” and “connections” between elements [1]. In this paper the notions “connections” and “elements” are interrelated in a sufficiently broad fashion. Thus, sometimes it is desirable to consider connections in the form of elements of a system; in this case, it is possible to find more “subtle” relations in the original system.

Representation of a system in the form of unique object – elements and connections between elements – makes it possible to attach a more precise meaning to the problem of revealing the structure of a system. A structure of a system is an organization of elements of the system into subsystems, which is put together in the form of a set of relations between subsystems. A structure can, for example, be a natural way of combining subsystems into a single system, which is determined on the basis of “strong” and “weak” connections between

[1] Analogous systems are called systems of interrelated elements in the literature.

elements of the system. A similar approach to the analysis of systems is described, for example, in [9], where the question of assembling systems of interrelated elements is considered. Assembling turns out to be a convenient macro-language foe expressing a structure of the system.

In the theory of systems, usually direct connections between elements are considered. Situation, however, sometimes requires considering indirect connections as well. This requirement is distinguished thus: that indirect connections are dynamic relations in the sense that “degree” of dependence is determined by a subsystem, in which this or that connection is considered. Below we describe and study a certain subclass of similar “dynamic” systems called monotonic systems.

The monotonicity property for systems allows us to formulate in a general form the concept of a kernel of a system as a subsystem, which in the originally indicated sense reflects the structure of the whole system in the large. A kernel represents a subsystem whose elements are “sensitive” in the highest degree to one of two types of actions (positive or negative), since “sensibility” to actions is determined by the intrinsic structure of the system. The definition of positive and negative actions reduces to the existence of two type of kernels – positive and negative kernels.

Existence of kernels (special subsystems) is guaranteed by the mathematical model described in this paper and the problem of “isolating” kernels is typical problem in the description of a “large” system in the language of a “small” system – kernel. In this sense, figuratively speaking, a kernel of a system is a subsystem whose removal inflicts “cardinal” changes the properties of that system: The system "gives up" the existing structure.

For exposition of the material terminology and symbolism, the theory of sets is used which requires no special knowledge. One should turn attention to the special notation introduced, since the apparatus developed in this paper is new.

2. Examples of Monotonic Systems [2]

1. In the n-dimensional vector space let there be given N vectors. For each pair of vectors x and y one can define in many ways a distance $\rho(x,y)$ between these vectors (i.e., to scale the space). Let us assume that the set of given vectors forms an unknown system W.

For every vector in an arbitrary subsystem of W we calculate the sum of distances to all vectors situated inside the selected subsystem. Thus, with the respect to each subsystem of W and each vector situated inside that subsystem, a characteristic sum of distances is defined, which can be different for different subsystems.

It is not difficult to establish the following property of the set of sums of distances. Because of removing a vector from the subsystem the sums computed for the remaining vectors decrease while because of adding a vector to the subsystem they increase. A similar property of sums for every subsystem of system W is called in this paper the monotonicity property and a system W having such a property is called a monotonic system.

2. For studying schools, directions in various branches of science, the so-called graphs of cited publications [10] are used. These are directed a-cyclic graphs, since each author can cite only those authors whose papers are already published. It is entirely reasonable to assume that the set of publications W forms a certain system, where the system elements (published papers) are exchanged with each other by information and by special way, namely, by the help of citation. If we consider a subset from an available survey of the set of publications W, then each publication can be characterized by the number of bibliographical tittles, taken only over the subset − subsystem − considered. It is clear that "removal" of publication from the subsystem only decreases the quantitative evaluation thus introduced for the degree of exchange of information in the subsystem while the "addition" of a publication in the subsystem only increases that evaluation for all publications in the subsystem. Thus, we have here a monotonic citation system given in the form of a graph.

---

[2] Kempner, Y., Mirkin, B., and Muchnik, I. B., "Monotone linkage clustering and quasi-concave set functions," Applied Mathematics Letters, **1997**, 4, 19-24, http://www.datalaundering.com/download/kmm.pdf ; B. Mirkin and I. Muchnik, "Layered Clusters of Tightness Set Functions," Applied Mathematics Letters, **2002**, v. 15, issue no. 2, pp. 147-151. http://www.datalaundering.com/download/mm012.pdf ; see also, A. V. Genkin (Moscow), I. B. Muchnik (Boston), "Fixed Approach to Clustering, Journal of Classification," Springer, **1993**, 10, pp. 219-240, http://www.datalaundering.com/download/fixed.pdf ; and latest connection, Kempner, Y., Levit V. E., "Correspondence between two antimatroid algorithmic characterizations," Dept. of Computer Science, Holon Academic Institute of Thechnology, July, **2003**, Israel, http://www.datalaundering.com/download/0307013.pdf .

In connection with the above example, it is interesting to note [11], where the author involuntarily considers an example of a monotonic system in the form of a directed graph.

3. Let us assume that there is a set W of telephone exchanges or points of connection that are joined by lines of two-sided connections. Under the absence of any connection between points in a system with communications, it is possible to organize a transit connection. If a functioning of a similar system is observed for a long time, then the "quality" of connection" between each pair of points can be expressed, independently of whether there exists a two-sided connection or not, by the average number of "denials" in establishing a connection between them in a standard unit of time. Generally speaking, if it is desired to characterize each point of the system W in the sense of "unreliability" of establishing connections with other points, then this second characteristic can be taken to be the average number of denials in establishing connection with at least one point of the system in a unit time. It is clear that these same numerical qualities (quality of connection, unreliability characteristic) can be defined only inside every subsystem of the system with communications W.

The proposed model has the following obvious properties. A gap in any line of two-sided connection increases the average number of denials among all other points of connection; introduction of any new line, in contrast decreases the average number of denials. This is related with the fact that load on the realization of a transit connection in a telephone communication network increases (decreases). In the case of curtailment of activity at any point of connection inside the given subsystem the unreliability of all points of subsystem increases while in case of addition of a point of connection to the subsystem the unreliability decreases.

Thus, there is a complete similarity with the examples of monotonic systems considered above and one can state that the model described for telephone communications is a monotonic system.

In the present paper a monotonic system is defined, to be a system over whose elements one can perform "positive" and "negative" actions. In addition, positive actions increase certain quantitative indicators of the functioning of a system while the negative actions decrease those indicators. In the second example considered above the positive action is the addition of an element to a subsystem while the negative action is removing an element from the subsystem; in the third example the converse holds.

In the second and third examples above, the kernel must have an intuitive meaning. Thus, in the citation graphs, a negative kernel must turn out to be the set of publications citing each other in a considerable degree (by authors representing a single scientific school) while a positive kernel must consist of publications citing each other to a lesser degree (representing different schools).

In telephone communications networks the intuitive sense of a kernel must manifest itself in the following. If we take as elements of a communication network the lines of connection, then a negative kernel is a collection of lines that give on the average a "mutually agreed upon" large number of denials while a positive kernel has the opposite sense – a collection of lines that give on the average less denials. In case the system elements are taken to be the connection points of a telephone communication network, a negative kernel is a set of mutually unreliable points while a positive kernel is a set of more reliable points.

The intuitive meaning given to kernels of citation graphs and communication network is not based on a sufficient number of experimental facts. The indicated properties are noted in analogy with available intuitive interpretation of kernels obtained for solutions of automatic-classification problems [12].

3. Description of a Monotonic System

One considers some system W consisting of a finite number of elements, [3] i.e., $|W| = N$, where each element α of the system W plays a well-defined role. It is supposed that the states of elements α of W are described by definite numerical quantities characterizing the "significance" level of elements α for the operation of the system as a whole and that from each element of the system one can construct some discrete actions.

We reflect the intrinsic dependence of system elements on the significance levels of individual elements. The intrinsic dependence of elements can be regarded in a natural way

[3] If W is a finite set, then $|W|$ denotes the number of its elements.

as the change, introducible in the significance levels of elements β, rendered by a discrete action produced upon element α.

We assume that the significance level of the same element varies as a result of this action. If the elements in a system are not related with each other in any way, then it is natural to suppose that the change introduced by element α on significance β (or the influence of α on β) equals zero.

We isolate a class of systems, for which global variations in the significance levels introduced by discrete actions on the system elements bears a monotonic character.

Definition. By a monotonic system, we understand a system, for which an action realized on an arbitrary element α involves either only decrease or only increase in the significance levels of all other elements.

In accordance with this definition of a monotonic system two types of actions are distinguished: type ⊕ and type ⊖. An action of type ⊕ involves increase in the significance levels while ⊖ involves decrease.

The formal concept of a discrete action on an element α of the system W and the change in significance levels of elements arising in connection with it allows us to define on the set of remaining elements of W an uncountable set of functions whenever we have at least one real significance function $\pi : W \to D$ (D being the set of real numbers).

Indeed, if an action is rendered on element α, the starting from the proposed scheme one can say that function π is mapped into π_α^+ or π_α^- according as a the action is of type ⊕ or ⊖. Significance of system elements is redistributed as action on element α changes from function π to π_α^+ (π_α^-) or, otherwise, the initial collection of significance levels $\{\pi(\partial) | \partial \in W\}$ changes into a new collection $\{\pi_\alpha^+(\partial) | \partial \in W\}$ [4]. Clearly, if we are given some sequence $\alpha_1, \alpha_2, \alpha_3, ...$ of elements of W (arbitrary repetitions and combinations of

[4] Functions π, π_α^+ and π_α^- are defined on the whole set W and, consequently, $\pi_\alpha^+(\partial)$ and $\pi_\alpha^-(\partial)$ are defined.

elements being permitted) and the binary sequence $+,-,+,...$, then by the usual means one can define the functional product of functions $\pi^+_{\alpha_1}, \pi^-_{\alpha_2}, \pi^+_{\alpha_3}$ in the form $\pi^+_{\alpha_1}\pi^-_{\alpha_2}\pi^+_{\alpha_3}$.

The construction presented allows us to write the property of monotonic systems in the form of the following basic inequalities:

$$\pi^+_{\alpha}(\partial) \geq \pi(\partial) \geq \pi^-_{\alpha}(\partial) \tag{1}$$

for every pair of elements $\alpha,\partial \in W$, including the pairs α,α or ∂,∂.

Let there be given a partition of set W into two subsets, i.e., $H \cup \overline{H} = W$ and $H \cap \overline{H} = \varnothing$. If we subject the elements $\alpha_1,\alpha_2,\alpha_3,... \in \overline{H}$ to positive actions only, then by the same token on set W there is defined some function $\pi^+_{\alpha_1}\pi^-_{\alpha_2}\pi^+_{\alpha_3}...$, which can be regarded as defined only on the subset H of W. [5]

If from all possible sequences of elements of set $\overline{H}$ we select a sequence $\left\langle \alpha_1,\alpha_2,...,\alpha_{|\overline{H}|} \right\rangle$[6], where α_i are not repeated, then on the set H function $\pi^+_{\alpha_1}\pi^+_{\alpha_2}...$ is induced univalently.

We denote this function π^+H and call it a standard function. We shall also refer to the function thus introduced as a weight function and to its value on an element as an α weight.

In accordance with this terminology the set $\{\pi^+H(\alpha) \mid \alpha \in H\}$, which is denoted by Π^+H is called a weight collection given on the set H or a weight collection relative to set H. Let us assume that we are given a set of weight collections $\{\Pi^+H \mid H \subseteq W\}$ on the set of all possible subsystems $P(W)$ of system W. The number of all possible subsystems is $|P(W)| = 2^{|W|}$.

Instead of considering a standard function for positive actions $\pi^+_{\alpha_1}\pi^+_{\alpha_2}...$ one can consider a similar function for negative actions π^-H. Thus, by exact analogy one defines single weight

---

[5] We are not interested in significance levels obtained as a result of operations on elements of $\overline{H}$ onto the same set $\overline{H}$.

[6] Here symbols $\langle , \rangle$ are used to stress the ordered character of a sequence of $\overline{H}$.

collection $\Pi^{-}H = \left\{ \pi^{-}H(\alpha) \,\middle|\, \alpha \in H \right\}$ and the aggregate of weight collections $\left\{ \Pi^{-}H \,\middle|\, H \subseteq W \right\}$.

Let us briefly summarize the above construction. Starting with some real function π defined on a finite set W and using the notion of positive and negative actions on elements of system W, one can construct two types of aggregate collections $\Pi^{+}H$ and $\Pi^{-}H$ defined on each of the H of subsets of W. Each function from the aggregate (weight collection) is constructed by means of the complement to H, equaling $W \setminus H$, and a sequence $\left\langle \alpha_1, \alpha_2, \ldots, \alpha_{|\overline{H}|} \right\rangle$ of distinct elements of the set $\overline{H}$. For this actions of types $\oplus$ and $\ominus$ are applied to all elements of set $\overline{H}$ in correspondence with the ordered sequence $\left\langle \alpha_1, \alpha_2, \ldots, \alpha_{|\overline{H}|} \right\rangle$ in order to obtain $\Pi^{+}H$ and $\Pi^{-}H$ respectively.

The concept of weight collections $\Pi^{+}H$ and $\Pi^{-}H$ needs refinement. The definition given above does not taken into account the character of dependence of function πH on the sequence of actions realized on the elements of set $\overline{H}$ [7]. Generally speaking, weight collection $\Pi^{+}H(\Pi^{-}H)$ is not defined uniquely, since it can happen that for different orderings of set $\overline{H}$ we obtain different function πH .

In order that weight collection $\Pi^{+}H(\Pi^{-}H)$ be uniquely defined by subset H of the set W it is necessary to introduce the notion of commutativity of actions.

Definition. An action of type $\oplus$ or $\ominus$ is called commutative for system W if for every pair of elements $\alpha, \beta \in W$ we have

$$\pi^{+}_{\alpha}\pi^{+}_{\beta} = \pi^{+}_{\beta}\pi^{+}_{\alpha}\,,\ \pi^{-}_{\alpha}\pi^{-}_{\beta} = \pi^{-}_{\beta}\pi^{-}_{\alpha}$$

In this case it is easy to show that the values of function πH on the set H do not depend on any order defined for the elements of the set $\overline{H}$ by sequence $\langle \alpha_1, \alpha_2, \ldots \rangle$. The proof can be conducted by induction and is omitted.

[7] In the sequel, if sign "–" or "+" is omitted from our notation, then it is understood to be either "–" or "+"

Thus, for commutative actions the function $\pi^{+}H(\pi^{-}H)$ is uniquely determined by a subset of W.

In concluding this section, we make one important remark of an intuitive character. As is obvious from the above-mentioned definition of aggregates of weights collection of type $\oplus$ and $\ominus$, the initial weight collection serves as the basic constructive element in their construction. The initial weight collection is a significance function defined on the set of system elements before the actions are derived from the elements. In other words, it is the initial state of the system fixed by weight collection ΠW. It is natural to consider only those aggregates of weight collections that are constructed from an initial $\oplus$ collection, which is the same as the initial $\ominus$ collection. The dependence indicated between $\oplus$ and $\ominus$ weight collections is used considerably for the proof of the duality theorem in the second part of this paper.

4. Extremal Theorems. Structure of Extremal Sets [8]

Let us consider the question of selecting a subset from system W whose elements have significance levels that are stipulated only by the internal "organization" of the subsystem and are numerically large or, conversely, numerically small. Since this problem consists of selecting from the whole set of subsystems $P(W)$ a subsystem having desired properties, therefore it is necessary to define more precisely how to prefer one subsystem over another.

Let there be given aggregates of weight collections $\left\{\Pi^{+}H \mid H \subseteq W\right\}$ and $\left\{\Pi^{-}H \mid H \subseteq W\right\}$. On each subset there are defined the following two functions:

$$F_{+}(H) = \max_{\pi \in H} \pi^{+}H(\alpha),\ F_{-}(H) = \min_{\pi \in H} \pi^{-}H(\alpha).$$

Definition of Kernels. By kernels of set W we dub the points of global minimum of function F_{+} and of global maximum of function F_{-}.

[8] See also, Muchnik, I., and Shvartser, L. (1990), "Maximization of generalized characteristics of functions of monotone systems," Automation and Remote Control, 51, 1562-1572, http://www.datalaundering.com/download/maxgench.pdf.

A subsystem, on which F_+ reaches a global minimum, is called a ⊕ kernel of the system W, while a subsystem on which F_- reaches a global maximum, is called ⊖ kernel.

Thus, in every monotonic system the problem of determining ⊕ and ⊖ kernels is raised.

With the purpose of intuitive interpretation as well as with the purpose of explaining the usefulness of the notion of kernels introduced above we turn once again to the examples of citation graphs and telephone commutation networks.

The definition of a kernel can be formulated with the help of a significance levels of system elements, that is: a ⊕ kernel is a subsystem of a monotonic system, for which a maximal level among significance levels stipulated only by internal organization of the subsystem is minimal, and a ⊖ kernel is a subsystem for which a minimal level among those same significance levels is maximal.

The definition of a kernel accords with the intuitive interpretation of a kernel in citation graphs and telephone commutation networks. Thus, in citation graphs a ⊕ kernel is a subset (subsystem) of publications, in which the longest list of bibliographical titles is at the same time very short; though not inside the subset, but among all possible subsets of the selected set of publications (among the very long lists). If in our subset of publications a very short list of bibliographical titles is at the same time very long among the very short ones relative to all the subsets, then it is a ⊖ kernel of the citation graph. It is clear that a ⊖ kernel publications cite one another often enough, since for each publication the list of bibliographical titles is at any rate not less than a very short one while a very short list is nevertheless long enough. In a ⊕ kernel the same reason explains why in this subset one must find representatives of various scientific schools.

In telephone commutation networks, one can consider two types of system elements – lines of connections and points of connections. In a system consisting of lines, a ⊖ kernel turns out to be a subset of lines, for which the lines with the least number of denials in that subset are at the same time the lines with the greatest number of denials among all possible sets of lines. This means that at least the number of denials stipulates only by the internal organization of a subnetwork of lines of a ⊖ kernel is not less than the number of denials for lines with the smallest number of denials and, besides, this number is large enough. Hence one can expect that the number of denials for lines of a ⊖ kernel is sufficiently large. Similarly one should expect a small number of denials for lines of a ⊕ kernel. Formulation for ⊕ and ⊖ kernels for points of connection is exactly the same as for the lines and is omitted here.

Before stating the theorems, we need to introduce some new definitions and notations.

Let $\overline{\alpha} = \langle \alpha_0, \alpha_1, ..., \alpha_{k-1} \rangle$ be an ordered sequence of distinct elements of set W, which exhausts the whole of this set, i.e., $k = |W|$. From sequence $\overline{\alpha}$ we construct an ordered sequence of subsets of W in the form

$$\Delta_{\overline{\alpha}} = \langle H_0, H_1, ..., H_{k-1} \rangle$$

with the help of the following recurrent rule

$$H_0 = W,\ H_{i+1} = H_i \setminus \{\alpha_i\};\ i = 0,1,...,k-2$$ [9]

Definition. Sequence $\overline{\alpha}$ of elements of W is called a defining sequence relative to the aggregate of weights collections $\{\Pi^- H \mid H \subseteq W\}$ if there exists in sequence $\Delta_{\overline{\alpha}}$, a subsequence of sets

$$\Gamma_{\overline{\alpha}} = \langle \Gamma_0^-, \Gamma_1^-, ..., \Gamma_p^- \rangle,$$

such that:

a) weight $\pi^- H_i(\alpha_i)$ of an arbitrary element α_i in sequence $\overline{\alpha}$, belonging to set Γ_j^- but not belonging to set Γ_{j+1}^- is strictly less than values of $F_-(\Gamma_{j+1})$ [10];

b) in set Γ_p^- there does not exist a proper subset L which satisfies the strict inequality

$$F_-(\Gamma_p) < F_-(L).$$

[9] Sign $\setminus$ denotes the subtraction operation for sets.

[10] Here and everywhere, for simplification of expression, where it is required, the sign "–" or "+" is not used twice in notations. We should have written $F_-(\Gamma_{j+1}^-)$ or $F_+(\Gamma_{j+1}^+)$.

A sequence $\overline{\alpha}$ with properties a) and b) is denoted by $\overline{\alpha}_-$. One similarly defines a sequence $\overline{\alpha}_+$.

Definition. Sequence $\overline{\alpha}$ of elements of W is called a defining sequence relative to the aggregate of weights collections $\left\{ \Pi^+ H \mid H \subseteq W \right\}$ if there exists in sequence $\Delta_{\overline{\alpha}}$, a subsequence of sets

$$\Gamma_{\overline{\alpha}} = \left\langle \Gamma_0^+, \Gamma_1^+, \ldots, \Gamma_q^+ \right\rangle,$$

such that:

c) weight $\pi^+ H_i(\alpha_i)$ of an arbitrary element α_i in sequence $\overline{\alpha}$, belonging to set Γ_j^+ but not belonging to set Γ_{j+1}^+ is strictly greater than values of $F_+(\Gamma_{j+1})$;

d) in set Γ_q^+ there does not exist a proper subset L which satisfies the strict inequality

$$F_+(\Gamma_q) > F_+(L).$$

A sequence $\overline{\alpha}$ with properties a) and b) is denoted by $\overline{\alpha}_-$. One similarly defines a sequence $\overline{\alpha}_+$.

Definition. Subset H_+^* of set W is called definable if there exists a defining sequence $\overline{\alpha}_+$ such that $H_+^* = \Gamma_q^+$.

Definition. Subset H_-^* of set W is called definable if there exists a defining sequence $\overline{\alpha}_-$ such that $H_-^* = \Gamma_p^-$.

Below we formulate, but do not prove, a theorem concerning properties of points of global maximum of function F_-. The proof is adduced in Appendix 1. A similar theorem holds for function F_+. In Appendix 1 the parallel proof for function F_+ is not reproduced. The corresponding passage from the proof for F_- to that of F_+ can be effected by simple interchange of verbal relations “greater than” and “less than”, inequality signs “≥” and “≤”, “>”, “<” as well as by interchange of signs “+” and “–”. The passage from definable set H_+^* to H_-^* and from definition of sequence $\overline{\alpha}_+$ and $\overline{\alpha}_-$, is effected by what has just been said.

Theorem 1. On a definable set H_-^* function F_- reaches a global maximum. There is a unique definable set H_-^*. All sets, on which a global maximum is reached, lie inside the definable set H_-^*.

Theorem 2. On a definable set H_+^* function F_+ reaches a global minimum. There is a unique definable set H_+^*. All sets, on which a global minimum is reached, lie inside the definable set H_+^*.

In the proof of Theorem 1 (Appendix 1) it is supposed that definable set H_-^* exists. It is natural that this assumption, in turn, needs proof. The existence of H_-^* is secured by a special constructive procedure. [11]

The proof of Theorem 2 is completely analogous to the proof of Theorem 1 and is not adduced in Appendix 1.

We present a theorem, which reflects a more refined structure of kernels of W as elements of the set $P(W)$ of all possible subsets (subsystems) of set W.

Theorem 3. The system of all sets in $P(W)$, on which function F_- (F_+) reaches maximum (minimum), is closed with the respect to the binary operation of taking union of sets.

The proof of this theorem is given in Appendix 2 and only for the function F_-. The assertion of the theorem for F_+ is established similarly.

Thus, it is established that the set of all $\ominus$ kernels ($\oplus$ kernels) forms a closed system of sets with respect to the binary operation of taking the unions. The union of all kernels is itself a large kernel and, by the statements of Theorems 1 and 2, is a definable set.

[11] This procedure will be presented in the second part of the article, since here only the extremal properties of kernels and the structure of the set of kernels are established.

APPENDIX 1

Proof of Theorem 1. We suppose that a definable set H_-^* exists.

(Conducting the proof by contradiction) let us assume that there exists a set $L \subseteq W$, which satisfies the inequality

$$F_-(H^*) \le F_-(L). \tag{A.1}$$

Thus two sets H_-^* and L are considered. One of the following statements holds:

1) Either $L / H_-^* \neq \varnothing$, which signifies the existence of elements in L, not belonging to H_-^*;

2) or $L \subseteq H_-^*$.

We first consider 2). By a property of definable set H_-^* there exists a defining sequence $\overline{\alpha}_-$ of elements of set W with the property b) (cf. the definition of $\overline{\alpha}_-$) such that the strict inequality $F_-(H^*) < F_-(L)$ does not hold and, consequently, only the equality holds in (A.1). In this case, the first and the third statements of the theorem are proved. It remains only to prove the uniqueness of H_-^*, whish is done after considering 1).

Thus, let $L / H_-^* \neq \varnothing$ and let us consider set H_t – the smallest of those H_i $(i = 0,1,...,k-1)$ from the defining sequence $\overline{\alpha}_-$ that include the set L / H_-^*. Then the fact that H_t is the smallest of the indicated sets implies the following: there exists element $\lambda \in L$, such that $\lambda \in H_t$, but $\lambda \notin H_{t+1}$.

Below, we denote by $i(\Omega)$ the smallest of the indices of elements of defining sequence $\overline{\alpha}_-$ that belong to the set $\Omega \subseteq W$.

Let Γ_p^- be the last in the sequence of sets $\langle \Gamma_j^- \rangle$, whose existence is guaranteed by the sequence $\overline{\alpha}_-$. For indices t and $i(\Gamma_p^-)$ we have the inequality $t < i(\Gamma_p^-)$.

The last inequality means that in sequence of sets $\langle \Gamma_j^- \rangle$ there exists at least one set Γ_s^-, which satisfies

$$i(\Gamma_{s+1}^-) \ge t + 1. \tag{A.2}$$

Without decreasing generality, one can assume that Γ_s^- is the largest among such sets.

It has been established above that $\lambda \in H_t$, but $\lambda \notin H_{t+1}$. Inequality (A.2) shows that $\Gamma_s^- \subset H_{t+1}$, since the opposite assumption $\Gamma_s^- \supseteq H_{t+1}$ leads to the conclusion that $i(\Gamma_s^-) \geq t+1$ and, consequently Γ_s^- is not the largest of the sets, for which (A.2) holds.

Thus, it is established that $\Gamma_{s-1}^- \supset H_t$. Indeed, if $\Gamma_{s-1}^- \subseteq H_t$, then for indices $i(\Gamma_{s-1}^-)$ and t we have $i(\Gamma_{s-1}^-) \geq t$.

Hence $i(\Gamma_{s-1}^-)+1 \geq t+1$ and the inequality $i(\Gamma_s^-) \geq i(\Gamma_{s-1}^-)+1$ implies $i(\Gamma_s^-) \geq t+1$. The last inequality once again contradicts the choice of set Γ_s^- as the largest set, which satisfies inequality (A.2).

Thus, $\lambda \notin \Gamma_s^-$, but $\lambda \in \Gamma_{s-1}^-$, since $\lambda \in H_t$, $H_t \subseteq \Gamma_{s-1}^-$. On the basis of property a) of the defining sequence $\overline{\alpha}_-$, we can conclude that

$$\pi^- H_t(\lambda) < F_-(\Gamma_s), \tag{A.3}$$

where $0 \leq s \leq p$.

Let us consider an arbitrary set Γ_j^- $(j = 0,1,\ldots,p-1)$ and an element $\tau \in \Gamma_j^-$, which has the smallest index in the sequence $\overline{\alpha}_-$. In other words, set Γ_j^- starts from the element τ in sequence $\overline{\alpha}_-$. In this case, set Γ_j^- is a certain set H_i in the sequence of imbedded sets $\langle H_i \rangle$. The definition of $F_-(H)$ and the property a) of defining sequence $\overline{\alpha}_-$ implies that

$$F_-(\Gamma_j) \leq \pi^- \Gamma_j(\tau) < F_-(\Gamma_{j+1}).$$

Hence

$$F_-(\Gamma_0) < F_-(\Gamma_1) < \ldots < F_-(\Gamma_p)$$

and as a corollary we have for $j = 0,1,\ldots,p$

$$F_-(\Gamma_j) \leq F_-(\Gamma_p) = F_-(H^*), \tag{A.4}$$

since $\Gamma_p^- = H_-^*$.

Let $\mu \in L$ and let weight $\pi^- L(\mu)$ be minimal in the collection of weights relative to set L. On the basis of inequalities (A.1), (A.3), and (A.4) we deduce that

$$\pi^- H_t(\lambda) < \pi^- L(\mu) = F_-(L). \qquad \text{(A.5)}$$

Above, H_t was chosen so that $L \subseteq H_t$. Recalling the fundamental monotonicity property (1) for collection of weights (the influence of elements on each other), it easy to establish that

$$\pi^- L(\lambda) \le \pi^- H_t(\lambda). \qquad \text{(A.6)}$$

Inequalities (A.5) and (A.6) imply the inequality

$$\pi^- L(\lambda) < \pi^- L(\mu),$$

i.e., there exists in the collection of weights relative to set L a weight which is strictly less than the minimal weight.

A contradiction is obtained and it is proved that set L can only be a subset of H_-^* and that all sets, distinct from H_-^*, on which the global maximum is also reached, lie inside H_-^*.

It remains to prove that if a definable set H_-^* exists, then it is unique. Indeed, in consequence of what has been proved above we can only suppose that some definable set $H_-^{'}$, distinct from H_-^*, is included in H_-^*.

It is now enough to adduce arguments for definable set $H_-^{'}$ similar to those adduced above for L, considering it as definable set $H_-^{'}$; this implies that $H_-^* \subseteq H_-^{'}$. The theorem is proved.

APPENDIX 2

Proof of Theorem 3. Let Ω be the system of set in $P(W)$, on which function F_- reaches a global maximum, and let $K_1 \in \Omega$ and $K_2 \in \Omega$.

Since on K_1 and K_2 the function F_- reaches a global maximum, therefore we might establish the inequalities

$$F_-(K_1 \cup K_2) \leq F_-(K_1),\ F_-(K_1 \cup K_2) \leq F_-(K_2). \qquad \text{(A.7)}$$

We consider element $\mu \in K_1 \cup K_2$, on which the value of function F_- on set $K_1 \cup K_2$, is reached, i.e.,

$$\pi^- K_1 \cup K_2(\mu) = \min_{\alpha \in K_1 \cup K_2} \pi^- K_1 \cup K_2(\alpha).$$

If $\mu \in K_1$, then by rendering $\ominus$ actions on all those elements of set $K_1 \cup K_2$, that do not belong to K_1, we deduce from the fundamental monotonicity property of collections of weights (1) the validity of the inequality

$$\pi^- K_1(\mu) \leq \pi^- K_1 \cup K_2(\mu).$$

Since the definition of F_- implies that $F_-(K_1) \leq \pi^- K_1(\mu)$ and by the choice of element μ we have $\pi^- K_1 \cup K_2(\mu) = F_-(K_1 \cup K_2)$, therefore we deduce the inequality

$$F_-(K_1) \leq F_-(K_1 \cup K_2).$$

Now from the inequality (A.7) it follows that

$$F_-(K_1) = F_-(K_1 \cup K_2).$$

If, however, it is supposed that $\mu \in K_2$, then $\ominus$ actions are rendered on elements of $K_1 \cup K_2$, not belonging to K_2; in an analogous way we obtain the equality

$$F_-(K_2) = F_-(K_1 \cup K_2),$$

which was to be proved.

L I T E R A T U R E   C I T E D

1. E. M. Braverman, N.E. Kiseleva, I.B. Muchnik, and S.G. Novikov, “Linguistic approach to the problem of processing large massifs of information,“ *Avtom. Telemekh., No. 11*, 73 (1974).

2. W.McCormick, P. Schweitzer, and T. White, “Problem decomposition and data recognition by clustering technique,“ *J. Oper. Res., 20*, No. 5, 993 (1972).

3. S. Deutsch and J.J. Martin, “An ordering algorithm for analysis of data arrays,“ *J. Oper. Res., 19*, No. 6, 1350 (1971).

4. T. Zahn, “Graph-thoeretical methods for detecting and describing gestalt clusters,“ *IEEE Trans. Comput., C-20,* No. 1, 68 (1971).

5. L.K. Vykhandu, “Investigation of multisign biological systems,“ *in: Application of Mathematical Methods in Biology, Vol. III*, Isd. LGU (1964), pp. 19-22.

6. P.V. Terent’ev, “Method of correlation galaxies,“ *Vestn. LGU, No. 9*, 137 (1959).

7. I.B. Muchnik, “Analysis of structure of extremal graphs,“ *Avtom. Telemekh., No. 9,* 62 (1974).

8. J. Hartigan, “Direct Clustering of Data Matrix,“ *J. Amer. Statist. Assoc., 67*, No. 337, 123 (1972).

9. E.M. Braverman, A.A. Dorofeyuk, V. Ya. Lumel’skij, and I.B. Muchnik, “Diagonalization of connection matrix and detection of latent factors,“ *in: Problems of Expanding Possibilities of Automata, No. 1*, Izd. In-ta Probl. Ipravl., Moscow (1971), pp. 42-79.

10. V.V. Nalimov and Z.M. Mul’chenko, Science-metry [in Russian], *Nauka* (1969).

11. A.O. Trybulets, “Document-graph method in the classification of sciences,“ in*: Application of Universal Computers in the Work of Data Organs* (*Symposium Proceedings*, Moscow, June, 1967), Izd. VINITI, (1970).

12. E. Ojaveer, J. Mullat, and L.Võhandu, “A study of infraspecific groups of the Baltic east coast autumn herring by two new methods based on cluster analysis,“ *in: Proceedings of Estonian Studies According to International Biology Program VI* [in Russian], Isd. Akad. Nauk ESSR, Tartu (1975).

[*i] The name “Monotonic System” at that moment in the past was the best match for our scheme. However, this name “Monotone System” was already occupied in “Reliability Theory” unknown to the author. Below we reproduce a fragment of a “monotone system” concept different from ours in lines of Sheldon M. Ross “Introduction to Probability Models”, Fourth Ed., Academic Press, Inc., pp. 406-407.

Example 2d (A four-Component Structure):

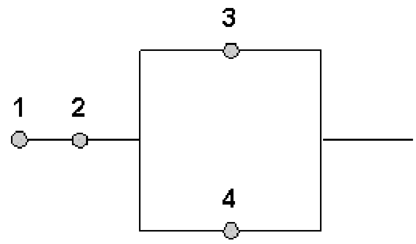

Figure. 9.4

Consider a system consisting of four components, and suppose that the system functions if and only if components 1 and 2 both function and at least one of components 3 and 4 function. Its structure function is given by

$$\phi(\mathrm{x}) = x_1 \cdot x_2 \cdot \max(x_3, x_4).$$

Pictorially, the system is shown in Figure 9.4. A useful identity, easily checked, is that for binary variables, (a binary variable is one which assumes either the value 0 or 1) x_i, $i = 1,...,n$,

$$\max(x_1,...,x_n) = 1 - \prod_{i=1} (1 - x_i)$$

When $n = 2$, this yields

$$\max(x_1, x_2) = 1 - (1 - x_1) \cdot (1 - x_2) = x_1 + x_2 - x_1 \cdot x_2 .$$

Hence, the structure function in the above example may be written as

$$\phi(\mathrm{x}) = x_1 \cdot x_2 \cdot (x_3 + x_4 - x_3 \cdot x_4) \ \diamond$$

It is natural to assume that replacing a failed component by a functioning one will never lead to a deterioration of the system. In other words, it is natural to assume that the structure function $\phi(\mathrm{x})$ is an increasing function of x, that is, if $x_i \le y_i$, $i = 1,...,n$, then $\phi(\mathrm{x}) \le \phi(\mathrm{y})$. Such an assumption shall be made in this chapter and the system will be called *monotone*.