Heaks kiidetud TPI nõukogus
15. märtsil 1977, protokoll nr. 9

ТАЛЛИНСКИЙ ПОЛИТЕХНИЧЕСКИЙ ИНСТИТУТ
Кафедра обработки информации

МЕТОДЫ ВСКРЫТИЯ СТРУКТУРЫ ДАННЫХ

Методическое руководство

Составитель И. Муллат

На эстонском языке

Koostanud J. Mullat

Vastutav toimetaja J. Vilipõld
da antud 11. IX 1978. Paber 60x84/16
2,0. Tingtrükipg. 1,86. Tiraaž 500
cint, Tallinn, Koskla 2/9. Tell. 769

Tasuta

Data Structure Opening Method: Methodological Guide *

Abstract. This methodological study delves into an extensively documented yet powerful monotonicity-based information processing technique that is often overlooked despite its widespread application and contemporary use. The focus is on the application of the category of monotonicity to formal systems for data analysis, a method with a simple algorithmic component in uncovering complex data structures and obtaining information in various fields, including sociology, economics, biology and demography. This methodology recognizes patterns in two basic data structures: frequency tables and graphs. Frequency tables arise as a common outcome of surveys when data are organized into categorical responses. The effectiveness of the method depends on converting these categories into frequency measures, which facilitates in-depth analysis based on numerical indicators. This preparatory step lays the foundation for robust analysis that allows researchers to gain detailed information about social trends, consumer behavior and economic models. The application of the method extends to the field of graph theory, where comprehensive patterns in complex networks are modeled. By emphasizing the construction of generalized models, this approach illuminates the fundamental characteristics of reality through visualization of so-called "encompassing pictures." This framework focuses on key metrics such as saturation levels and the presence or absence of important components such as triangles and cycles in graphs. By carefully studying these graph structures, researchers can unravel complex relationships, identify emergent phenomena, and elucidate the underlying mechanisms governing system behavior.
Keywords: data matrix; layering algorithm; graph; tournament

1. INTRODUCTION

If one decides to collect data, the following questions must first be answered:

- What information is needed?
- Why is this information needed?
- To what extent are the reasons for gathering information?
- How can decisions be made based on the information gathered and thus influence the investigation process?

If answers are available, then the set of collected *"objects"*, those data, is also defined. For example, information may concern people living in a city, families in a given country, electronic equipment, factories made up of basic production units (objects in the terminology of the guide) etc.

Population information can be composed of a series of indicators that describe the population as a whole, such as the scales against which income is measured. In productive area, indicators determine the technical environment in which, e.g. equipment was manufactured and operated. Naturally, estimates based on the information collected differ from actual estimates. Thus, the researcher may draw incorrect conclusions if the error of the estimate is too large. This guide looks at one possible way to avoid the errors associated with the so-called stratification concept.

* The original version in Estonian (Mullat, 1977), Protocol No. 9, approved by the TPI Council (Tallinn Polytechnic Institute) on March 15, 1977. TPI currently stands for Tallinn University of Technology – TalTech.

Let's give an example of the importance of this concept in information processing: in USA a presidential elections were held in 1932. Literary Digest sent postcards to voters with questions to predict Roosevelt's election to the presidency. Some 10 million postcards had been sent out. The results showed that the forecast made on the basis of the information collected was accurate within 1%. However, the prediction made using exactly the same technique in 1936, contained an error of almost 20%.

There is a general perception that the "postcard method" introduced a disproportion among voters who return postcards. It turned out that people with higher education and better conditions tended to return more postcards. People with a higher standard of living tended to prefer Roosevelt's competitor during the readiness period, and the forecast of results shifted away from the real thing.

This example shows that when the population is stratified (for example, only voters with higher education and better conditions are observed), a big mistake cannot be avoided. That is, in order to avoid such an error, the researcher must know in advance the subgroups of the population (classification), but usually the identification of subgroups is a complex and voluminous effort, which in turn is associated with the collection of information.

The guide looks at population stratification (classification) methods that currently exist in three types:

a) Methods that take into account the researcher's subjective opinion of the population. This means that classification with exact properties are known or simply assumed;
b) Methods to be used in the absence of any data or hypotheses about existing strategies and their attributes;
c) Methods, which are intended only to visualize a sample of the population in order for the researcher to be able to make a decision on the available strata.

Among methods a), b) and c), only the so-called monotonic layering (Mullat, 1971-1995) or known since then as the “monotonic linkage method” (Kempner et al, 1997) is considered. The last chapters are devoted to the theoretical study of these monotone systems and methods of monotone layering, in particular, on graphs. We do not discuss issues related to the use of standard statistical methods and algorithms. The additional tools and technologies needed for the monotone layering of data, the accompanying terminology and strict nomenclature are explained in the course of the narrative and defined where necessary.

The article consists of an introduction and a section that discusses the main concepts, a total of 8 sections. Section 3 discusses the different types of metric distances between objects to measure the difference between objects in classification problems. Section 4 describes the method itself at an informal level. Section 5 provides a more accurate construction at a precise mathematical level. In Sections 6-7, we consider the application of the method to the study of graphs, in particular, to determine the groups of strong players in tournaments as opposed to weak players. Concluding remarks are provided in Section 8.

2. KEY DEFINITIONS

First, we introduce the reader to the terminology and basic concepts used. The basic concept of data processing is a data matrix. The data X is a $n \times m$ matrix (n row and m columns), each row of which is called an object; one column of the matrix is called an attribute. This means that the data matrix is:

$$X = \begin{Vmatrix} x_{1,1}, x_{1,2}, .., x_{1,m} \\ x_{2,1}, x_{2,2}, .., x_{2,m} \\ ..... \\ x_{n,1}, x_{n,2}, .., x_{n,m} \end{Vmatrix}$$

and $x_{i,j}$ is the value of the j-th attribute of the i-th object. It is natural that the question immediately arises as to what the numerical values of the attribute in the data matrix reflect? There must be brands that the attributes may differ substantially. For example, the air temperature may be a characteristic when electric lamps are lit; the shoe number of the person; gender (male or female), etc. As the processing is formally based on mathematical apparatus, three types of attributes are distinguished in order to be able to interpret the final results and use them according to the purpose:

a) Attributes on a continuous scale (Interval scale), such as body credential, height, temperature (quantitative);

b) Attributes on a discrete ordinal scale, such as the grades a student receives in some subjects: unsatisfactory, satisfactory, good, and very good. At this point, the values of the attribute are considered ordered (in Points or ranked);

c) Attributes with discrete values that are not ranked (nominal scale or even qualitative attributes), For example, eye color, gender (male or female).

2.1. Quantitative attributes

The quantitative expression of an attribute is usually referred to as the value of the attribute can be compared. Questions about how many times the value of one attribute is greater than another can be answered. At first glance, the question does not seem to be very complicated, although a deeper examination in turn raises the question: "What is natural to compare?" Let's look at some more examples before answering this question.

Let us choose the cars that are described by the price tag. Undoubtedly, the attribute "price" is quantitative, the a car with the price of 1O.OOO€, is twice as expensive as the b car with the price of 5.000€. The characteristic "price" or "value" expressed by the function $f(a)$ can also be expressed by the function $\kappa \cdot f(a))$ (κ is a positive number). Every other type of conversion changes the price ratio of cars. The allowed transformations of the attribute “price” are multiplication by the constant κ. This property of the price makes it possible to determine how many times $f(a)$ is greater than $f(b)$ — the ratio $\kappa \cdot f(a) \big/ \kappa \cdot f(b)$ does not depend on κ of the choice, and if κ is fixed, we can thus say how much is $f(a)$ greater than $f(b)$. This class of transformations allows for the universal presentation of concepts related to quantitative as well as other types of attributes. However, the determination of a unit of measurement requires only quantitative attributes.

2.2. Definition

The permissible transformation of the value of an attribute $f(a)$ in the set of attributes $\mathcal{A}$ is called the function $\varphi(x)$ if the function $\varphi(f(a))$ $(a \in \mathcal{A})$ shows the same attribute. If the values of the characteristic f are given together with the number of allowed conversions F, then we say that the measurements of the characteristic were performed on the F-type scale.

In the example of passenger cars $F_o\{\kappa \cdot x \mid \kappa > 0\}$ and on the scale F_0 it is usually said that the measurements are made on a ratio scale. An interval scale is a scale where the number of transformations allowed is $F_x = \{\kappa \cdot x + o \mid \kappa > 0\}$; the specific scale F_x is determined by the quantities κ and o, which give the unit of measurement and the starting point of the scale.

In most cases, the measurement results are presented in the form of a matrix, if after each allowed transformation the measurement results do not change. However, it should be noted that the results of matrix measurements do not allow them to be immediately used in arithmetic calculations. For example, the relationship $f(a) + f(b) > f(c)$ does not make sense in the scale with origin $o > 0$, since $\kappa \cdot [f(a) + f(b)] + 2 \cdot o$ is greater than $\kappa \cdot f(c) + o$ only for some κ and o values. Indeed, absolute zero is the natural and unambiguous presence of the zero point o that cannot be changed: °0-Kelvin is absolute zero on the scale, which characterizes the absence of the measured feature. However, °0-Celsius or °0-Fahrenheit are not. Two arbitrary physical phenomena are taken here: melting of ice, or an equal mixture of water, ice and salt at -21.1°C. Comparing the mean values of the interval scale is another matter.

Expression

$$\frac{1}{n}\cdot\sum_{i=1}^{n} f(a_i) > \frac{1}{m}\sum_{j=1}^{m} f(b_j) \qquad (1)$$

remains unchanged after using the allowed conversion. Namely

$$\frac{1}{n}\cdot\sum_{i=1}^{n} \kappa\cdot f(a_i)+o > \frac{1}{m}\sum_{j=1}^{m} \kappa\cdot f(b_j)+o \text{ iff}$$

$$\frac{\kappa}{n}\cdot\sum_{i=1}^{n} f(a_i)+\frac{o\cdot n}{n} > \frac{\kappa}{m}\sum_{j=1}^{m} f(b_j)+\frac{o\cdot m}{m}$$

and the latter is equivalent to inequality (1). It makes sense to compare the absolute differences in the values of the attributes, namely

$$\frac{|f(a)-f(b)|}{|f(c)-f(d)|}=\frac{|(\kappa\cdot f(a)+o)-(\kappa\cdot f(b)+o)|}{|(\kappa\cdot f(c)+o)-(\kappa\cdot f(d)+o)|}.$$

Now we ask the question what determines the number of allowed transformations $f(x)$? Usually the choice is related to other attributes with the possibility of forecasting. Formally expressed laws of science allow all these forecasting transformations not to change the law. For example, Clipperon's law $P\cdot V/T = \text{const}$ connects the scales of temperature T, volume V and pressure P of a given gas, allows transformation, leaving the law unchanged. Also in economics, in functional models, the price is determined fixed to within a multiplier.

Unknown patterns of relationships, characteristic of sociological or psychological research, allow transformations between objects in the form of empirical relationships, for example, by stratification methods. In these studies, however, interval or ratio scales are unacceptable.

2.3. Point or ordinal scales.

Pupil assessment aims to test the degree of skill acquisition and achievement of primary education goals on a point scale: Fail (IN – Insuficiente); Pass (SU – Suficiente), Good (BI – Bien), Very Good (NT – Notable), Excellent (SB – Sobresaliente). Point scale gradations are limited by equal intervals of discrete numerical values. Expert judgments are often recorded as a sequence of natural numbers arranged symmetrically to the o point $(o, \pm 1, ...)$.

A distinction should be made between two types of point estimates. In the first case, the assessments reflect some well-known standards. The more opportunities you have to describe and characterize standards, the more accurately you can, for example, determine the deviation from the standard. Thus, the teacher depending on his work experience and personal experience forms the pedagogical level of high school students' performance. On the other hand, refining a benchmark helps predict attribute values; for example, a student who is very good at geometry usually also scores higher in algebra.

The second type of points occurs when there are no well-known standards or even the existence of an objective criterion is questionable, which may be reflected in subjective judgments, for example, the taste of culinary products. This type is also called an ordinal or ordered scale. The set of allowed transformations F consists of all monotonically increasing functions. The ordered values of the attributes are compared only on the basis of the relation "higher-lower". It is meaningless to compare the differences between the values of the attribute. For example:

if $f(a)=10$, $f(b)=2$, $f(c)=1$, $f(a)-f(b)=8$, $f(b)-f(c)=1$, $f(a)-f(b)=8>f(b)-f(c)=1$, Then, using the monotonic transformation φ, where $\varphi(1)=1$, $\varphi(2)=20$, $\varphi(10)=30$ gives a contradiction $10=\varphi(f(a))-\varphi(f(b))>\varphi(f(b))-\varphi(f(c))=19$.

It is, nonetheless, realistic to fix the values of original attributes using non-numerical terms. Eligible elements for each ordered set, such as alphabet, etc.

(c) The nominal scale. The scales of the above attributes — quantitative, point and ordinal scales — have general attributes. All scales define the binary relation B on the set of objects X. The relation is defined by the following rule: $(a,b)\in B$ then and only then when $f(a)>f(b)$. Quantitative and point measurements are informatively more voluminous than ordinary measurements.

In practice, we can often only be interested in the information contained in the binary relation B. The researcher's conclusions about the functioning of the socio-economic system are usually qualitative (for example, stratification or ranking of objects in a sample).

It is natural to ask the question: is qualitative information not enough to draw conclusions? Qualitative information is easier to measure and more reliable. We do not have the means to accurately measure $f(a)$ and $f(b)$, while we can be sure that $f(a)>f(b)$.

On the other hand, the complex examination of data requires the transformation of the measurement results of individual assessments and objective indicators into a common type of data: quantitative or qualitative.

By limiting the number of transformations F allowed, complex data analysis is usually performed by quantifying all measurements. By limiting the number of transformations allowed sophisticated data analysis is usually performed by quantifying all measurements. Qualitative measurements can "suffer" in this way. When examining qualitative data, it is also possible to do the opposite: to transform quantitative measurements into qualitative ones. It is possible that even then the data will "suffer". However, if the results using quantitative methods are consistent with the results of qualitative data processing methods, the investigator is more likely to be sure of the conclusions reached,

Let the equivalence relation $\mathcal{J}$ be given for the cross product of objects $X \times X$. We assign to each object $x \in X$ the number of the i-th class of X, which contains the object X. Let's say that the measurements are made on a nominal scale, if the value of the attribute is the number of the i-th equivalent relation. Number of conversions allowed by F_n are unique functions. Thus the pair $(a,b) \in \mathcal{J}$ then and only then when attributes values $f(a) = f(b)$. Measurement on a nominal scale is the "weakest" measurement step, as it is only determined whether the equation $f(a) = f(b)$ truly applies.

3. METHODS FOR MEASURING DIFFERENCES BETWEEN OBJECTS

All of the methods that we will discuss in Sections 4-7 relate to some degree to the concept of distance or metric. This means that the task of stratification can be performed accurately only if the distance between objects is determined. Choosing a distance means also comparing distances that measure the similarity of two objects. The higher this number, the more the objects themselves differ, and vice versa.

The distance $\rho(x,y)$ between objects x and y is called a function that satisfies three conditions:

(a) for each x object $\rho(x,x) = 0$;

(b) for each pair (x,y) of objects $\rho(x,y) = \rho(y,x)$;

(c) there is a relationship for each of the three objects (x,y,z) that $\rho(x,y) + \rho(y,z) \geq \rho(y,z)$.

The following is a list of metrics or distances used. The notations are as follows: We denote the i-th, $i = \overline{1,n}$, object of the data matrix X as $x_i = \langle x_{i,1}, x_{i,2}, ..., x_{i,m} \rangle$, where $x_{i,j}$, $j = \overline{1,m}$, is the j-th attribute of the object i. The distance between two objects x_k and x_ℓ herein as said is nominated as $\rho(x_k, x_\ell)$.

Here are some of the most commonly used metrics.

Cubic distance: $$\rho(x_k, x_\ell) = \max_{j=1,m} \left| x_{k.j} - x_{\ell,j} \right|,$$

where $|\cdot|$ indicates an absolute value.

Octahedral distance: $\rho(x_k, x_\ell) = \sum_{j=1}^{m} \left| x_{k.j} - x_{\ell,j} \right|$.

Euclidean distance, $\rho(x_k, x_\ell) = \sqrt{\sum_{j=1}^{m} \left(x_{k.j} - x_{\ell,j} \right)^2}$

These three metrics are mostly useful for an interval scale. The following distance is useful when attributes are measured in points or on an ordinal scale:

$$\rho(x_k, x_\ell) = \sum_{j=1}^{m} \left| x_{k.j} - x_{\ell,j} \right| \Big/ \sum_{j=1}^{m} \max_{k,\ell} \left(x_{k,j}, x_{\ell,j} \right).$$

There are distances that are valid when the attributes are binary. Binary is a sign of "marital" status, e.g. if there can be only two answers — "married-yes" or "married-no". These distances are valid even if the scale is nominal.

3.1. Hamming distance

The notation is borrowed from set theory because objects can be interpreted as subsets of attributes. A value of 1 can be viewed as an indicator $x_{i,j}$ of whether the original attribute j belongs or does not belong to subset x_i. The object x_i is thus a Boolean vector $x_i = \left\langle x_{i,1},..., x_{i,m} \right\rangle$, where $x_{i,j}$ is the "1"-one or "0"-zero type, $j = \overline{1,m}$.

The absolute distance $\rho(x_k, x_\ell)$ is defined as follows: $\rho(x_k, x_\ell) = m - \left| x_k \cap x_\ell \right|$, which equals the number of missing matches in the objects x_k, x_ℓ. In this case, $\left| x_k \cap x_\ell \right|$ is the number of attributes matches in the data matrix, which takes into account 1-s in both objects x_k, x_ℓ, indicating the same attributes. The relative distance looks like $\rho(x_k, x_\ell) = 1 - \left| x_k \cap x_\ell \right| / \left| x_k \cup x_\ell \right|$, where $x_k \cup x_\ell$ is a set of only those attributes that are present in both x_k, x_ℓ objects, but do not necessarily indicate the same attributes.

The list of distances between objects can be continued, since the possibilities for determining the distances are not limited. It should only be noted that the choice of distances is a process that is difficult to formalize and is usually performed by a researcher based on his/her own experience. Measuring the differences or distances between attributes further complicates matters and differs from the above list. Inter-trait, or correlation coefficient between features/attributes is the most commonly used measure that shows the relative linearity of the change in a second identifier when the first identifier changes. The correlation coefficient c between attributes α, β can be determined using the following expression:

$$c_{\alpha,\beta}=\frac{\sum_{i=1}^{n}x_{i,\alpha}\cdot x_{i,\beta}-\left(\sum_{i=1}^{n}x_{i,\alpha}\cdot\sum_{i=1}^{n}x_{i,\beta}\right)\Big/n}{\sqrt{\sum_{i=1}^{n}x_{i,\alpha}^{2}-\left(\sum_{i=1}^{n}x_{i,\alpha}^{2}\right)^{2}\Big/n}\cdot\sqrt{\sum_{i=1}^{n}x_{i,\beta}^{2}-\left(\sum_{i=1}^{n}x_{i,\beta}^{2}\right)^{2}\Big/n}}$$

In the case of the attributes "no", "yes", it is useful to apply a binary (Pirson's φ) correlation r between objects κ, ℓ in the form of:

$$r_{\kappa,\ell}=\frac{\left|x_k\cap x_\ell\right|\cdot\left|\overline{x}_k\cap\overline{x}_\ell\right|-\left|x_k\cap\overline{x}_\ell\right|\cdot\left|\overline{x}_k\cap x_\ell\right|}{\sqrt{\left|x_\kappa\right|\cdot\left|\overline{x}_\ell\right|\cdot\left|\overline{x}_\kappa\right|\cdot\left|x_\ell\right|}},$$

where $\overline{x}$ is a complement of x; $\left|x_\kappa\right|\cdot\left|\overline{x}_\ell\right|\cdot\left|\overline{x}_\kappa\right|\cdot\left|x_\ell\right|>0$. Before selecting the distance/correlation between objects, one must perform a Class F independence check of the permitted transformations.

4. DATA LAYERING ALGORITHM

The reader is probably aware that many models of automatic stratification or objective classification are given and described in the literature. We also know that quite a lot of algorithms of this type have been developed, but due to the lack of access to such knowledge, we independently developed and studied here only one, possibly new for many, method. This method is primarily intended for sociological data, but it can also be used to process the general data matrix X.

Let the information gathered be presented in a form that can depict a large graph. For example, some cities are divided into many quarters. The researcher collects information from the city's residents on movements from one quarter to another. Thus, quarters occur on top of a graph (graph) on the vertices of a graph. The arcs of Graph indicate where the local movements of the population are directed in the city. The task is to find out the movements global trends. So the task is basically in that not to stratify city quarters, but stratify possible directions of movement.

Let's match the number to each arrow (arc) in the graph indicating how many transit paths of length 2 the arrow around gives. Graphically, this means that the number of triangles attached to the arc of the graph has been enumerated, (Fig. 1).

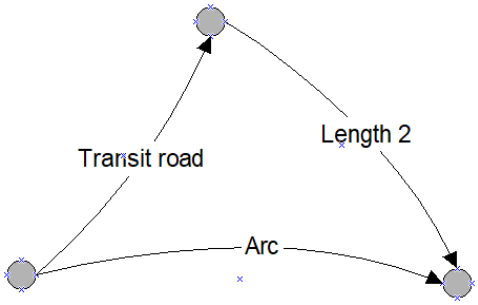


Figure 1

When this is done, the stratification of the arrows (arcs) is completed using the following algorithm. Mullat developed this algorithm in 1971-1977. Everywhere, if necessary, we will call this algorithm using the abbreviation KSF — "Kernel Searching Routine".

1. **Zero step**

   Find the arc with the least number of triangles on the graph and set it to the value of the parameter u at the level u_0. The arc is removed from the graph. It may be that the removal operation at this point affects some other arcs in the graph and the number of triangles viewed on them changes, so that some other arcs with credentials become less than or equal to u_0. These arcs are also removed. This removal of arc or set of arcs shall be repeated until there are no more arcs whose credentials satisfy the condition: less than or equal to u_0,

2. **Recursive k-th step**

a) From the graph that developed in the previous $k-1$ steps when used, a new minimum credential arc, such as an arc with a minimum number of triangles but higher than previous u_{k-1} is found. The parameter u level u_k, $u_{k-1} < u_k$ of the credential of this arc remembers the level. The arc or arcs found is or are removed from the graph.

b) It may be that the removal operation in current step k affects some more arcs and that their credentials become less or equal to u_k. We repeat this "peeling" until there is no more arcs with credentials less or equal to u_k. All arcs are on some p-th step removed/reset from the graph. This terminates the algorithm.

As a result of the algorithm, all arcs of the graph are distributed into groups or layers, each of which is linked with the corresponding size (threshold) u_k, $k = \overline{(0,p)}$. Observing these groups from the last, p-th group, the researcher can draw conclusions about the global or major movement directions on the graph.

Example Let this graph be in Fig. 2,

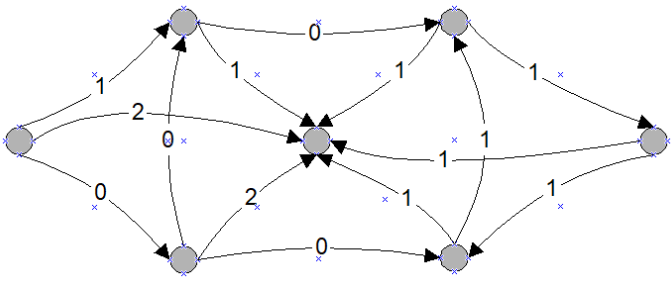


Figure 2.

This figure shows the transit number of routes defined above by Fig. 1 around with the arc in Fig. 2. According to the algorithm the performance of the zero step is the shape of the graph as shown in the Fig. 3.

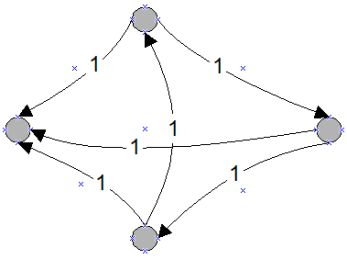


Figure 3

So, above in Fig. 2 it is determined that the given graph has three different 0-arcs. If it were a traffic intensity graph, then there should be two different u_0, u_1 values, or two different traffic layers: 0 and 1, in fact, meaning that the main traffic is possible only for the traffic shown in Fig. 3.

Another way to use the layering algorithm is more complex. An analogous algorithm can also be applied to the % processing (layering) of the data matrix. Only a few new concepts should be defined.

Based on data matrix X, we can create two frequency tables: the rows table and columns table, which will indicate the possible values of the attributes in a nominal scale. The maximum possible number **atr** of different attributes in the data matrix determines the nominal scale width or expansion.

By scanning the cells and at the same time summing the 1-s in the additional tables the two frequency tables $\boldsymbol{c}$ and $\boldsymbol{r}$ are progressively filled out. First, let's look at the corresponding cell of the κ-th object and its ℓ-th attribute in X. The $x_{\kappa,\ell}$ of this cell determines in which additional column $x_{\kappa,\ell}$ to the right of X, and in which additional row $x_{\kappa,\ell}$ at the bottom, in relation to X, the 1-s in cells of $r_{k,x_{k,\ell}}$ and 1-s in cells of $c_{x_{k,\ell},\ell}$ are summed up correspondingly. Namely, in relation to X, here $x_{\kappa,\ell}$ is the column No to the right, but also the row No at the bottom, in additional tables $\boldsymbol{r}$ and $\boldsymbol{c}$. We assume that table X (see example below) is filled with integer attributes or labels 1,2,1,3,… When filling out frequency tables, we initially look at the first object, then the next, and so on.

Table 1	´1	´2	´3	´4	´5	´6	´7	´8	´1	´2	´3	
´1	1	1	1	2	1	1	2	0	*5*	*2*	*0*	**7**
´2	1	1	1	3	1	1	3	3	*5*	*0*	*3*	**8**
´3	3	2	2	1	3	0	2	2	*1*	*4*	*2*	**7** ⇐ $\boldsymbol{\varkappa}$
´4	1	1	1	2	1	1	3	3	*5*	*1*	*2*	**8**
´5	1	1	1	0	1	1	2	1	*6*	*1*	*0*	**7**
´1	*4*	*4*	*4*	*1*	*4*	*4*	*0*	*1*				
´2	*0*	*1*	*1*	*2*	*0*	*0*	*3*	*1*				
´3	*1*	*0*	*0*	*1*	*1*	*0*	*2*	*2*				
	5	**5**	**5**	**4**	**5**	**4**	**5**	**4**				

$\boldsymbol{c} \Rightarrow$

In more compact form, the data cell (κ,ℓ) attribute determines the column No- $x_{k,\ell}$ of frequency $c_{x_{k,\ell},\ell}$ location in the table $\boldsymbol{c} = \left\| c_{t,\ell} \right\|$, $t = \overline{1,\text{atr}}$, while the cell (κ,ℓ) also determines the frequency $r_{k,x_{k,\ell}}$ location but in the row No- $x_{k,\ell}$ of table $\boldsymbol{\varkappa} = \left\| r_{\kappa,t} \right\|$; i.e. the cell (κ,ℓ) of the data matrix X, points at frequencies: $r_{k,x_{k,\ell}}$ and $c_{x_{k,\ell},\ell}$. Consider the following credentials: $\pi_{\kappa,\ell} = r_{k,x_{k,\ell}} + c_{x_{k,\ell},\ell} + \sum_{t=1}^{\text{atr}} r_{\kappa,t} + \sum_{t=1}^{\text{atr}} c_{t,\ell}$, where **atr** already has been determined as the nominal scale expansion or width.

Zero step. For all credentials $\pi_{\kappa,\ell}$ the minimum must be found and remembered using the auxiliary variable u_0. In the data matrix X the entry, where the minimum was found, — the κ-th row and ℓ-th column cell of the data table X is reset to zero or marked as processed. Thus, it usually happens that the corresponding cells to κ-th row and ℓ-th column in additional frequencies tables $\boldsymbol{c}$ and $\boldsymbol{\varkappa}$ change.

Recursive step. Thus, the reset operation may affect some of the other credentials $\pi_{\kappa,\ell}$ of the data matrix X cells, so that the credentials corresponding to those cells become less than or equal to the minor value u_k. Repeat the current step or steps for matrix X cells with this credential level u_k until no entries (cells) are found in the matrix X that satisfy the reset (zeroing) condition at the k-th step.

It is analogous to the zero steps in the graph alignment algorithm. Examples of 5×8 matrix see the Table 1 above. The credential matrix corresponding to the data matrix is as follows:

Table 2	´1	´2	´3	´4	´5	´6	´7	´8
´1	21	21	21	15	21	20	17	11
´2	22	22	22	16	22	21	18	17
´3	15	17	17	13	15	11	19	16
´4	22	22	22	15	22	21	17	16
´5	22	22	22	11	22	21	16	18

After the algorithm has been implemented against Table 2, it performs a transformation of Table 2 to Table. 3 (the reset cells are marked with the number 99):

Table 3	**´1**	**´2**	**´3**	**´4**	**´5**	**´6**	**´7**	**´8**
´1	18	18	18	99	18	18	99	99
´2	18	18	18	99	18	18	99	99
´3	99	99	99	99	99	99	99	99
´4	18	18	18	99	18	18	99	99
´5	18	18	18	99	18	18	99	99

If the result needs to be interpreted essentially, the algorithm offers the researcher, after further investigation, the following interpretation: An area exists inside the data table X or block filled with 3-s labels, which consists of rows $1,2,4,5$ and columns $1,2,3,5,6$.

A similar algorithm can be used for the following two cases. Let's choose the credentials π as a cell value of the data matrix X in the κ-th row and ℓ-th column, which will be

$$\pi_{\kappa,\ell} = \sum_{t=1}^{atr} t \cdot r_{\kappa,t} + \sum_{t=1}^{atr} t \cdot c_{t,\ell}\,.$$

These types $\pi_{\kappa,\ell}$ of indicators in mechanics are called moments. The credential consists of row moment and column moment sum. We can act exactly according to the algorithm presented earlier.

Another example. The entropy of an object κ can be calculated by formula:

$$H(\kappa) = -\frac{1}{\sum_{t=1}^{atr} r_{\kappa,t}} \sum_{t=1}^{atr} r_{k,t} \cdot \log_{atr}\left(r_{k,t} \Big/ \sum_{t=1}^{atr} r_{k,t} \right),$$

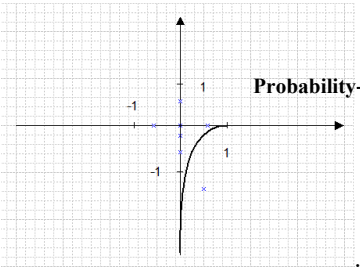
as well as similar formula $H(\ell)$ for an attribute ℓ.

The quantities $H(\kappa)$ and $H(\ell)$ are the contributions of the κ-th object and ℓ-th attribute to the total entropies $\sum_{\kappa=1}^{n} H(\kappa)$ or $\sum_{\ell=1}^{m} H(\ell)$ of the data matrix X, which according to Shannon can be expressed as the sum of the entropies of individual objects or attributes respectively.

The maximum entropy in the frequency table is reached when the distribution of distribution in the data matrix X becomes uniform. To clarify the last statement, we draw a graph of the function:

$$-\log_{atr}\left(r_{\kappa,t}\Big/\sum_{t=1}^{atr} r_{\kappa,t}\right):$$

$$-\log_{atr}\left(r_{\kappa,t}\Big/\sum_{t=1}^{atr} r_{\kappa,t}\right)$$



$$r_{\kappa,t}\Big/\sum_{t=1}^{atr} r_{\kappa,t}$$

.

The maximum entropy of the data matrix in the row direction is computed when the probabilities on the x-axis allocate a uniform frequency distribution, resulting in $\max\limits_{\kappa} H(\kappa) \approx 1$. Indeed, the value $-\log_{atr}\left(r_{\kappa,t}\Big/\sum_{t=1}^{atr} r_{\kappa,t}\right)$ is at its maximum when $r_{\kappa,t}\Big/\sum_{t=1}^{atr} r_{\kappa,t} \approx 1/atr$. In case $r_{k,t} = 0$ then this zero value is not taken into account. Based on the maximum entropy, we get the actual information about the object κ equal to $1 - H(\kappa)$. Thus, the complete information contained in the data matrix X is calculated by the formula: $n - \sum_{\kappa=1}^{n} H(\kappa)$. The above layering algorithm can now be used.

For the credential of an individual object, we choose the entropy value $H(\kappa)$. Thus, the set of objects $x_1, x_2, ..., x_n$ is to be stratified. It is only necessary to keep in mind that after removing an object from the data matrix, changes occur in the frequency table (frequency bands). The changes consist in the fact that when using the values of the ℓ-th attribute $x_{\kappa,\ell}$ of the κ-th object, in the corresponding cells $r_{\kappa,x_{\kappa,\ell}}$ and $c_{x_{\kappa,\ell},\ell}$ of the frequency tables r and c, 1 is subtracted from the frequencies: $r_{\kappa,x_{\kappa,\ell}} = r_{k,x_{\kappa,\ell}} - 1$ and $c_{x_{\kappa,\ell},\ell} = c_{x_{\kappa,\ell},\ell} - 1$.

We will consider the properties of the stratification algorithm using the mentioned monotone systems in the next section, where the positive $\oplus$ and negative effects of elements are used. In graphs, the negative $\ominus$ effect on the arc was its removal. For data matrix, this is the reset of the ℓ-th attribute of the κ-th object or a series of $\ominus$ effects until the object will be completely removed by the entropy level u_k assessment.

5. MONOTONE SYSTEMS

We will continue our story about monotone systems now at a more precise level according to the publication (Mullat, 1976-1977) in Autom. and Telemechanics. A monotonous system manifests itself in the relationship between elements in the fact that if an element of the system is "positively influenced", then this effect is also positively reflected on its interrelated elements. It's the same with negative effects.

The monotonicity property as a central property allows us to formulate the concept of the system kernel or core in a general form. By the core, we mean a subset of the elements of "strongly attracting" or "strongly pushing" each other the elements of the system.

Consider any system W consisting of a finite set of elements, i.e., $|W| = n$ |. Quantities or credentials that indicate the level of "importance" of the element $\alpha \in W$ for the functioning of the system as a whole characterize the states of the elements of a system W.

It proves necessary to reflect the internal dependence of the elements of the system at the level of importance of the elements. In view of the fact that the elements of the system are interconnected, it is possible to take into account the effect of element α on other elements related to the change in the properties of element β. We assume that the level of importance of the element α itself also changes due to its effect. If elements α and β are in no way related in the system, it is natural to assume that the change caused by element α to the importance of element β is zero.

In the system W, we consider as an effect on the element α of two types of effects: $\oplus$ and $\ominus$ type effects ($\oplus$- and $\ominus$-effects). In the first case, the properties of element α are considered to improve as its importance to the system increases; in the second case, the properties of element α deteriorate as its level of importance in relation to the system decreases.

Now we can also provide a definition of a monotonic system. A monotonic system is a system in which the positive effect of $\oplus$ on any system element α causes the positive effect of $\oplus$ on all other elements of the system and the effect of the $\ominus$ type causes the effect of $\ominus$ type respectively.

5.1. System monotonicity conditions

The observed important concept — the effect on the element α of the system W and the accompanying effect on the other elements of the system — allows the set W to determine an infinite number of functions, since we have at least one actual function of the importance of the elements W of the system: $\pi : W \to \Re$, where $\Re$ is a set of real numbers.

If element α is affected, then it can be said that the function π is reflected in the function π_{α}^{+} for the effect of $\oplus$ and in the function π_{α}^{-} for the effect of $\ominus$ respectively. As a result of the effects $\oplus$ and $\ominus$ on the element implementation, the credentials of the system elements are redistributed from the function π to the functions $\pi_{\alpha}^{+}\pi_{\alpha}^{-}$ or the initial set of values $\{\pi \mid \pi(\delta \in W)\}$ is transferred to a new set $\{\pi \mid \pi_{\alpha}^{+}(\delta \in W)\}$ and $\{\pi \mid \pi_{\alpha}^{-}(\delta \in W)\}$ respectively. The functions π, π_{α}^{+}, π_{α}^{-} are defined on the whole set W and thus are also defined $\pi_{\alpha}^{+}(\alpha)$ and $\pi_{\alpha}^{-}(\alpha)$. It is clear that if there is given a sequence α_1, $\alpha_2, \alpha_3 \ldots$from the W set of elements (all repetitions and combinations of elements are allowed), and e.g. the a binary sequence $\oplus$, $\ominus$, $\oplus$,… then can be easily determined the combined effect in the form of a functional product of $\pi_{\alpha_1}^{+} \cdot \pi_{\alpha_2}^{-} \cdot \pi_{\alpha_3}^{+} \cdot \ldots$

The presented construction allows writing the monotonicity property of the systems as two main inequalities:

$$\pi_{\alpha}^{+}(\beta) \geq \pi(\beta) \geq \pi_{\alpha}^{-}(\beta)$$

for each element pair $\alpha, \beta \in W$, including pairs (α, α) and (β, β).

5.2. Identification of the system kernel

To determine the kernel of the system, consider the two subsets of W, namely H and $\overline{H}$, so that $H \cup \overline{H} = W$ and $H \cap \overline{H} = \varnothing$.

If only elements α_1, $\alpha_2, \ldots, \in H$ are positively affected then it determines for the set W a certain function $\pi_{\alpha_1}^{+} \cdot \pi_{\alpha_2}^{+} \cdot \ldots$, which can be considered determined only for the subset H. If we choose one of all possible sequences of a set H, namely $\left\langle \alpha_1, \alpha_2, \ldots, \alpha_{|\overline{H}|} \right\rangle$ where α_i does not repeat, then the function $\pi_{\alpha_1}^{+} \cdot \pi_{\alpha_2}^{+} \cdot, \ldots, \pi_{\alpha_{|\overline{H}|}}^{+}$ is denoted unambiguously on the set H function and call it a standard function. The function thus introduced is called the credential function on the set H and the individual value of the function on the element α is the credential. These credentials $\{\pi^{+}H(\alpha) \mid \alpha \in H\}$ we denote by $\Pi^{+}H$ and call this set of credentials specified for a given set H, i.e., for the set of credentials with respect to the set H.

Suppose that the set of credentials sets $\{\Pi^{+}H \mid H \subseteq W\}$ for all possible subsystems 2^{W} of system W — the number of all possible subsystems is $2^{|W|}$.

Instead of the plus effects of the standard function, we can look at the analogous $\ominus$ effects function $\pi^-_{\alpha_1}\cdot\pi^-_{\alpha_2}\cdot,...,\pi^-_{\alpha_{|\overline{H}|}}$. Similarly to the function $\pi^+H(\alpha)$, we also determine, the set of credentials $\{\pi^-H(\alpha)\,|\,\alpha\in H\}$ and also the collections of sets of credentials $\{\Pi^-H\,|\,H\subseteq W\}$. In addition, to obtain a process of type $\ominus$ effects — an analogous process Π^-H is performed. All elements of the set H are affected in sequence according to the ordered list $\left\langle\alpha_1,\alpha_2,...,\alpha_{|\overline{H}|}\right\rangle$.

On the subsets or arrays $\{\Pi^+H\,|\,H\subseteq W\}$ and $\{\Pi^-H\,|\,H\subseteq W\}$ of credentials given on the sets $H\subseteq W$, the following two functions can be defined for each subset H:

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

By the kernels of W we call the global minimum of the function $F_+(H)$ and the global maximum of the function $F_-(H)$. The subsystem $H^{\oplus}$ that reaches the global minimum of the F_+ function is called the system $\oplus$-kernel, and the subsystem $H^{\ominus}$ that reaches the global maximum of the F_- function is called the $\ominus$-kernel, respectively.

Definition. The defining set considered in monotone systems theory is the last set in the layer algorithm with level u_p (see the section 3 above), where the sequence $\overline{\alpha}=\left\langle\alpha_1,\alpha_2,...,\alpha_{|\overline{H}|}\right\rangle$ of system elements by which such a defining set is found is called the defining sequence.

Theorem 1. The defining set $H^{\ominus}$ is the set where the F_- function reaches the global maximum. There is only one defining set $H^{\ominus}$ set. All other subsets if they exist where F_- reach the global maximum are within the defining set $H^{\ominus}$.

Theorem 2. For the definite set of $H^{\oplus}$, the function F_+ reaches a global minimum. There is only one defining set $H^{\oplus}$. All sets that reach the global minimum are enclosed in the defining set $H^{\oplus}$.

The existence of defining sets $H^{\ominus}$ and $H^{\oplus}$ is ensured by a special constructive routine. The defining sets are kernels of Monotone Systems, because on these sets the functions F_{-} and F_{+} reach the global maximum (minimum) accordingly. Theorems 1 and 2 guarantee that all kernels are located in one "large" kernel — the defining set.

6. MONOTONE SUBSYSTEMS ON GRAPHS

Let us have a "*big*" graph G and a "*small*" graph g. It is necessary to select a part of the "*big*" graph G (a set of arcs or edges) so that this set is the most *"saturated"* with "*small*" graphs g. For example, we can assume that one part of the graph is more saturated than the other if the first contains more small graphs g than the second.

With some complexity, saturation can also be approached as follows. Consider the arcs, edges or vertices of G that belong to the part we are interested in. We now count in integers: how many there are small graphs g, separately those g graphs that are located "near" each vertex, arc or edge. By this integers is meant the number of graphs g that contain a given vertex, arc or edge, and are thus expressed as an integer. By doing this, we get exactly such an integer or credential that characterizes the part of G we is interested in. Each such integer reflects a certain "*local*" saturation of the graph G with the graphs g.

Based on the obtained integers, several variants open to determine the saturation of the G part of the graph. The mean, variance, etc., of these numbers can be calculated. We consider the simplest credential magnitude, namely the entity of small graphs g, which are located in a separate part of a large graph G, i.e., the smallest value of the local parts. Figuratively speaking, this number of sub-graphs is in the most "empty" location of the graph G, which we should further on remove by $\ominus$ type actions.

Below we give an exact representation of the problem of determining the most saturated parts of the graph G with small graphs. We set the problem as follows: From all possible parts (or a large number of parts) of a graph G we find the one with the maximum value of the smallest number of local sets of small graphs g.

It is natural that in this method many small graphs g can be placed in a part in the usual way, because the number of small sub-graphs g on each vertex or arc is not less than on the vertex or arc on which it is minimal. At the same time, however, this minimum number in the extreme part is quite large, because we specifically chose the part where the local number of graphs condition reaching the global maximum of the minimum would be satisfied,

Similarly, we can set the task of finding the part of the graph G that is least saturated with small graphs g. The number of sub-graphs g at the vertex or arc where this number is maximal characterizes then each part of the graph. Instead of looking for the part of the graph where the minimum local number of graphs is the maximum, we look for the part where the maximum local number is the minimum. In this case, the number g of the sub-graphs of each vertex or arc is not greater than the "maximum" vertex or arc, and the latter has a default due to the global minimum condition.

The extreme parts of a graph are usually uniformly saturated or unsaturated with small graphs. In a saturated extreme part, no single vertex or arc can usually have very few graphs g, because without the arc of this vertex the part of the graph is probably more saturated at the top or arc with sub-graphs g in the more complex sense mentioned above.

7. GENERAL MODEL OF KERNEL EXTRACTION ON GRAPHS

If a graph G is given, then with $V(G)$ or by V we denote the set of vertices of the graph. We denote the set of arcs of an oriented graph G by $U(G)$ or U and the set of edges of an unoriented graph by $E(G)$ or E.

In graph theory, the concept of a sub-graph of a given graph G is used. A graph G' is a sub-graph of the graph $[V(G), U(G)]$ if $V(G') \subset V(G)$ and $U(G')$ is the set of arcs of all and only those that bind the pair from $V(G')$. Similarly, we can define a sub-graph of an undirected graph if the term edge is used instead of the arc.

Sometimes the term part G of a graph is also used. We call graph G a part of the graph $G[V, U]$ if $V(G'') \subseteq V(G')$ and $U(G'') \subseteq U(G')$. In terms of the oriented graph, some arcs of the graph G are simply missing. Similarly, an undirected sub-graph is determined.

The design of concepts described in the previous two sections of this guide must begin with the identification of the elements of the system W. Two structural units can be separated from graphs — a vertex and an arc. Let us consider first the case where the vertex of the graph G is chosen as an element of the system. We now determine the effects of the $\oplus$- and $\ominus$-effects on the vertices, i.e., on the elements of the system W. Determining the effects of $\oplus$ and $\ominus$ requires the addition of a special significance function π to the vertices of the graph G. The action has already been mentioned in the previous two sections of the guide, that the credentials in the system must increase as a result of the $\oplus$ effect and decrease as a result of the $\ominus$ effects.

We need to define saturation indicators, or whatever we call them, credentials for the elements α of each subset of H from W. To get this, we need to set up an initial set of credentials for W, as well as a framework how to express $\oplus$ and $\ominus$ effects.

An initial set of credentials $\{\pi(\alpha) \mid \alpha \in W\}$ can be specified, for example, as follows. Let g be a small graph given a large graph G. We count the number of different sub-graphs of graph G that are isomorphic to graph g and whose vertices include vertex α. We set the just obtained number to the initial credential level $\pi(\alpha)$. To underline the introduced dependence of the level $\pi(\alpha)$ on the small graph g, we use the expression — the credential of the vertex α of the graph G with respect to g. Next, we consider two operations for obtaining new graphs from G, namely the $\oplus$ and $\ominus$ operations.

Let a graph G be given and an empty graph Λ (a graph that has no arcs but has $|V(G)|$ vertices). We assume that $V(\Lambda)$ is an exact copy of $V(G)$. And when we talk about the vertex α, we mean the vertex of a graph G, which appears in two forms — like the vertex of a graph G and like the vertex of a graph Λ.

A $\ominus$-type operation of a graph G with a vertex α is to carry out removing all the arcs or edges leading to that vertex. On an empty graph Λ, however, the $\oplus$-type operation is a recovery operation for all edges leading to that vertex α. It appears that if a $\oplus$-type operation is applied to a vertex, the credentials of all other vertices (relative to the small graph g) either decrease or, in some cases, remain the same. When performing a $\oplus$-type operation, a natural question arises: what should be considered the credential of the vertex after restoring the vertex?

The solution to this question lies in the following construction. Let us count the credentials of the vertices of the graph Λ (with respect to the small graph g) and add the credentials of the vertices of the graph G. We consider the obtained amounts as the total credentials of the vertices. In this case, the opposite effect can be observed: as a result of the $\oplus$-type operation, the total credentials increase or, like the $\ominus$-type credentials, remain at the same level. Generally speaking, the initial credential set $\{\pi(\alpha) \mid \alpha \in W\}$ (the credential set before any $\oplus$-type operation) of the vertices of graph G can be considered as a general credential set to be built since any part of graph G is initially empty. At this stage, minimizing the maximum credentials means some options for the vertices of graph G to be isolated. In this approach, the monotonicity condition is satisfied.

When constructing sets of credentials in system W, it must be demonstrated how the initial set of credentials $\{\pi(\alpha) \mid \alpha \in W\}$ found is redistributed due to $\oplus$ and $\ominus$ operations.

Let be given a certain sequence of vertices $\overline{\alpha} = \langle \alpha_1, \alpha_2, ... \rangle$, which forms a set of $\overline{H} \subseteq W$. We express the effect of $\oplus$ on the vertices of G according to their occurrence in the sequence. As a result, a sub-graph of G is formed on the graph $V(\Lambda)$. At the vertex of each resulting sub-graph we can count the number of isomorphic sub-graphs with a small graph g, so we get the credentials of a set of H (the complement of $\overline{H}$ to W) elements. Consistent with the above theory, we can state that the set H determines a new significance function in the form,

$$\pi^{+}_{\alpha_1} \cdot \pi^{+}_{\alpha_2} \cdot ... \tag{2}$$

obtained from the initial credential collection $\{\pi(\alpha) \mid \alpha \in W\}$.

Thus, if a sequence of vertices $\overline{\alpha} = \langle \alpha_1, \alpha_2, ... \rangle$ is given that promotes the set $\overline{H}$, then the set H forms a set of credentials determined by (2) or (3). We denote this set by $\Pi^{+}H$, and we call the set of credentials by the set of vertices induced on H. The sets of induced credentials form the set $\{\Pi^{+}H \mid H \subseteq W\}$. Sometimes it is appropriate to use the expression of $\oplus$-collection of sets with respect to the small graph g.

The collection or array $\{\Pi^{-}H \mid H \subseteq W\}$ of sets of credentials is determined analogously. The collection $\Pi^{-}H$ of the credentials is determined by the function

$$\pi^{-}_{\alpha_1} \cdot \pi^{-}_{\alpha_2} \cdot ... \tag{3}$$

given in part G of the graph, which remains after the application of the $\ominus$-activities to the sequence of vertices forming $\overline{\alpha} = \langle \alpha_1, \alpha_2, ... \rangle$. It only needs to be emphasized that each subset $H \subseteq W$ of the set of credentials is in fact the set of the remaining part, but not the total, i.e., not the part given by the set of graph Λ, which actually is an empty graph.

Next, let's take the arc as the system element. The system is defined as the set of interconnected arcs $U(G)$ of the graph G, determining the $\oplus$ and $\ominus$ effects again requires setting the values of the initial function π.

Let be given a small graph of g. We count the number of different subgraphs of the graph G that are isomorphic to the graph g and whose arcs or edges include this arc or edge. The resulting integer is taken as the significance level of the arc α of the graph G. This is called the credential of the arc α with respect to the graph g.

Similarly to those described at the vertices of G, the concepts of $\oplus$ and $\ominus$ activities are also determined by the arcs or edges of the graph G. Arcs or edges are now removed or restored instead of vertices.

Let's look at the $\ominus$ operation first. It is obvious that as a result of removing the arc (edge), the initial set of credentials with respect to the small graph g may decrease or remain the same. A decrease in importance of credentials indicates that the $\ominus$ operation is equivalent to defining $\ominus$ activity for system elements.

Let $\langle \alpha_1, \alpha_2, ... \rangle$ be a sequence of different arcs on G, including arcs forming $\overline{H} \subseteq U(G)$. We perform $\ominus$-actions sequentially on the arcs of the graph G according to the given sequence. As a result, we get a certain part of the graph G, the elements of which are arcs (edges) belonging to the set $H \subseteq U(G)$. For each arc $\alpha \in H$, count the number of isomorphic graphs with the graph g, which is considered to be the credential or significance of the element α with respect to the set H.

According to the notations used, the method for determining the given credentials creates a function on the elements of the set H of arcs. Similarly to the case where the number of sets of credentials was assigned to the vertices of a given graph, arcs (edges) are created that belong to the set of credentials $\{\pi^- H(\alpha) \mid \alpha \in H\}$, which we denote again $\Pi^- H$. We proceed in a similar way to find the set of credentials $\{\Pi^- H \mid H \subseteq U(G)\}$. On an empty graph Λ, defining the $\oplus$-activity on the basis of the $\oplus$-operation requires a more detailed analysis.

Let again the sequence of arcs $\overline{\alpha} = \langle \alpha_1, \alpha_2, ... \rangle$ in the given graph G (said arcs form the set $\overline{H}$), we perform $\oplus$-operations on the set $\overline{H}$ arcs sequentially. As a result, the set of vertices $V(\Lambda)$ forms a part of a graph G whose list of arcs is equal to $\overline{H}$. For the vertex model, we calculated the total credential of each vertex $\alpha \in V(G)$. In this case, too, we try to do the same and find the total credential of the arcs forming H.

The arcs belonging to the set H are not present in the graph g and the question is how to count the number of sub-graphs isomorphic to the graph g and containing the arc α (which is not present in the graph Λ). Proceed as follows: we read that this arc α is fictitious only at the moment of counting the sub-graphs. In this case, the set of arcs H forms certain integers that depend on both the graph and the part of the graph formed on the empty graph g.

In the method described above, the function $\pi^{+}_{\alpha_1} \cdot \pi^{+}_{\alpha_2} \cdot \ldots$ is determined from the quantity H, which creates a set of $\oplus$-credentials $\left\{\pi^{+}H(\alpha) \,\middle|\, \alpha \in H\right\}$.

In this case, even in the case of a $\oplus$-operation, the set of credentials of the $\oplus$-activities can be determined with respect to a small graph. The use of the term "$\oplus$-activity" is perfectly legal here, as the total credentials of those elements that are not yet subject to $\oplus$-activity may increase or remain the same.

7.1. Illustrative Examples on Directed Graphs

A graph G of partial ordering is defined as a binary relation G with the following properties:

a) Reflexivity, i.e., if $\alpha \in V(G)$, then $\alpha G \alpha$. The graph G has a loop at the vertex α.

b) Transitivity, if there exists an arc (α,β) and (β,γ), then the graph G has an arc (α,γ), or from $\alpha\, G\, \beta$ and $\beta\, G\, \gamma$ it follows that $\alpha\, G\, \gamma$.

A complete order is defined as a graph of partial ordering in which any pair of vertices α and β is connected by an arc.

It is possible to formulate the following problem: in a given directed graph it is required to find the (in certain sense) most “saturated” regions that are “close” to a graph of partial ordering or to graphs of complete ordering. This problem will be solved by a method of organization (on a graph) of a monotonic system with subsequent determination of kernels.

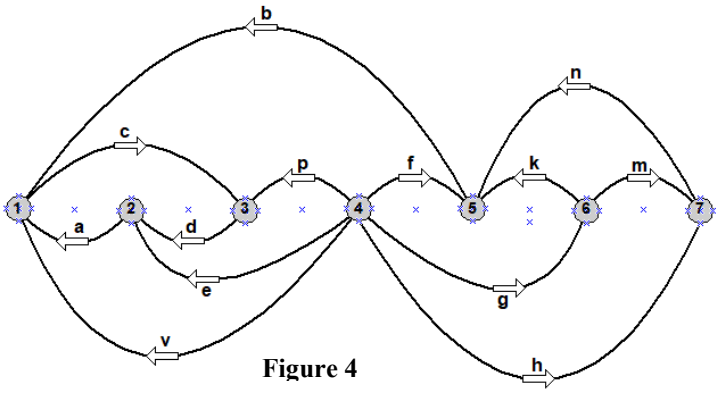


Figure 4

In accordance with the scheme of organization of a monotonic system on graphs described in the previous section, it is necessary to assign a small graph g. Suppose that this graph consists of three vertices x, y, z, and it is such that $U(\Gamma) = \{(x,y),(y,z),(x,z)\}$. The graph has a total of three arcs (a transitive triple).

Now let us consider the assignment of collection of credentials arrays at the vertices of a graph shown in Fig. 4. The loops on this graph have been omitted.

According to the scheme of assignment of collections of credential arrays at the vertices of a graph, it is required to determine an initial array of credentials $\{\pi(\alpha)\}$, where $\alpha = 1,2,3,...,7$. According to the method of calculation of the values $\pi(\alpha)$ with respect to the graph g (a transitive triple), we obtain $\pi(1) = 3$, $\pi(2) = 2$, $\pi(3) = 2$, $\pi(4) = 7$, $\pi(5) = 4$, $\pi(6) = 3$, $\pi(7) = 3$. As an example, let us determine a credential array on a subset of vertices $H = \{1,2,3,4,5\}$. By successively performing $\ominus$ actions on the set $\overline{H} = \{6,7\}$, we obtain on the set H a new credential array $\pi(1) = 3$, $\pi(2) = 2$, $\pi(3) = 2$, $\pi(4) = 4$, $\pi(4) = 4,$, $\pi(5) = 1$.

The values of the function $\pi_6^+\pi_7^+$ can be obtained in a similar way, but for this purpose it is necessary to use the assignment of collections of total $\oplus$ arrays with respect to a transitive triple. According to Fig. 5, the values of this function in their order at the vertices $\{1,2,3,4,5\}$ are as follows: $\pi(1) = 3,$, $\pi(2) = 2$, $\pi(3) = 2$, $\pi(4) = 8$, $\pi(5) = 4$. In exactly the same way we can determine on any subset H of vertices $V = \{1,2,3,4,5,6,7\}$ a proper credential array of $\oplus$ or $\square$ actions with respect to a transitive triple.

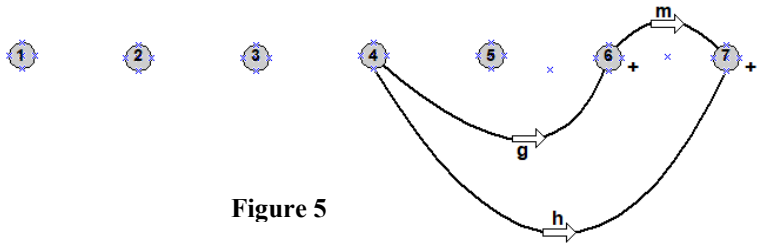


Figure 5

Now let us consider a construction that is assigned not on vertices, but on the arcs of the graph presented on Fig. 4. In this case the set of elements of the system W will be $U(G) = \{a,b,c,...,n,m\}$. As the small graph g we shall take the same graph as above, with a set $U(g) = \{(x,y),(y,z),(x,z)\}$.

By analogy with the foregoing, we realize the construction in the same succession. We determine an initial credential array $\{\pi(\alpha) \mid \alpha \in U\}$ on the arcs of the graph G in accordance with the general scheme.

We find that

$$\pi(a)=1, \pi(b)=1, \pi(c)=1, \pi(d)=1, \pi(e)=2, \pi(f)=3,$$
$$\pi(g)=2, \pi(h)=2, \pi(k)=2, \pi(n)=2, \quad \pi(m)=1, \pi(v)=3, \pi(p)=2$$

As an example, let us now perform ($\oplus$ and $\ominus$ actions on the arcs f, k and m, i.e., on the set $H=\{f,k,m\}$. On the set H we hence obtain

$$\pi(a)=1, \pi(b)=0, \pi(c)=1, \pi(d)=1, \pi(e)=2,$$
$$\pi(g)=0, \pi(h)=0, \pi(n)=0, \pi(v)=2, \pi(p)=2.$$

In accordance with the adopted system of notations this array of numbers will be denoted by $\Pi^{-}H$. For obtaining an $\Pi^{+}H$ array, we must calculate the total credentials. The dashed lines in Fig. 6 represent the arcs of graph Λ that experience the effect of □ actions performed on the arcs f, k and m.

According to Fig. 6, the total credential array will be as follows:

$$\pi(a)=1, \pi(b)=1, \pi(c)=1, \pi(d)=1, \pi(e)=2,$$
$$\pi(g)=3, \pi(h)=2, \pi(n)=3, \pi(v)=2, \pi(p)=2.$$

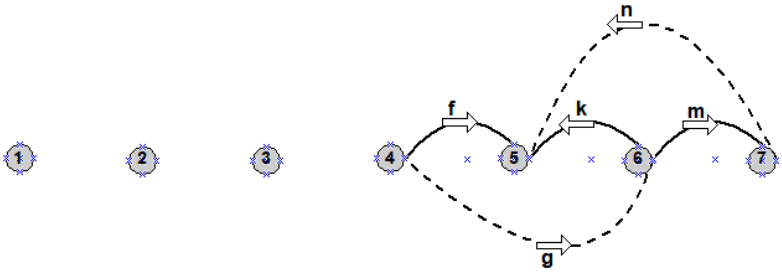


Figure 6

Thus on any subset H of arcs of the graph shown in Fig. 4 we can construct the credential arrays $\Pi^{-}H$ and $\Pi^{+}H$.

Next we describe the procedures of construction of determining sequences of $\oplus$ or $\ominus$ actions, at first for vertices, and then for arcs of the graph shown in Fig. 4. The construction is carried out for the purpose of illustrating the concepts of $\oplus$ or kernels of the monotonic system and also for ascertaining the effect of the duality theorem formulated by Mullat (1976-1977).

Let us consider an example in which $\ominus$ credential arrays are assigned at vertices with respect to a transitive triple. According to the scheme prescribed in Mullat's routine of construction of a determining $\oplus$ and $\ominus$ sequence of vertices of a graph on the basis of $\oplus$ and $\ominus$ actions. For the graph shown in Fig. 4, the Kernel-Searching Routine consists of two steps: the zero-th and the step one. It yields two subsets $\Gamma_0^-, \Gamma_1^- \subseteq V(G)$, where

$$\Gamma_0^- = V(G) = \{1,2,3,...,7\},\ \Gamma_1^- = \{4,5,6,7\},$$
and the thresholds $u_0 = 2$, $u_1 = 3$.

The determining sequence of vertices constructed with the aid of $\ominus$ actions is as follows: $\overline{\alpha}_- = \langle 3,2,1,4,5,6,7 \rangle$. Thus on the basis of: a) according to Theorems 1,3 (Mullat, 1971) and b) according to Theorem 1 (Mullat, 1976) about KSR, it can be argued that the set $\{4,5,6,7\}$ is the definable set of vertices of the graph shown in Fig. 4, and, therefore, this set is also the largest kernel $K^{\ominus}$.

Now let apply the KSR for constructing a $\oplus$-determining sequence. We find that $\overline{\alpha}_+ = \{4,5,6,7,1,2,3\}$. The routine terminates at the third step, and it consists of four steps, namely the zero-th, the first, the second and the third. According to the construction of $\oplus$ sequences prescribed in the KSR, we produce the sets Γ_j^+: $\Gamma_0^+ = \{4,5,6,7,1,2,3\}$, $\Gamma_1^+ = \{5,6,7,1,2,3\}$, $\Gamma_2^+ = \{6,7,1,2,3\}$, $\Gamma_3^+ = \{2,3\}$ and a sequence of thresholds $u_0 = 7$, $u_1 = 4$, $u_3 = 3$, $u_3 = 2$. As in the case of a $\oplus$ sequence, we conclude on the basis of Theorems 2 and 3 of a) Mullat, and of Theorem 1 of b) Mullat, that $\{2,3\}$ is the largest $K^{\oplus}$ kernel of the system of vertices of the graph in Fig.1.

A careful analysis of Fig.1 shows that the $K^{\oplus}$ kernel is in fact completely ordered set, i.e., $\langle 4,5,6,7 \rangle$. On the other hand the $K^{\oplus}$ indicates from the point of view of the "structure" of a graph that the region, in which the vertices are least ordered, it is ordered itself as well. This is in agreement with the our formulation of the problem of finding kernels as representatives of "saturated" or "unsaturated" regions (parts of a graph) with small graphs g

Now let us use the KSR for constructing determining sequences of arcs of the graph in Fig.1. The graph has a total of 13 arcs. After applying the KSR, we obtain on the basis of $\ominus$ actions the following sequence:

$$\overline{\alpha}_- = \langle a,b,c,d,v,e,p,f,k,n,m,h,g \rangle.$$

The routine terminates at first step and it consists of two steps, namely the zero-th step and the first step. At the zero-th step we have $\Gamma_0^- = U(G)$, and at the first step we have $\Gamma_1^- = \{f,k,n,m,h,g\}$, with the thresholds $u_0 = 1$ and $u_1 = 2$ respectively. Summing up, we can assert on the basis of the results of a), b) Mullat, that this is a definable set and at the same time the largest $K^{\ominus}$ kernel in the system of arcs.

From the point of view of the graph structure, the application of the KSR to arcs in the construction of a $\ominus$ determining sequence does not yield anything new compared to the application of the KSR to vertices. We obtain the same complete order $\langle 4,5,6,7 \rangle$ represented in the form of a string of arcs, and it also corroborates our assertions concerning the saturation of a $K^{\ominus}$ kernel by transitive triples. On the other hand the use of KSR for constructing $\oplus$ determining sequence of arcs yields a $K^{\oplus}$ kernel

$$\Gamma_1^+ = \{k,m,n,g,h,e,p,b,a,c,d\},$$

whose meaning with regard to "non-saturation" with transitive triples cannot be determined.

Below we shall illustrate the peculiar features of using the duality theorem from b) Mullat (1976) for finding $K^{\oplus}$ and $K^{\ominus}$ kernels of a monotonic system specified by vertices or arcs of a directed graph.

At first let us consider the monotonic system of vertices of the graph in Fig.1. The sequence of sets $\langle \Gamma_j^+ \rangle$ specified by the KSR on the basis of $\oplus$ actions uniquely determines the sets $V \setminus \Gamma_1^+ = \{4\}$, $V \setminus \Gamma_2^+ = \{4,5\}$, $V \setminus \Gamma_3^+ = \{1,4,5,6,7\}$. Above we have found that $F_+(\Gamma_2^+) = u_2 = 3$. From the construction of a determining sequence $\overline{\alpha}_-$ of vertices of a graph we know that $F_-\{4,5,6,7\} = 3$. Hence by virtue of Corollary 1 of Theorem 1 of b) Mullat, we can assert already after the second step of construction of a $\overline{\alpha}_-$ sequence that the set $\{1,4,5,6,7\}$ contains the largest $K^{\ominus}$ kernel. Thus we have shown that the sufficient conditions of the duality theorem of b) Mullat, are satisfied in the example of the graph represented in Fig. 1.

Now let us consider the set $V \setminus \Gamma_1^- = \{1,2,3\}$. As was shown above, inside this set there exists a set $\Gamma_3^+ = \{2,3\}$ such that $F_+(\Gamma_3^+) = 2$; $F_-(\Gamma_1^-) = 3$ on the other hand. By virtue of Corollary 4 of the duality theorem we can assert that set $\{1,2,3\}$ contains the largest $K^{\oplus}$ kernel of the system of vertices of the graph (Fig.1); this likewise confirms that existence of the conditions governing the theorem.

At last let us consider a collection of credential arrays on the arcs of the graph. The determining $\overline{\alpha}_+$ sequence of arcs specifies a set $\Gamma_1^+ = \{k,m,n,g,h,e,p,b,a,c,d\}$. It is easy to see that inside the set $U \setminus \Gamma_1^+$ there does not exist a set H as required by the conditions of Corollaries 1 and 2 of the duality theorem in Mullat (1976). This shows that in comparison to arrays on vertices, credential arrays on arcs do not satisfy the duality theorem.

7.2. Monotonic systems on special classes of graphs

In contrast to the previous section, we do not carry out here a detailed construction of collections of credential arrays and determining sequences and kernels on any illustrative example. Here we shall show how to select a small graph g and $\oplus$ and $\ominus$ actions so as to match the selection of these elements with the desired "saturation" of the investigated graph. The desired saturation of a graph can be understood as the saturation desirable for the investigator who usually has a working hypothesis with respect to the graph structure. In view of this, we shall consider the following classes of graphs: tournaments, a-cyclic (directed) graphs, and (directed or undirected) trees.

Let us recall the definitions of these classes of graphs. A tournament is a directed graph in which each pair of vertices (x,y) is connected by an arc, cf. Harari (1969). A none-cyclic graph is a graph without cycles (in case of an undirected graph), and a graph without circuits (in case of a directed graph). None-cyclic undirected graphs are trees, and we shall consider the most general class of trees, as well as the class of directed trees.

In tournaments it is appropriate to consider regions of vertices that are "saturated" with cyclic triples. A cyclic triple is a graph g such that $V(g) = \{x,y,z\}$, $U(g) = \{(x,y),(y,z),(x,z)\}$. It can be assumed that a tournament in which there exists such a region represents a structure of the participants of the tournament. This structure is non-uniform; i.e., there exists a central region (set) of participants who can win against the other players, but they are in neutral position with respect to one another.

For solving the above problem, we propose the following exact formulation in the language of monotonic systems. In Section 2 we have considered credential arrays on vertices and arcs of a graph. Now let us consider the above models on vertices or arcs in a certain order. In both models we take a cyclic triple as the small graph g with respect to which the π function is calculated. Suppose that the methods of assignment of collections of credential arrays on vertices are the same as in Section 2. It is possible to modify this scheme by taking as a $\ominus$-action on the vertex α the removal of all arcs of a tournament that originates at α, whereas $\oplus$-action is the restoration of all the arcs in the graph Λ that originate at α. In Section 2 we performed the opposite operation, i.e., the removal of incoming arcs and the restoration of these same incoming arcs.

The assignment of credential arrays on arcs of a tournament graph must be carried out in accordance with a scheme similar to that described in Section 2. Within the framework of the theory it is apparently impossible to decide whether the scheme of determination of kernels on arcs of a tournament is preferable to the scheme using vertices; therefore, it is necessary to carry out computer experiments. There exists only one heuristic consideration. If in a tournament there can exist several central regions saturated with cyclic triples, it will be preferable to use the scheme of determination of kernels on the arcs of tournament, since these regions can be found. The model based on vertices makes it possible to find a kernel that consists also of regions, but it does not permit finding an individual region. We do not possess a string of arcs representing these regions.

None-cyclic directed graphs are a convenient language for describing operation systems (Kendal, 1940). An operation system can be regarded as a system of modules and interpreted as a library of programs. Each working program is a path in a none-cyclic graph, or, in other words, the set of modules of a library needed at a given instant. The modules are called one after another if not all of them can be stored in the main memory. In case of a library of a large size, there naturally arises the question of fixing the modules on information carriers. Prior to solving this problem, it is appropriate to ascertain the "structure" of a none-cyclic graph of a library of modules.

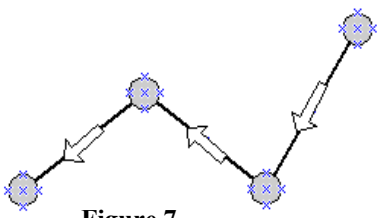


Figure 7

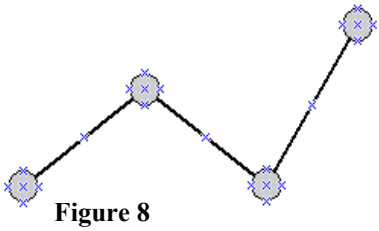


Figure 8

For ascertaining the structure of a graph and for just-mentioned task of fixing the modules, we have to find the principal (nodal) vertices or arcs. The nodes are the "bottlenecks" of graphs or, in other words, the modules that occur in many working programs.

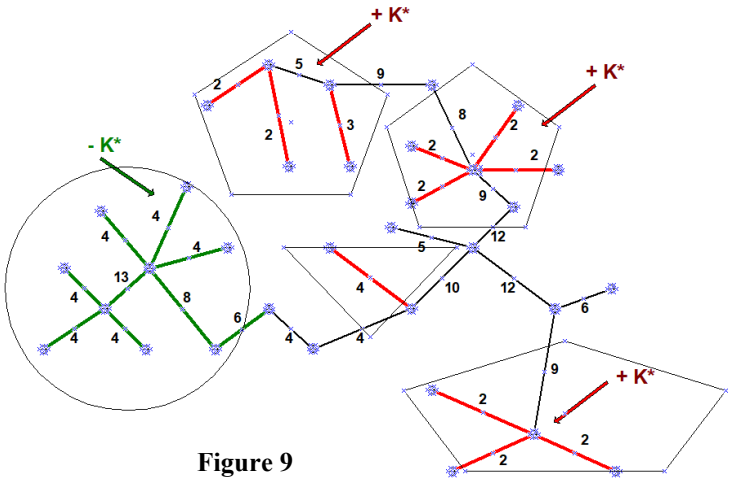


Figure 9

We shall now formally describe this problem with the aid of a model of organization of a monotonic system on a graph. As a small graph we shall take directed graph in Fig.7. The structure of this graph is in accordance with the above definition of bottlenecks of the none-cyclic graph under consideration. It is possible to construct a monotonic system also on the arcs of a none-cyclic graph of a library of modules. With the respect to the graph on Fig.7, the collection of credential arrays and $\oplus$ and $\ominus$ actions, in accordance with the general scheme of Section 2, must be defined. After this it is necessary to use the routine of finding vertex kernels or arc kernels, which in conjunction must indicate the bottlenecks in accordance with the above definition. As in case of tournaments, which a monotonic system is preferable of arcs or vertices requires experimental checking.

In comparison to the two previous examples, the last example does not have the aim of associating the application or description of any actual problem with tees. Our aim is to try and find in a tree a region, which in some sense is more similar to "cluster" than any other part of the tree.

At first let us consider undirected trees. We shall use a model of organization of a monotonic system on the branches of a tree. As a small graph g we shall take the graph shown on Fig. 8. As in the case of assignment of collections of $\oplus$ and $\ominus$ credential arrays on arcs, we assign the corresponding $\oplus$ and $\ominus$ arrays with respect to the graph shown in Fig.9. The $\ominus$ arrays appear as a result of $\ominus$ actions (removal of edges), whereas the $\oplus$ arrays result from $\oplus$ actions (restoration of edges on empty graph Λ by calculating the total credentials of the tree G and its copy on Λ. As an example we presented the $\oplus$ and $\ominus$ kernels in Fig.9 of this tree. Together with each edge we indicated the number of sub-graphs g that contain this edge and which are isomorphic to the graph shown in the Fig.8.

Now let us consider directed trees. If it is of interest to separate “clusters” in a directed tree, we shall proceed as follows. Let us consider the following small graphs: g_1, g_2 and g_3 (see Fig. 10).

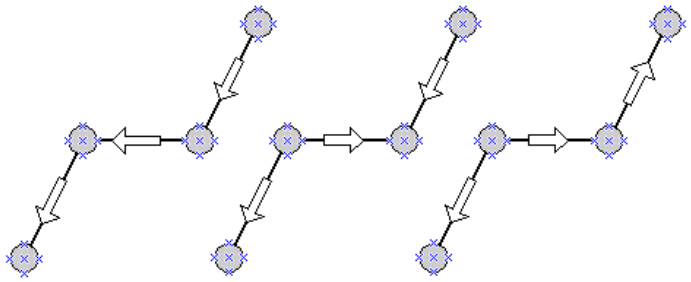


Figure 10

The credential function π on a directed tree can be calculated separately with respect to each small graph g_1, g_2 and g_3; then the values of all these three functions can be added up (a linear combination), thus yielding the overall function with respect to the graphs g_1, g_2 and g_3. In the same way we can assign a monotonic system on arcs of a tree if $\ominus$ action signifies the removal of an arc of a tree, $\oplus$ action the restoration of an arc on a copy of given tree on Λ. Thus we can pose on directed trees a similar problem of finding cluster kernels. Let us note that we use in the last example with trees a more general model of assignment of collections of credential functions with respect to a series of small graphs. The model in Section 2 has been presented for one graph g. A collection of credential arrays with respect to a series of graphs has also the property of monotonicity, and apparently such a model is more interesting in solving problems of determination of “saturated” parts of graphs.

Let us consider how the g, $\oplus$ and $\ominus$ activities of a small graph can be selected to coordinate the selection of these elements with the desired "saturation" of the graph under study. The desired saturation of a graph can be understood as desirable from the researcher's point of view, because the researcher has a certain working hypothesis about the structure of the graph.

For the small graph g for which the functions π were calculated, we choose a cyclic triangle. We use the method described in the previous subsections to create a set of credentials. The removal of all the arcs in the tournament $\langle x \text{ wins } y \rangle$ from the vertex x is the $\ominus$ action on the vertex x and the $\oplus$-action on the graph Λ is the restoration of all pairs where x wins y. The set of credentials on the graph tournament arcs must be created analogously to the previous sections.

The question of which is more preferable, whether the scheme is done on the arcs of the tournament (a game between two participants) or on the vertices of a graph, cannot be solved within the theory. It can only be said that if there are several central regions in the tournament that are saturated with cyclic triplets, the scheme of separating the kernel by arcs will be better, because these regions can be separated. A model that uses vertices separates the kernel that consists of these regions, but does not allow a single region to be found. We don't have a list of arcs that represent these areas.

Non-cyclic oriented graphs are a suitable tool for describing operating systems. The operating system can be thought of as a system of modules and interpreted as a library of programs. Each work program is a set of modules activated from a library, or in other words, in a non-cyclic graph of the path form. The modules call each other in sequence if they are not all in RAM or for some other reasons.

If the library is large, the natural idea is to place the modules on data carriers. Before solving this task, it is reasonable to explain the structure of the non-cyclic graph of the library of modules. The latter can be understood as the separation of the main sub-vertices or arrows. Vertices are very important places in the graph, in this case they are modules that are available in many work programs.

This task can be formally described in a graph by a monotonic system organization model. The question of the preference of monotonic systems formed by arrows or vertices again requires experimental control. Looking at the trees, we try to separate them from an area that is in some way more like a "bush" than the rest of the tree.

8. DISCUSSIONS AND SUMMARY

Usually, information is collected in to draw the necessary conclusions on issues related to human collectives, economic activity, production processes, demography, etc. If you are more interested in the verbal history itself, then the numerical experiments in Tables 1-3 can still be interesting of themselves. Indeed, with the help of these tables, the main feature of the analysis method is manifested, namely, the independence from any prior knowledge or specific information that is necessary for data analysis. This is especially true of the usual practice of personal and expensive interviews in sociological research. In this regard, the algorithm described in the manual for decomposing the data matrix into layers can be called "blind eye of statistical evaluation or scoring", which is what we need (Võhandu, 1979, 1989). This methodological guide looked at this information processing method that often has been used.

Although the main component of this methodological guide was prepared and presented for publication many years ago, as it seems to us everything that is given here is still relevant. It's not a secret that with the development of information technologies, methods for analyzing data extracted from our environment not only become more complicated, but also their volume has grown to enormous sizes when you have to deal with databases whose size reaches many gigabytes in the amount of collected information. One thing is that all the information in such well-known applications as Facebook and the like are always reflected in some graphs of mutual relations between the participants, whether it is Linkedin or Twitter, etc. Many do not even suspect that our technology for analyzing relationships reflected in these applications are fully adapted to the analysis of such information. The problem here is that such information must be collected and presented either in tabular form or in the form of graphs. Graphs, however, must again be presented in tabular form, which, as we have already indicated, is the main form of data to be analyzed.

The algorithm for decomposing data into layers given in this tutorial turned out to be effective in many specific problems as we can apply here in the form of data viewing technology. Moreover, as already indicated throughout the book, the entire analysis process begins with the construction of the so-called defining sequence, whether it be elements of graphs or data tables, when it is required to find a local maximum at which the global maximum is reached when moving along the defining sequence from weak elements in the direction of strong ones. It turns out that a more effective method of searching for the core or kernel of a monotonic system is to move from top to bottom, from strong to weak elements. Such a search for the kernel is much more economical than the one that was proposed at that time in the original of this methodological manual.

On the other hand, the model of a monotonic system turned out to be a more complex than the author had assumed, who initiated the theoretical and practical use of monotonic systems. The fact is that on graphs when arcs of a graph or edges are taken as elements of the system, it is required to formulate very precisely what are $\oplus$ and $\ominus$ actions. If the $\ominus$ action is to remove or $\oplus$ is add both arcs and edges of the graph together with arcs and edges adjacent to an arc or edge, then monotone systems of a special type arise when the layering algorithm does not always lead to an optimal layer in the global sense. This white area has not yet been sufficiently studied, and here it is quite possible to discover some new features of monotonic systems of the indicated unusual type. We have already indicated this feature earlier in the article on how to organize a party in order to make the optimal combination of participants.

LITERATURE

Harari, F. (1969). Graph Theory, Addison-Wesley, Reading; Mass.

Kempner Y., Mirkin B.G. and I.B. Muchnik. (1997), Monotone Linkage Clustering and Quasi-Convex Set Functions, Appl. Math. Letters, 10, No. 4, pp. 19-24,

Mullat, J.E. a) (1971). On certain maximum principle for certain set-valued functions, Tallinn Tech. Univ. Proc., Ser. A, 313, 37-44;
b) (1976). Extremal subsystems of monotonic systems," I, II and 1977. III. Avtomatika and Telemekhanika, I, 5, pp. 130-139; II, 8, pp. 169-177; III, 1, pp. 109-119;
c) (1978). Andmestruktuuri Avamismeetodid, metoodiline juhend, TPI, Informatsioonitöötlemise kateeder, TRÜ Arhhivikogu (in Estonian);
d). (1979). Stable Coalitions in Monotonic Games," Automation and Remote Control, 1469-1478, Avtomatika and Telemekhanika, 40, 84-94;
e) (1981). Counter Monotonic Systems in the Analysis of the Structure of Multivariate Distributions, Automation and Remote Control, 42, 986-993;
f) (1995). A Fast Algorithm for Finding Matching Responses in a Survey Data Table, Mathematical Social Sciences, 30, 195-205.

Võhandu, L.K. a) (1979). Express methods of data analysis, Trans. of Tallinn Technical Univ, No. 464. pp. 21-35 (in Russian);
b) (1989). Fast methods in exploratory data analysis, Trans. of Tallinn Technical University, No. 705, pp. 3-19.